Outer Hair Cells Amplify the Fluid Traveling Wave by Changing Organ-of-Corti area in the Short-Wave Region

John J. Guinan Jr[1,2, a)]

[1] *Eaton-Peabody Laboratories, Mass. Eye and Ear, 243 Charles St., Boston MA 02114, USA.*
[2] *Harvard Medical School, Dept. of Otolaryngology Head and Neck Surgery, Boston MA, USA.*
[a)] john_guinan@meei.harvard.edu

Abstract. There are two kinds of cochlear amplification: "non-propagating amplification" that is widespread but doesn't couple to the traveling wave, and "traveling-wave amplification" (TWA) that enables wide-band outer hair cell (OHC) motility to produce sharply tuned amplification. TWA has been thought to be produced by an "OHCs-push-on-the-basilar-membrane" mechanism. Here it is hypothesized that in the traveling-wave short-wave region, OHC motility causes longitudinal cortilymph flow in the organ-of-Corti (OoC) tunnels, producing OoC area changes that drive scala media fluid motion and amplify the traveling wave.

INTRODUCTION

Cochlear amplification has been thought to be produced by outer hair cell (OHC) somatic motility acting on the basilar membrane (BM) at the right phase to add energy cycle-by-cycle to the traveling wave. However, this "OHCs-push-on-the-BM" hypothesis has several obstacles: OHCs do not directly contact the BM, OHCs have no anchor for their push on the BM, and how OHC contractions are properly timed to add energy to the traveling wave is unknown. Adding to the conundrum, the motions of many organ-of-Corti (OoC) structures, e.g., the reticular lamina (RL), are amplified at frequencies that are far from frequencies amplified on the BM. These issues are addressed by dividing cochlear amplification into two types, and by a new method for traveling-wave amplification.

Traveling-wave properties and two types of amplification. Although the traveling wave is often talked about as if it were mainly on the BM, the energy of the traveling wave is carried by motion of the fluids in scala tympani, scala media and scala vestibuli (ST, SM and SV) [1]. BM motion is produced by the pressure difference across the cochlear partition (SV-SM). Traveling-wave pressure has been measured in the scalae [2-4] but fluid motion in the scalae has not been measured (although the motions of the structures abutting the scalae provide measures of the motions in the adjacent fluids). Starting at the cochlear base, the traveling wave propagates apically with a speed that decreases as it approaches CF; near the stapes the wavelength is long, but near a tone's best-frequency (BF) region the wavelength is much shorter. These are denoted the long-wave and short-wave regions [5].

The long- and short-wave regions demarcate the approximate locations of the two types of cochlear amplification: "traveling-wave amplification" and "non-propagating amplification" (Fig. 1). Traveling-wave amplification is manifest by increases in BM motion and by increases in traveling-wave pressure. A crucial aspect of traveling-wave amplification is that increases (and decreases from two-tone suppression (2TS)) are carried apically

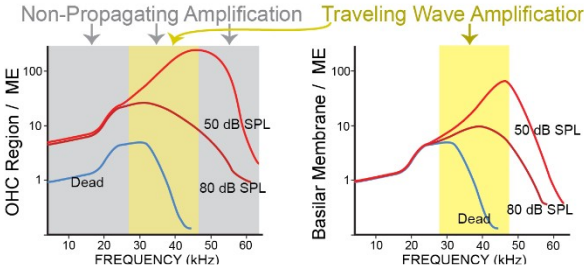


FIGURE 1. Frequency regions that show traveling-wave amplification (yellow) and non-propagating amplification (gray) on outer-hair-cell (OHC) region motion (left) and basilar-membrane (BM) motion (right). Lines are smoothed versions of mouse data from [6]. Both amplifications are present in the center of the left panel.

by the traveling wave so that amplification (or 2TS) produced in each OoC cross-section *accumulates* as the traveling wave moves forward (e.g., [7,8]). In contrast, non-propagating amplification is present everywhere that the traveling wave amplitude is enough to substantially deflect OHC stereocilia and cause OHC contractions/elongations, but it is local and not coupled to the traveling wave. This conclusion is supported by 2TS of non-propagating amplification having an effect only locally [9,10]. Also, solutions that block OHC motility and

remove non-propagating amplification far basal of a tone's BF have no effect on otoacoustic emissions, showing that the motion increases from non-propagating amplification don't couple to backward traveling waves [11]. In the Theory section, we will consider why OHC motility produces traveling-wave amplification in some regions and non-propagating amplification in others.

THEORY

Short-wave-region longitudinal cortilymph flow. Although OHCs contract/elongate wherever there is substantial traveling-wave motion, OHC length changes only produce fluid flow into and along the OoC tunnels in the short-wave region. There are no measurements of cortilymph flow in intact cochleae, but in excised preparations electrical stimulation of OHCs produced cortilymph flow along the tunnel of Corti [12]. In addition, model simulations indicate that cortilymph flow along the tunnels extends for more than the short-wave-region wavelength of the traveling wave [12-14]. In the short-wave region, contractions of the OHCs at one phase of the traveling wave squeeze the OoC and force some cortilymph out of the region being squeezed and into the OoC tunnels in adjacent sections. In the same cochlea, one-half wavelength along the cochlea, OHCs are expanding and sucking fluid out of the tunnels. The net effect of this is to produce cortilymph flow from contracting regions along the tunnels to expanding regions (Fig. 2C), which results in an oscillatory flow along the tunnels. For a dimension example, in the gerbil 20 kHz region near BF, half of the traveling-wave wavelength is ~100 μm [15] which is approximately the distance between centers of the two tunnels. In contrast, far basal of BF, in the long-wave region, there is relatively little cortilymph flow along the tunnels because the long distance means a large mass of fluid must be moved along the tunnels (Fig. 2A). The result is that OHC contractions can be expected to produce longitudinal cortilymph flow mainly in the short-wave region.

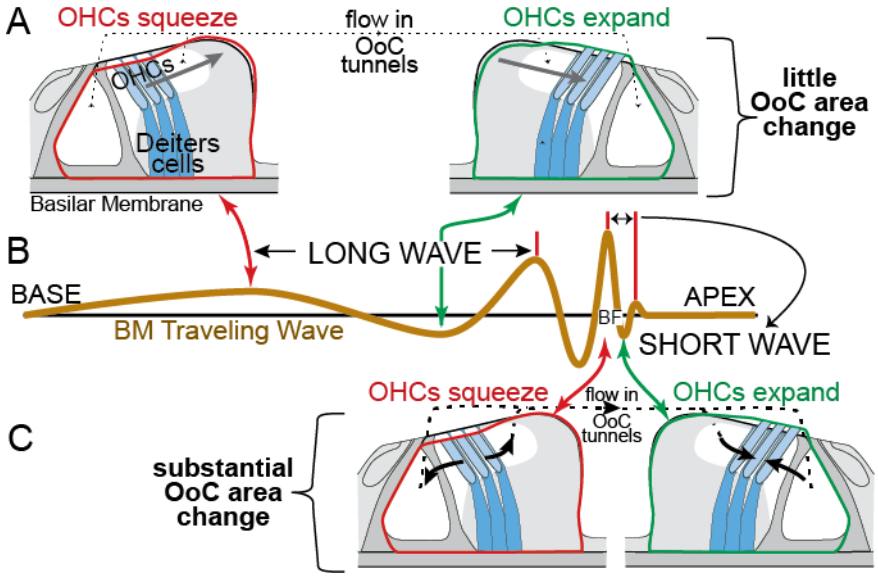


FIGURE 2. Along the traveling, wave contracting and expanding OHCs cause longitudinal flow of cortilymph in the short-wave region but very little in the long-wave region. **A**. Organ of Corti (OoC) transverse sections at a point of the traveling wave with OHC contractions (left) and expansions (right). **B**. A snapshot of BM deflections from a tone showing the best frequency (BF) and example long-wave and short-wave wavelengths. **C**. OoC transverse sections as in A but in the short-wave region where there is a short distance from OHCs squeezing to OHCs expanding and substantial flow along the tunnels between these OHC regions (shown by black arrows and dashed lines). Measures of OoC area are shown by the red outlines (left) and green outlines (right) which indicate little net area change in A and substantial area change in C. Gray arrows in A indicate fluid and tissue movement that keeps the OoC area little changed by OHC contractions/expansions in the short-wave region. The figure shows flows at a time instant but in reality the flows are oscillating at the tone frequency.

The role of Deiters cells. The Deiters cells (DC) connect OHCs to the BM, and the DC phalangeal process (PhP) connects the bottom of the OHC to the RL at a location slightly more lateral than the OHC and several OHCs more apical [16,17]. The Deiters-cell connection of the OHC and the BM is the basis of the OHCs-push-on-the-BM hypothesis (e.g., [18] Fig. 2). In feed-forward cochlear models, the PhP has been hypothesized to apply OHC forces at a location apical to the OHC (e.g., [19,20]). A problem for these hypotheses is that at frequencies far below CF,

the OHC-DC junction (OHC-DCj) has transverse motion that is many times larger than BM motion (e.g., [21]). This implies that the DC doesn't couple OHC motion effectively to the BM, i.e., that DCs are relatively elastic. An alternate hypothesis for the role of Deiters cells and their PhP is that they and the OHCs act like a bellows that enables OHC contractions/expansions to change the area of the fluid space enclosed and help to force cortilymph around the OHCs into the tunnels (Fig. 3).

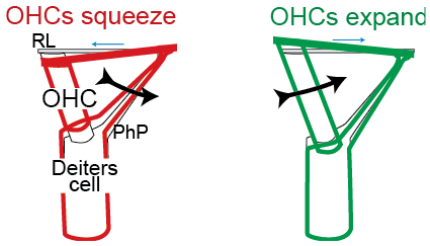


FIGURE 3. Cartoons of an OHC and attached Deiters cell contracting (left) or expanding (right) and forcing fluid in the space enclosed by the OHC, the Deiters-cell phalangeal process (PhP), and the reticular lamina (RL) out of, or into, the enclosed space. Thin, blue arrows at top show the direction the OHC squeezing/contracting and consequent tilting of the RL might have on scala media fluid flow adjacent to the RL.

Effects of OoC area changes. An important consequence of cortilymph flowing out of a cochlear cross section and into the OoC tunnels of adjacent OoC cross sections, is that this changes the cross-sectional area of the OoC: where OHCs contract the OoC area is smaller than at rest, and where OHCs expand, the OoC area is larger than at rest (Fig. 2C). The change in OoC area is primarily along the SM edge of the OoC, not the ST edge, because the RL is much less stiff than the BM and the RL moves more. This OoC area change creates pressures in, and/or movements of, the SM fluids. In the long-wave region the distance from regions with OHCs contracting to regions with OHCs elongating is very far making it difficult to force cortilymph along the tunnels, so the squeezing/expansion of the OHC region produces expansion/contraction locally in the outer tunnel and movement toward the Hensen cells, so that the net OoC area change is very little in the long-wave region (Fig. 2A). Opposite-phase movements of the RL and Hensen cell edges have been documented [9, 22, 23]

What are the consequences of OoC area changes? Could OoC area changes produce traveling-wave amplification? First, we consider the phase of OHC contractions that is ideal for producing cochlear amplification by the OHCs-push-on-the-BM method (Fig. 4A-E). This phase is not far from the phase measured by Lee et al. [24] and the phase considered by Dong and Olson [25]. At this phase of OHC contractions relative to BM motion, the OoC area change, acting on a compliant SM, would produce a pressure in SM that is in phase with the SM pressure (Fig. 4E). With this view, the area change (Fig. 4E) would increase the SM pressure of the traveling wave (Fig. 4C), which would increase the amplitude of the traveling wave, i.e., produce traveling-wave amplification.

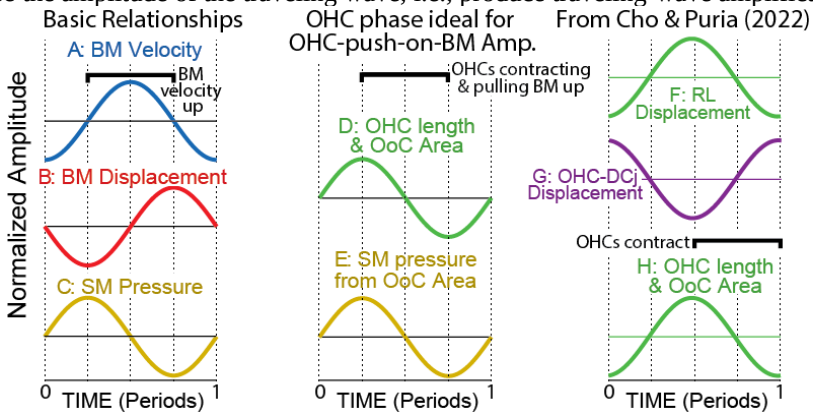


FIGURE 4. One cycle amplitude vs. time plots for cochlear responses to a tone. **Left column**: Traveling wave relationships. **Middle column**: Response phases for OHC length changes that were set to be ideal for amplifying by the OHCs-push-on-BM hypothesis, i.e., OHCs pull the BM up when the BM is moving upward – denoted by black brackets at the column tops. **Right column**: phases set relative to BM displacement at ½ octave below BF from measurements of Cho and Puria [21] for the RL over the third row of OHCs (F), and for the OHC-Deiters-cell junction (OHC-DCj) (G). Note that the OHC contraction time (bracket in H) is half while the BM is moving up and half while the BM is moving down, so the effects cancel.

As an alternate to setting the phase to be ideal for OHCs-push-on-the BM amplification (Fig. 4, middle), we used the measurements Cho and Puria [21] that showed OoC motion at ½ octave below BF to have a phase that was delayed by ~¼ period from the ideal (Fig. 4, right – delayed from Fig. 4 middle). With the phases of Figure 4, right, OHC contractions do not produce traveling-wave amplification by the OHCs-push-on-the-BM method! While OHCs are contracting BM velocity is first positive then negative (compare brackets in Fig. 4A and H) so the net

effect is zero. The phases shown in Figure 4, right, are perhaps an oversimplification because the measurements of Cho and Puria [21] are more complex and frequency dependent, but overall these measurements are not favorable for OHCs-push-on-BM cochlear amplification. However, the Cho & Puria data showed considerable traveling-wave amplification, which implies that the amplification came from the OoC-area-change applying forces to the SM fluid.

How does the OoC-area-change produce traveling-wave amplification? One way to think about the force exerted by the OoC-area-change on the SM fluid is in a cochlear cross section. With this view, the relationship between RL movement and the pressure or forces it exerts on the SM fluid depends on whether the OoC area change is acting on a RL compliance, resistance, or mass, and on the direction of the fluid motion adjacent to the RL. A good understanding of this requires measurements of motion along the SM perimeter of the OoC (which is within current capabilities) and accurate, 3D, models to get the fluid motion and discern when energy is added (or subtracted) from the traveling wave. However, even without such a detailed analysis, it is clear that the motion of the RL exerts forces on the adjacent fluid, and with correct phasing these can amplify the traveling wave.

The two different phase scenarios in Figure 4, both based on experimental data, indicate that how traveling-wave amplification is produced may be complex and not the same in all cases. The analysis in Figure 4 was based on considering traveling-wave amplification as solely due to forces exerted by the OoC-area change *within a given cochlear cross section*. Such forces are shown in Figure 5 by the red/green double tipped arrows. For this kind of motion to amplify the fluid traveling wave, the forces must be at the correct phase (or close to the correct phase). However, another way that the OoC area change may add energy to the traveling wave is by the *longitudinal* motion of the corrugations in the RL surface produced by the OoC-area-changes (Fig. 5, blue arrows pointing apically). This OoC-area-change-wave would push, and/or drag, the SM fluid apically as the traveling wave moves apically, and by the push/drag add energy to the traveling wave.

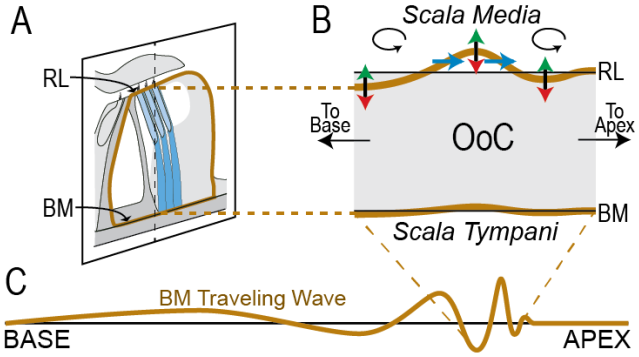


FIGURE 5. Organ-of-Corti (OoC) area changes acting on Scala Media (SM) fluid around the peak of the traveling wave. **A**. An oblique OoC cross-section showing the locations of the reticular-lamina (RL) and basilar membrane (BM) points diagramed in B. **B**. A *longitudinal* section through the OoC (at the dashed line in A) showing motions at one point in time in the RL abutting SM fluid and in the BM abutting scala tympani fluid. The arrows at top show motion directions for OoC expansion (green, up) and contraction (red, down) that act on SM fluid within a given OoC cross section, and blue horizontal arrows that indicate the direction of action of the forward travel of RL corrugations on SM fluid. The motions shown in B represent the net effect of RL motion on SM fluid in the short-wave region where there are changes in OoC area, but in the long-wave region where there are minimal OoC area changes, the motions of more lateral RL regions have opposite phases so the net effect on SM fluids is small. **C**. A cartoon traveling wave with dashed lines showing the region considered in B.

Human traveling-wave amplification. The anatomy and motion of the cochlear partition is different in human temporal bones than in smaller laboratory animals. The cochlear partition is much longer in humans than that in lab animals and has a non-body "bridge" between the bone of the spiral lamina and the medial edge of the BM [26]. In addition, in response to tones, the point of greatest motion in the human temporal bone is at the bridge-BM boundary, not near the center of the BM as in lab animals [26]. An important consequence of this is that when the BM moves up, in human temporal bones the OoC rotates clockwise whereas in lab animals it rotates counter-clockwise [26]. With many OHCs-push-on-BM theories, the OoC rotation in humans would lead to cochlear amplification at all frequencies, but a wide range of data indicate that the pattern of cochlear amplification is similar in humans and lab animals (e.g., [27,28]). This seeming anomaly would be removed if traveling-wave amplification came an OoC-area-change drive (as in Figs. 2, 5) and not from OHCs-push-on-the-BM.

Another way humans are somewhat different from lab animals is in frequency selectivity, which is much sharper in humans when measured in frequency bandwidth, but not when measured in the spread of tuning along the BM [29, 30]. However, in addition to the human cochlea being much longer than cochleae of lab animals, humans have

much longer otoacoustic emission delays [30]. These two things together make the wavelength of the traveling wave near corresponding BFs to be similar in humans and lab animals. If, as is proposed here, the wavelength controls traveling-wave amplification, then the similarity of the sharpness of tuning in humans and lab animals, when expressed in distance along the BM, may be due to their similar wavelengths near BF.

DISCUSSION

Active processes within the cochlea increase the motion response of the structures within the cochlea. This increase is called “cochlear amplification.” When the only cochlear structure that could be measured was the BM and cochlear amplification occurred only at BF and slightly basal, the meaning of “cochlear amplification.” was unambiguous. However, optical coherence tomography (OCT) allowed motion measurements of structures within the OoC and revealed more extensive cochlear amplification than was seen on the BM. This has produced confusion about the relationship of the various amplifications. Here we propose that two kinds of cochlear amplification should be distinguished because of their different properties and consequences. Two kinds of cochlear nonlinearities were distinguished by Strimbu et al. [31] that may correspond to the division made here, but identifying these as amplification, not just nonlinearity, seems important for understanding their origins. Our division has allowed us to form a series of hypotheses to explain the most consequential amplification, traveling-wave amplification.

The most important hypotheses put forth here are: (1) OHC motility causes longitudinal flow of cortilymph along the OoC tunnels in the short-wave region, but little cortilymph flow in the long-wave region, (2) the longitudinal cortilymph flow causes changes in OoC area, and (3) the changes in OoC area drive traveling wave cochlear amplification by applying forces to SM fluid. Exactly how the OoC-area-change adds energy to the traveling wave is not yet clear; there may be multiple ways (e.g., Fig. 5). The phase relationship between basilar-membrane motion and OHC contractions appears to be substantially different between the mouse 9 kHz region [24] and the gerbil 45 kHz region [21]. Although methodological differences cannot be ruled out (e.g., the measurements were not actually from the structures the data were assigned to, and/or the data were contaminated by transverse motion), these differences across species provide a reason for thinking that the action of the OoC-area-pump strictly in the transverse plane (which requires good matching of OoC-area push and SM pressure/motion) may not explain the data. Perhaps instead of an “OoC-area pump” we should think of an OoC-area *longitudinal wave* as being what adds energy to the traveling wave (Fig. 5, blue arrows). It appears to have the advantage of not needing an exact match between the phase of the OoC-area change and the local SM pressure/motion. In favor of OoC longitudinal motion having a role in traveling-wave amplification is that cochlear longitudinal motion may be larger than the motions in other directions [23,32].

The hypotheses put forth here need testing. Some possible tests have been suggested (Guinan, 2022). The longitudinal drive from an OoC-area-change RL wave may be difficult to experimentally measure. However, confirming that there is an OoC-area change at a fixed time point in the traveling wave (e.g., where OHCs contract most) would be enough to show that this bulge/crinkle is carried apically by the traveling wave. The degree to which a forward (apically) moving RL wave adds energy to the traveling wave must be approached by models, as with all hypotheses for traveling-wave amplification,.

ACKNOWLEDGEMENTS

Supported in part by NIH NIDCD R01 DC007910 (to Sunil Puria).

REFERENCES

1. M. J. Lighthill, Energy flow in the cochlea. *J Fluid Mech* **106**, 149-213 (1981).
2. E. S. Olson, Direct measurement of intra-cochlear pressure waves. *Nature* **402**, 526-529 (1999).
3. W. Dong, E. S. Olson, Detection of cochlear amplification and its activation. *Biophys J* **105**, 1067-1078 (2013).
4. S. S. Kale, and E. S. Olson, Intracochlear Scala Media Pressure Measurement: Implications for Models of Cochlear Mechanics. *Biophysical journal* **109**, 2678-2688 (2015).
5. R. Patuzzi, "Cochlear Micomechanics and Macromechanics" in The Cochlea, P. J. Dallos, A. N. Popper, R. R. Fay, Eds. (Springer-Verlag, New York, 1996), pp. 186-257.`
6. T. Ren, W. He, and D. Kemp, Reticular lamina and basilar membrane vibrations in living mouse cochleae. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 9910-9915 (2016).

7. N. P. Cooper, Two-tone suppression in cochlear mechanics. *J Acoust Soc Am* **99**, 3087-3098 (1996).
8. J. A. Fisher, F. Nin, T. Reichenbach, R. C. Uthaiah, and A. J. Hudspeth, The spatial pattern of cochlear amplification. *Neuron* **76**, 989-997 (2012).
9. J. B. Dewey, B. E. Applegate, and J. S. Oghalai, Amplification and Suppression of Traveling Waves along the Mouse Organ of Corti: Evidence for Spatial Variation in the Longitudinal Coupling of Outer Hair Cell-Generated Forces. *The Journal of neuroscience* **39**, 1805-1816 (2019).
10. E. Fallah, C. E. Strimbu, and E. S. Olson, Nonlinearity and amplification in cochlear responses to single and multi-tone stimuli. *Hearing research* **377**, 271-281 (2019).
11. S. S. Goodman, C. Lee, J. J. Guinan, Jr., and J. T. Lichtenhan, The Spatial Origins of Cochlear Amplification Assessed by Stimulus-Frequency Otoacoustic Emissions. *Biophys J* **118**, 1183-1195 (2020).
12. K. D. Karavitaki, and D. C. Mountain, Evidence for outer hair cell driven oscillatory fluid flow in the tunnel of corti. *Biophys J* **92**, 3284-3293 (2007).
13. B. F. Zagadou, and D. C. Mountain, Analysis of the Cochlear Amplifier Fluid Pump Hypothesis. *J Assoc Res Otolaryngol* 10.1007/s10162-011-0308-x (2012).
14. B. F. Zagadou, P. E. Barbone, and D. C. Mountain, Significance of the microfludic flow inside the organ of Corti. *J Biomech Eng* **142** (2020).
15. M. Van der Heijden, and N. P. Cooper, "Wave propagation in the mammalian cochlea" in To the ear and back again: Advances in auditory biophysics. Proc. 13th Mechanics of Hearing Workshop, C. Bergevin, S. Puria, Eds. (AIP Conference Proceedings, Melvile, NY, 2018), vol. 1965, pp. 080002.
16. N. B. Slepecky, "Structure of the Mammalian Cochlea" in The Cochlea, P. J. Dallos, A. N. Popper, R. R. Fay, Eds. (Springer-Verlag, New York, 1996), pp. 44-129.
17. J. A. Soons, A. J. Ricci, C. R. Steele, and S. Puria, Cytoarchitecture of the mouse organ of corti from base to apex, determined using in situ two-photon imaging. *J Assoc Res Otolaryngol* **16**, 47-66 (2015).
18. J. J. Guinan, Jr., A. Salt, and M. A. Cheatham, Progress in cochlear physiology after Bekesy. *Hear Res* **293**, 12-20 (2012).
19. Y. J. Yoon, C. R. Steele, and S. Puria, Feed-forward and feed-backward amplification model from cochlear cytoarchitecture: an interspecies comparison. *Biophys J* **100**, 1-10 (2011).
20. H. Motallebzadeh, J. A. M. Soons, and S. Puria, Cochlear amplification and tuning depend on the cellular arrangement within the organ of Corti. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 5762-5767 (2018).
21. N.-H. Cho, and S. Puria, Motion of the cochlear reticular lamina varies radially across the outer-hair-cell rows. *Biorvix* submitted March 1, 2022 (2022).
22. K. D. Karavitaki, and D. C. Mountain, Imaging electrically evoked micromechanical motion within the organ of Corti of the excised gerbil cochlea. *Biophys J* **92**, 3294-3316 (2007).
23. N. P. Cooper, A. Vavakou, and M. van der Heijden, Vibration hotspots reveal longitudinal funneling of sound-evoked motion in the mammalian cochlea. *Nature communications* **9**, 3054 (2018).
24. H. Y. Lee *et al.*, Two-Dimensional Cochlear Micromechanics Measured In Vivo Demonstrate Radial Tuning within the Mouse Organ of Corti. *J Neurosci* **36**, 8160-8173 (2016).
25. W. Dong, E. S. Olson, Detection of cochlear amplification and its activation. *Biophys J* **105**, 1067-1078 (2013).
26. S. Raufer, J. J. Guinan, Jr., and H. H. Nakajima, Cochlear partition anatomy and motion in humans differ from the classic view of mammals. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 13977-13982 (2019).
27. A. Oxenham, How We Hear: Perception and Neural Coding of Sound. *Annu Rev Psychol* **69**, 27-50 (2018).
28. C. J. Sumner *et al.*, Mammalian behavior and physiology converge to confirm sharper cochlear tuning in humans. *Proceedings of the National Academy of Sciences U.S.A.* **115**, 11322-11326 (2018).
29. C. A. Shera, J. J. Guinan, Jr., A. J. Oxenham, Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. *Proceedings of the National Academy of Sciences U.S.A.* **99**, 3318-3323 (2002).
30. C. A. Shera, J. J. Guinan, Jr., and A. J. Oxenham, Otoacoustic estimation of cochlear tuning: validation in the chinchilla. *J Assoc Res Otolaryngol* **11**, 343-365 (2010).
31. C. E. Strimbu, Y. Wang, and E. S. Olson, Manipulation of the Endocochlear Potential Reveals Two Distinct Types of Cochlear Nonlinearity. *Biophys J* **119**, 2087-2101 (2020).
32. S. W. Meenderink, and W. Dong, Organ of Corti vibrations are dominated by longitudinal motion in vivo. *Research Square* doi.org/10.21203/rs.3.rs-1405408/v1, 1-12 (2022).

33. J. J. Guinan, Jr., Cochlear Amplification in the Short-Wave Region by Outer Hair Cells changing Organ-of-Corti area to Amplify the Fluid Traveling Wave. *Hear Res* submitted (2022).

The Shape of Noise to Come: Signal vs. Noise Amplification in the Active Cochlea

Alessandro Altoè[a)] and Christopher A. Shera[b)]

Caruso Department of Otolaryngology, University of Southern California, Los Angeles, CA

[a)]*Corresponding author: altoe@usc.edu*

[b)]*Also at Department of Physics and Astronomy, University of Southern California, Los Angeles, CA*

Abstract. According to the dominant view, the mammalian cochlea spatially amplifies signals by actively pumping energy into the traveling wave. That is, signals are amplified as they propagate through a region where the medium's resistance is effectively negative. While signal amplification has been extensively studied in active cochlear models, the same cannot be said for amplification of internal noise. According to transmission-line theory, signals are amplified more than internal noise in regions where the net resistance is negative. Here we generalize this finding by showing that a distributed system composed of cascaded "noisy" amplifiers boosts signals more rapidly than the internal noise; the larger the amplifier gain the larger the signal-to-noise ratio (SNR) of the amplified signal. We further show that this mechanism is present in existing active cochlear models: the cochlear amplifier increases the SNR of cochlear responses, and thus enhances cochlear sensitivity. When considering also that the cochlear amplifier narrows the bandwidth of the "cochlear filters", activation of the cochlear amplifiers dramatically increases the SNR (by about one order of magnitude in our simulations) from the tail to the peak of the traveling wave. We further demonstrate that the tapered ear-horn-like cochlear geometry significantly improves the SNR of basilar-membrane responses.

INTRODUCTION

Long before cochlear models included an active amplifier, K.K.N. Chang [1] analyzed the noise factor—a measure of degradation of signal-to-noise ratio (SNR)—in transmission lines including both positive and negative resistances. He concluded that:

> "For a transmission line with both positive and negative distributed resistances as well as distributed random noise sources, low-noise amplification can be achieved via the line provided that (1) the line is operated at sufficiently high gain, (2) the line is well matched at both the input and the output, and (3) the net negative resistance and/or the net negative conductance is large; i.e., that gain per unit length is large. These conditions, which have been deduced from a quantitative analysis, are the necessary ones for low noise amplification."

In particular, the analysis of Cheng reveals that active transmission lines with net negative damping (and small scattering) spatially amplify the signal more rapidly than they amplify internal incoherent noise sources. Strikingly, theoretical work points out that the mammalian cochlea fulfills the three conditions for low-noise amplification identified by Chang [2, 3, 4, 5, 6]. That is, existing active cochlear models are expected to spatially amplify signals more than the intracochlear noise.

Additionally, a common noise reduction technique consists of band-pass filtering signals to remove noise outside the frequency range of interest. The cochlear amplifier narrows the bandwidth of the so-called "cochlear filters", and therefore the cochlear frequency-selective amplification is expected to boost the signal-to-noise ratio (SNR) of cochlear responses, thereby improving cochlear sensitivity. Furthermore, we previously proposed that the earhorn-shaped geometry of the cochlea helps maintain the signal intensity as this propagates along the cochlea. Based on this observation, we conjectured that the tapered geometry produces a significant improvement of SNR relative to the fictitious "rectangular" cochlea often depicted in models [7].

In a nutshell, physical intuition suggests that the cochlea is well-equipped with active (cochlear amplification) and passive (geometrical) mechanisms to minimize signal degradation by internal noise. Here, we test this conjecture using active cochlear models. As a first step, we analyze how distributed ("cascaded") amplification improves the SNR of the amplified signal.

DISTRIBUTED AMPLIFICATION BOOSTS SIGNAL MORE THAN INTERNAL NOISE

Consider the generic model of a spatially distributed signal amplifier presented in Fig. 1. This model consists of cascaded amplifiers, each one introducing noise in the signal path. In particular, the operation of each amplifier is that of multiplying the incoming signal by a factor g while adding noise to the amplified signal. We included an equivalent noise source at the input of the first amplifier in the model, which simplifies the mathematical treatment. All noise sources are incoherent, and, for simplicity, of equal RMS amplitude γ. Note that in this simple model, the first amplifier can be regarded as the n-th amplifier embedded in a larger system, where the signal and noise at its input represent the amplified signal and noise due to preceding amplifiers, respectively.

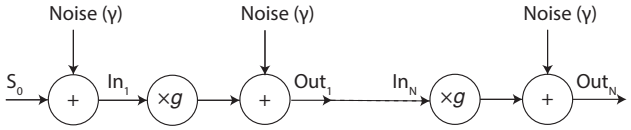


FIGURE 1. Generic model of distributed, noisy amplification. Each amplifiers in the model consists of a multiplier that boosts the signal at its input by a factor g, while adding noise to the resulting amplified signal.

We start by considering the output of the first section of the transmission line (Out_1 in Fig. 1). Results that apply at the output of the first section apply to the n-th section— the rigorous mathematical demonstration is somewhat cumbersome, so it is not detailed here. Consider the case when $g = 1$, so that the amplifier is "switched off" and the system of Fig. 1 represents lossless propagation of a signal in a noisy medium. In this case, the amplitude of the signal at the output of the first port is identical to the input signal,

$$O_{1,\text{signal}} = S_0, \tag{1}$$

where $O_{1,\text{signal}}$ and S_0 indicate the RMS amplitudes of the signal at the output of the first stage (Out_1 in Fig. 1) and at the system's input, respectively. In this case, the RMS noise amplitude at the output of the first stage is

$$O_{1,\text{noise}} = \sqrt{2}\gamma. \tag{2}$$

We now turn on the amplifier, so that $g > 1$. At the first amplifier output the signal is amplified by a factor g, while the noise output is the (incoherent) sum of the amplified noise present at the input with the noise at the output:

$$O_{1,\text{signal}} = gS_0, \tag{3}$$

$$O_{1,\text{noise}} = \sqrt{g^2+1}\gamma. \tag{4}$$

That is, turning on the amplifier boosts the signal by a factor $G_{\text{signal}} = g$ and the internal noise by a factor $G_{\text{noise}} = \sqrt{g^2+1}/\sqrt{2}$ (at the output of the first amplifier). Because $G_{\text{noise}} < G_{\text{signal}}$ for $g > 1$, the simple but generic model reveals that the first amplifier boosted the signal more than the internal noise. The larger the gain, the more the signal is amplified relative to the internal noise. Thus, turning on the amplifier (and cranking up the gain) improved the SNR at the first stage. Applying this reasoning inductively suggests that cascaded amplification boosts the signal-to-noise ratio as the signal propagates along the transmission line. As a corollary, when the amplifier in Fig. 1 is replaced by a "joy-killer", so that $g < 1$, the mathematical model reveals that $G_{\text{noise}} > G_{\text{signal}}$, suggesting that the signal is attenuated more rapidly than the internal noise and that the SNR progressively degrades along the transmission line.

Now consider signal and noise at the output of the n-th amplifier in the chain. We have that

$$O_{n,\text{signal}} = g^n S_0, \tag{5}$$

$$O_{n,\text{noise}} = \sqrt{\sum_0^n g^{2n}}\gamma = \begin{cases} \sqrt{\frac{g^{2(n+1)}-1}{g^2-1}}\gamma & \text{for } g \neq 1, \\ \sqrt{n+1}\gamma & \text{for } g = 1. \end{cases} \tag{6}$$

Changing the gain of the amplifiers from $g = 1$ to $g \neq 1$ modifies the signal amplitude by a factor $G_{\text{signal}} = g^n$, and the noise by a factor $G_{\text{noise}} = \sqrt{\frac{g^{2(n+1)}-1}{(n+1)g^2-1}}$. For $g > 1$ (i) the signal is amplified more than the internal noise

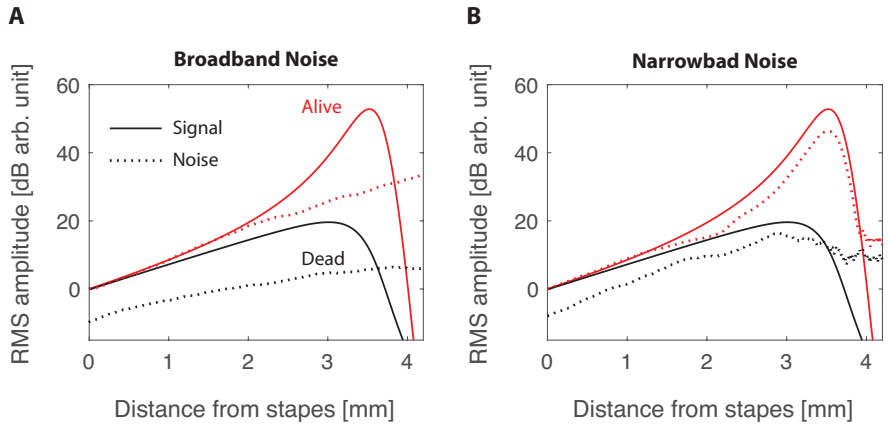


FIGURE 2. Comparison of BM responses to a 10 kHz sound stimulus and to internal broadband (A) and narrowband (B) noise. The simulations where performed in a 2-D finite-difference box model of the mouse [9], where noise is added by including incoherent pressure sources on the BM. The black and red line depict models of the active alive cochlea—where the cochlear amplifier is turned on—and of the passive "dead" cochlea—where the amplifier is turned off. The amplitude of the noise sources was selected so that at the cochlear entrance the signal-to-noise ratio is 0 dB in the active model, in order to better visualize the differential effects of amplification on noise and signal. The RMS amplitude of noise was obtained as the average RMS amplitude for 20 differently randomized simulations, in order to obtain smooth curves. The narrowband noise in panel B was simulated by including pressure sources of the same frequency as the stimulus, with gaussian distributed amplitude and phase .

($G_{\text{signal}} > G_{\text{noise}}$) and (ii) the ratio between signal and noise amplification increases with n and g—i.e., amplification progressively boosts the SNR as the signal propagates along the transmission line of an amount that depends on the amplifiers gain. Conversely, when $g < 1$ (i.e., the amplifiers become attenuators) the SNR progressively degrades along the line.

FREQUENCY-SELECTIVE AMPLIFICATION—A COMMON STRATEGY TO IMPROVE SENSITIVITY

A common practice to maximize the sensitivity of a detector is to selectively amplify the received signal in the frequency range where the signal to be detected is expected to occur. Indeed, according to signal-processing theory, an optimal detector embeds filters that maximize the SNR of the received signal (i.e., optimal filtering [8]). While optimal filters don't strictly need to amplify the signal, telecommunication systems universally employ active filters not only because they outperform passive filters in terms of frequency-selectivity, but also because amplification of the received signal can greatly improve the SNR by rendering the signal amplitude much larger than the internal noise of the receiver. Indeed, a common noise reduction technique consists of pre-amplifying the signal in frequency bands where the receiver (or medium over which the signal is transmitted) is especially noisy (pre-emphasis), to then de-amplifying it (de-emphasis).

THE COCHLEAR AMPLIFIER BOOSTS THE SNR

In the previous sections we showed that distributed and frequency-selective amplification are two mechanisms to improve the SNR and hence the sensitivity of a receiver. The cochlear amplifier, relying on the spatially distributed piezoelectric action of the OHCs that boost the traveling wave in a restricted frequency range near the characteristic frequency (CF), appears to implement both of these strategies simultaneously. To test whether the cochlear amplifier boosts signals more than internal noise, we performed simulations in a model including distributed noise sources with the amplifier turned "off" and "on".

Figure 2A compares the BM responses to a 10 kHz "airborne" tone and to broadband internal noise sources, for an

"alive" (amplifier on) and a "dead" (amplifier off) cochlear model. Although amplification in this model boosts the noise on the BM at all frequencies—a consequence of the increase of BM admittance magnitude following activation of the amplifier—the combination of cascaded and frequency-selective amplification produces greater amplification of signal than noise towards the peak region of the traveling wave. In particular, as waves propagate from the cochlear entrance to the CF location, the signal is spatially amplified by 41 dB while the noise increases by only 23 dB—i.e., signal amplification is an order of magnitude larger than noise amplification.

In the peak-region of the traveling wave the SNR increases (Fig. 2A) partially because the amplifier enhances the frequency-selectivity of the cochlea—thereby reducing noise via active filtering. In order to highlight the contribution of spatial amplification (as opposed to frequency-selective amplification) to the SNR, Fig. 2B compares BM responses to sound and to internal incoherent sources at the same frequency as the sound signal—these incoherent sources represent the internal noise component at the signal frequency. According to the theory of cascaded amplification detailed above, the signal is spatially amplified more than the internal noise. Although in these simulations the noise sources are large enough that signal and noise have the same amplitude near the cochlear entrance, the SNR at the peak of the traveling wave increases by $\sim$6 dB. In the passive case, where the amplifier is switched off, the signal is attenuated more rapidly than the noise, and the SNR from the tail to the peak (now shifted basally) of the traveling wave decreases by about 5 dB. Switching on the amplifier boosts the BM response (relative to post-mortem) at the CF location by 41 and 34 dB for the signal and narrowband noise, respectively.

COCHLEAR TAPERING BOOSTS THE SNR

We noted that in the box model of the cochlea the differential effect of noise and signal amplification varies systematically with frequency (data not shown). In particular, we found that the SNR near CF decreases at lower frequencies. The "propagating signal" that is spatially amplified is the pressure difference across the cochlear partition. In box models, the transpartition pressure inherently decreases with distance [7], with the consequence that the pressure wave entering the amplification region is somewhat attenuated (Fig.3A). The tapering of the cochlear duct counteracts this effect (Fig. 3B) and the pressure in the peak region of a realistically tapered model (geometry described in [11]) is indeed larger than in an "equivalent" box model (see [7] and Figs. 3A,B).

According to the theory of distributed amplification described above, the different amounts of noise and signal amplification depend critically on how much the medium modifies the signal amplitude as it propagates. Additionally, intuition suggests that "if the signal amplitude drops, the signal-to-noise ratio must drop too". Therefore, cochlear geometry is expected to play a leading-order role for signal vs. noise amplification. Figures 3C,D compare signal and narrowband noise amplification (as in Fig. 2B) in the box and tapered models. According to the intuitive reasoning, the SNR is larger in the tapered than in the box model, both in the passive and the active case. In particular, the signal is spatially amplified by $\sim$ 12 dB more than the internal noise in the tapered model, while only by $\sim$ 6 dB in the box model. At the CF location, turning on the amplifier boosts the signal (relative to post-mortem) by 15 and 7 dB more than the noise in the tapered and box models, respectively.

DISCUSSION AND CONCLUSIONS

Our analysis reveals that the cochlea is well-equipped with mechanisms to boost its sensitivity. These mechanisms include (i) spatially-distributed amplification, (ii) frequency-selective amplification, and (iii) a cochlear geometry that minimizes drops of signal level. These three mechanisms are analogous to those universally used in telecommunication system: (i) Repeaters that re-amplify signals en route—although they inevitably introduce some amount of noise, they greatly boost the SNR at distant receivers (thereby allowing us in Los Angeles to have a good chat over the phone with a friend in New York); (ii) Frequency-selective amplification—the mechanism that allows us to listen to a radio station although the background electromagnetic noise (e.g., from other radio stations) is orders of magnitude larger than the signal near the antenna; and (iii) avoiding unnecessary drops of signal level in the first place.

Interestingly, our model points out that even when signal and internal noise are of the same amplitude at the cochlear entrance, spatially distributed amplification (Fig. 2B) will produce a positive SNR in the peak region of the traveling wave, theoretically enabling detection of sounds near or just below background noise levels . When accounting for frequency-selective amplification and the broadband nature of noise (Fig. 2A), the model suggests that the cochlea can theoretically detect signals that are roughly one order of magnitude smaller than the noise at the cochlear entrance. Our results, taken together with transmission-line theory [1] and existing technology, contradict the pessimistic and

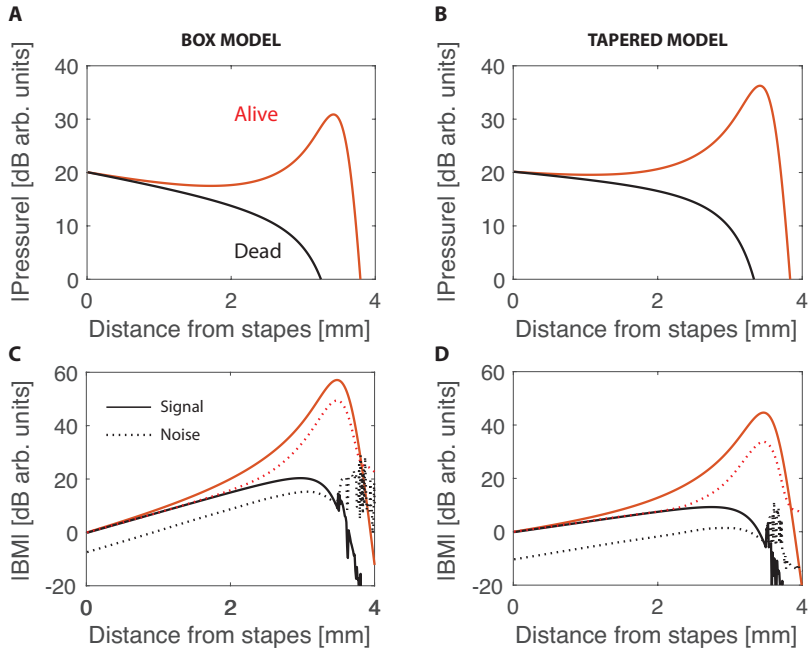


FIGURE 3. Comparison of box and tapered cochlear models. The simulations shown here are performed using semi-analytical models—described in [7]—of a mouse cochlea, with same parameters as the finite-difference model used in Fig. 2, and a signal frequency 9 kHz. (A,B) Pressure response to a 9-kHz tone in a box (A) and tapered (B) cochlear model. (C,D) Signal and narrowband noise amplification in a box (C) and tapered (D) model. Note that the prediction of the box-model used here are practically identical to those obtained using the finite-difference model (Fig. 2), except for the presence of numerical noise apically than the CF location—where analytical approximations break down [10].

folkloristic tale—recently revived by [12]—that cascaded amplification cannot improve cochlear sensitivity because it inevitably amplifies the internal noise and thus degrades the SNR. Indeed, cascaded amplifiers boost signals more than they boost internal noise, improving the SNR at every link of the chain. Our analysis of distributed systems further suggests that replacing the cascaded amplifier with a chain of attenuators (brakes) would serve primarily to increase the susceptibility to internal noise and hence minimize cochlear sensitivity.

Finally, much about cochlear noise remains unknown—both empirically and theoretically. For example, it is not clear what sources dominate the intracochlear noise. Our simple modeling certainly cannot answer this question, and further work is necessary to establish whether and how different noise sources (e.g., those on the partition or in the fluids far away from it) would produce different noise responses on the BM. Our analysis of the distributed amplifier points out that the noise at the n-th link of the chain will (likely) be dominated by the amplified noise at the "entrance" of the amplification region; this observation might be exploited to deduce the origin of intracochlear noise from the experimental data [13].

ACKNOWLEDGMENTS

This work was supported by NIH/NIDCD R21 DC019712 (AA) and R01 DC003687 (CAS).

REFERENCES

1. K. Chang, "Theory of a negative-resistance transmission line amplifier with distributed noise generators," J Applied Phys **31**, 871–875 (1960).
2. S. T. Neely and D. Kim, "A model for active elements in cochlear biomechanics," J. Acoust. Soc. Am. **79**, 1472–1480 (1986).
3. G. Zweig, "Finding the impedance of the organ of Corti." J. Acoust. Soc. Am. **89**, 1229–1254 (1991).
4. C. A. Shera and G. Zweig, "A symmetry suppresses the cochlear catastrophe," J. Acoust. Soc. Am. **89**, 1276–1289 (1991).

5. S. Puria and J. B. Allen, "A parametric study of cochlear input impedance," J. Acoust. Soc. Am **89**, 287–309 (1991).
6. E. de Boer and A. L. Nuttall, "The mechanical waveform of the basilar membrane. III. Intensity effects," J. Acoust. Soc. Am. **107**, 1497–1507 (2000).
7. A. Altoè and C. A. Shera, "The cochlear ear horn: Geometric origin of tonotopic variations in the sharpness of tuning," Sci. Rep. **10**, 20528 (2020).
8. N. Wiener, *Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications*, Vol. 113 (MIT press Cambridge, MA, 1949).
9. A. Altoè and C. A. Shera, "Nonlinear cochlear mechanics without direct vibration-amplification feedback," Phys. Rev. Res. **2**, 013218 (2020).
10. C. A. Shera, A. Tubis, and C. L. Talmadge, "Coherent reflection in a two-dimensional cochlea: Short-wave versus long-wave scattering in the generation of reflection-source otoacoustic emissions," J. Acoust. Soc. Am. **118**, 287–313 (2005).
11. A. Altoè, K. K. Charaziak, J. B. Dewey, A. Moleti, R. Sisto, J. S. Oghalai, and C. A. Shera, "The elusive cochlear filter: Wave origin of cochlear cross-frequency masking," J. Assoc. Res. Otolaryngol. **22**, 623–640 (2021).
12. M. van der Heijden and A. Vavakou, "Rectifying and sluggish: outer hair cells as regulators rather than amplifiers," Hear Res , 108367 (2021).
13. P. M. Quiñones, S. W. Meenderink, B. E. Applegate, and J. S. Oghalai, "Unloading outer hair cell bundles in vivo does not yield evidence of spontaneous oscillations in the mouse cochlea," Hear. Res. , 108473 (2022).

The dramatically dispersive character of the traveling wave

Marcel van der Heijden[1, a)], Anna Vavakou[1], Kostas Kondylidis[1] and Nigel P. Cooper[1]

[1]*Erasmus MC, Rotterdam, The Netherlands*
[a)] *marcel[dot]vdh[dot]work[at]gmail[dot]com*

Abstract. Using optical cocherence tomography (OCT) vibrometry we studied wave propagation by recording at multiple longitudinal postions in the gerbil and guinea pig cochlea. Our aims are to (1) fully portray the unusual dispersive character of the wave; (2) derive from these properties the underlying physics of the wave and the way it is regulated by physiological processes; (3) challenge cochlear modelers to report and analyze wave propagation in their models. In the current study we report two unexpected features of the spatial profiles of sound-evoked vibrations of the basilar membrane. The profiles of low-frequency components show very little systematic phase accumulation, whereas the slow waves near the best frequency often exhibit abrupt, place-fixed transitions in propagation velocity.

INTRODUCTION

The advent of OCT vibrometry has shaken up the field of cochlear mechanics. It does not require the use of reflectors and allows for vibration measurements across many different intracochlear positions. In contrast, the bulk of pre-OCT data consisted of single-point basilar membrane (BM) recordings. Consequently, decades of modeling efforts have concentrated on reproducing single-point BM data. Likewise, the conceptual framework of most models has been "BM centered", e.g. seeking to determine "BM impedance", or describing how "outer hair cells move the BM", etc. The complex micromechanics unveiled by OCT vibrometry expose the severe shortcomings of both the predictive power of existing cochlear models and the underlying conceptual framework. For instance, vibrations many octaves below the characteristic frequency display physiologically vulnerable nonlinearities and complex motion patterns ("hotspots"), contradicting the basic tenets of all BM-centered models. Amidst the (healthy) confusion, it is reassuring to find an unchallenged tenet of cochlear mechanics: the traveling wave.

Apart from the ability to look beyond the BM, OCT vibrometry has also facilitated the study of traveling waves in the cochlea, because the choice of recording locations is no longer restricted by the need to place reflective microbeads on the structures of interest.

In the current study we exploit this freedom of choice to obtain accurate, densely spaced longitudinal profiles of the traveling wave in the base of the gerbil cochlea. In our opinion, it is important to have such an accurate, quantitative characterization of cochlear waves. Despite many decades of theoretical analyses, the physics of these waves is still obscure. From previous experimental work [e.g. 1,2] we already know that the cochlear traveling wave is highly dispersive, i.e., that at a given location the propagation properties are dependent on frequency. A systematic quantitative description of wave propagation and its dependence on frequency, however, is still lacking.

METHODS

Sound-evoked vibrations were recorded from the basilar membrane of deeply anesthetized gerbils. The bulla was opened and the cochlear partition was viewed through the intact round window membrane. Images (B scans) and vibration measurements (M scans) were obtained using a Thorlabs Telesto III Optical Coherence Tomography (OCT) system at a sample rate of 111.6 kHz. The optical wavelength was 1300 nm and the effective lateral and axial resolution were ~10 μm and ~6 μm, respectively. For detailed methods, see [3]. The acoustic stimuli were irregularly spaced tone complexes (zwuis, [4]). Vibration responses were analyzed in the Fourier domain by extracting the stimulus-related spectral components and subjecting them to to a Rayleigh test [5]. Only components

with (p<0.001) were admitted to the analysis. Recordings were made from a densely spaced array of points spaced longitudinally. The positioning of the recording beam was done on the basis of a large (>20) set of radially oriented B-scan images covering the longitudinal stretch of interest, thus making sure that the exact same anatomical structures were targeted across longitudinal positions. This is important: not only are there are huge variations in vibration magnitude and phase between adjacent structures (e.g., BM versus Deiters' cells, [3]), there are also small but systematic differences between adjacent points of the same structure (e.g., radially spaced points of the BM, [6]).

RESULTS

Figure 1 shows a collection of normalized frequency response curves obtained at an array of 13 BM locations spanning an 860-µm longitudinal stretch. The recording beam was aimed at the junction of the pectinate and arcuate zone.

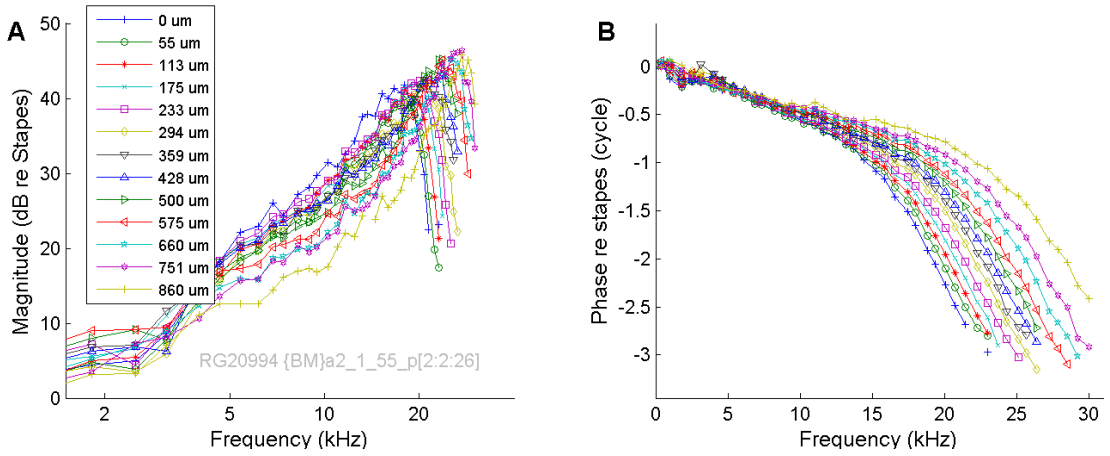


FIGURE 1. Collection frequency response curves recorded at 13 points on the BM spanning a longitudinal distance of 860 µm. The stimulus was a zwuis tone complex presnted at 55 dB SPL per component. **A**. Magnitude normalized to stapes motion. **B**. Phase re stapes. Best frequencies ranged from 16-28 kHz.

In Fig, 2, the same data are displayed a function of distance along the BM, a representation that underscores the traveling wave character. Each curve represents a spatial profile for a single frequency of the tone complex, and the display is split in two parts, covering the low (<10 kHz): high frequencies, respectively. Note that 10 kHz is more than 0.5 octave below the 16-kHz best frequency of the most apical location, so all of the low-frequency data in Fig. 2 belong to the "low tail" of the frequency response curves (Fig. 1). In particular, the low-frequency data in Fig. 2 represent linear, "passive" responses of the BM.

The low-frequency magnitude curves (Fig. 2A) show a systematic growth from base to apex. The average spatial rate of growth for the low-frequency responses was 6±4 dB/mm (N=14). Within the group of low-frequency responses, the frequencies above 5 kHz show a steeper growth, having an average spatial rate of growth of 9.6±1 dB/mm (N=7). It is well documented that the magnitude in the low-frequency tail is robust and independent of the physiological condition. The most likely origin of the spatial growth is therefore the stiffness gradient of the BM [7]. Note that the irregularities in the curves are not due to limitations in measurement accuracy, which is ~1 dB in our recordings.

The high-frequency magnitude curves (Fig. 2B) were non-monotonic, except at the lowest frequencies (<12 kHz) of this group which still fall within the low-frequency tail range of the BM range shown. The basal portion of the magnitude curves grew toward the peak with spatial slopes ranging between 10-20 dB/mm. The spatial location of the peaks follows a tonotopic order (shifting toward the base with increasing frequency), and each peak is followed by a steep decay at the apical side.

Many of the low-frequency phase curves (Fig. 2C) fail to show a systematic accumulation towards the apex which one would expect from waves that travel from base to apex. In fact the phase differences across the frequencies (offsets between the curves) are more pronounced and consistent than the spatial phase variations within each curve. As with the magnitudes, this is not due to limitations in the accuracy of the phase data; the "irregularities" in Fig. 2C are robust and reproducible. It is only at the highest frequencies (>8 kHz) within the low-frequency group that the irregular phase variations disappear and give way to a systematic phase accumulation (lowest two curves in Fig. 2C). We found this "non-wave like" behavior in all 8 gerbil ears tested.

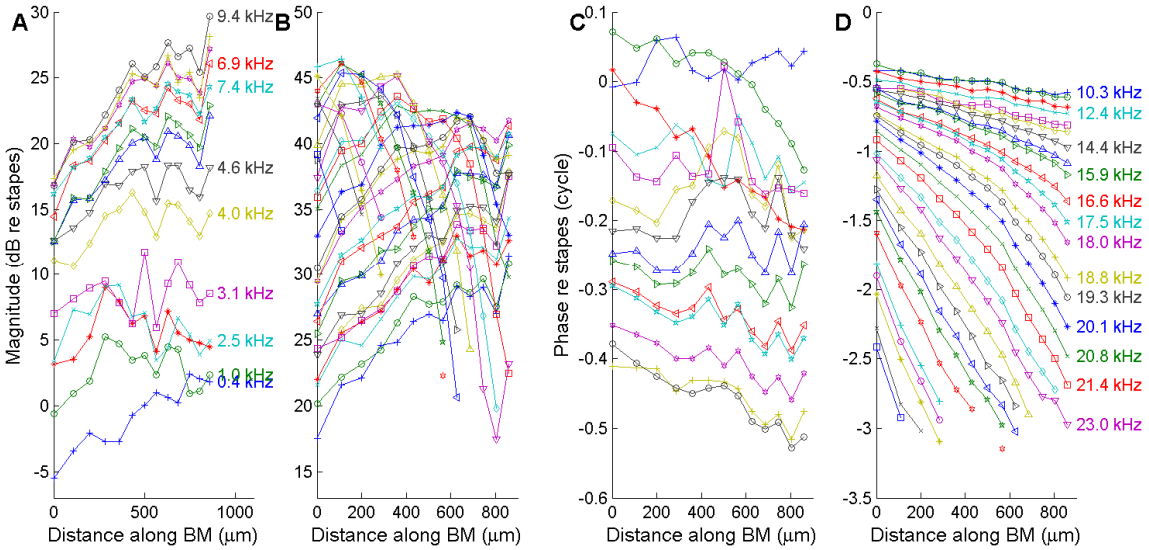


FIGURE 2. Longitudinal patterns of vibration magnitude and phase on the BM. Same data as show in fig. 1. **A**. Magnitude for frequencies below 10 kHz. Individual frequencies are indicated next to the curves **B**. Magnitude for frequencies above 10 kHz. Frequencies of individual curves, see panel D. **C**. Phase for frequency below 10 kHz. Frequencies of individual curves, see panel A. **D**. Phase for frequencies above 10 kHz.

The high-frequency phase curves (Fig. 2D) show a systematic pattern. The lowest frequencies within the group (upper curves) show a systematic, shallow phase accumulation towards the apex, consistent with relatively fast wave propagation. With increasing frequency (lower curves), the curves become steeper, reflecting slower wave propagation. Although the trend is systematic, it is not as regular as may be expected from "scaling," a mathematical form of tonotopy often used in cochlear model studies [8]. Specifically, the 19.3-kHz curve (*black circles*) shows a definite kink near X=550 μm, indicative of an abrupt deceleration of the wave. If scaling symmetry were valid, a similar kink should appear for other frequencies, and the location of those kinks should vary systematically with frequency. The data do not follow this pattern. Frequencies below 17.5 kHz lack a clear kink, while a few of the higher frequencies (20.1-21.4 kHz) do show a kink, but it occurs at the same location instead of shifting tonotopically as dictated by strict scaling. At still higher frequencies, the phase curves become so steep that it is hard to discern any abrupt changes in their slope.

Figure 3 shows a further analysis of this apparent violation of scaling symmetry. In Fig. 3A, the spatial phase profiles of the high-frequency group (Fig. 2D) are reproduced. In order to bring out the difference in shapes between the curves, we determined the average spatial slope (in cycle/mm) over the most basal 400 μm of the spatial range (*vertical dashed line* in Fig. 4). Figure 4B shows the same phase data, but now with the "average basal slope" subtracted over the entire 860-μm range. The resulting normalized phase curves display two distinct features:

1) The spatial phase gradients are virtually flat for $X < 400$ μm. This means that the phase gradient of the raw data is fairly constant in the basal portion, and that the phase velocity and wavelength of the waves change little over the first 400 μm.
2) The sharp deviation from flatness occurs at approximately the same location (X~550 μm) for all frequencies above 17 kHz.

The second observation in particular is a striking violation of scaling symmetry. In a scaling-symmetric cochlea, any abrupt deceleration should occur in a spatially regular fashion, yielding a staggered forking of the normalized phase curves. The frequency range of 17.5-24 kHz of components that show a phase kink (Fig. 3B) correspond, via scaling, to a ~400-µm long stretch of BM, so the staggered forking should be distributed over this length. The actual phase pattern (Fig. 3B) does not indicate any spatial distribution and the phase kinks can be described as "place fixed."

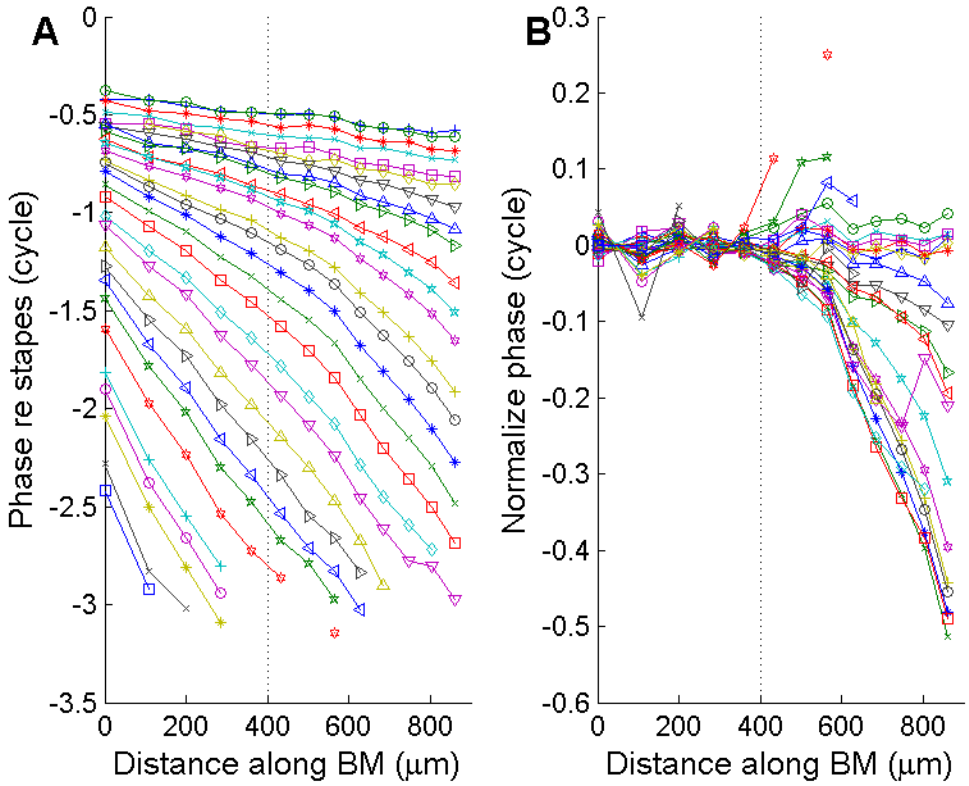


FIGURE 3. Spatial phase profiles analyzed for localize kinks. **A**. Spatial phase profiles reproduced from Fig. 2D. *Vertical dashed line* marks the region used for phase normalization (see text). **B**. Phase profiles comensates for their individual average slope over the region X<400 µm.

DISCUSSION

The detailed quantitative analysis of longitudinal spatial profiles of BM vibration has yielded two unexpected findings. (1) Vibrations evoked by frequencies in the low tail (more than ½ octave below best frequency) do not display a regular phase accumulation expected from traveling waves. (2) Vibrations evoked by frequencies near the best frequency show distinct spatial phase kinks, indicative of an abrupt slowdown of the wave. These abrupt transitions do not obey tonotopic regularity, but instead occur at a fixed location for a range of frequencies. Not only is this a clear violation of scaling symmetry, it also suggests the presence of significant structural irregularities in the mechanical properties of the cochlear partition. We observed such kink patterns in most of the ears we tested, which suggests that they are a common feature in the gerbil cochlea. Is it a bug or a feature? That is, are the irregularities harmless imperfections (maybe even species-dependent ones) that do not hamper cochlear function, or do they reflect a mechanism that is at the heart of cochlear function?

Work supported by Netherlands Organization for Scientific Research, ALW 823.02.018, and an EU Horizon 2020 Marie Sklodowska-Curie Action Innovative Training network, H2020-MSCA-ITN-2016, LISTEN – 722098.

REFERENCES

1. T. Ren, W. He, E. Porsov, “Localization of the Cochlear Amplifier in Living Sensitive Ears,” Plos One **6**, e20149 (2011).
2. M. van der Heijden and N. P. Cooper, "Wave propagation in the mammalian cochlea", AIP Conference Proceedings **1965**, 080002 (2018) https://doi.org/10.1063/1.5038486
3. N. P. Cooper, A. Vavakou, and M. van der Heijden, “Vibrational hotspots reveal longitudinal funneling of sound-evoked motion in the mammalian cochlea,”, Nat. Commun. **9**, 1–12 (2018).
4. M. van der Heijden, P. X. Joris, “Cochlear Phase and Amplitude Retrieved from the Auditory Nerve at Arbitrary Frequencies,” J. Neurosci., **3** 9194-9198; (2003) https://doi.org/10.1523/JNEUROSCI.23-27-09194
5. C. P. C. Versteegh and M. van der Heijden, “Basilar membrane responses to tones and tone complexes: nonlinear effects of stimulus intensity,” J Assoc Res Otolaryngol **13**:785–798 (2012).
6. N. P. Cooper, M. van der Heijden, “Spatial profiles of sound-evoked vibration in the gerbil cochlea,” AIP Conference Proceedings **1965**, 080001 (2018); https://doi.org/10.1063/1.5038485
7. G. Emadi, C.-P. Richter, P. Dallos, “Stiffness of the Gerbil Basilar Membrane: Radial and Longitudinal Variations,” J Neurophysiol **91**, 474–488 (2004).
8. G. Zweig, “Finding the impedance of the organ of Corti.” J. Acoust. Soc. **89**, 1229-1254 (1991).

Perturbing the Cochlea

C. Elliott Strimbu,[1, a)] Elika Fallah,[1, b)] and Elizabeth S. Olson[1, 2, c)]

[1)]*Department of Otolaryngology Head and Neck Surgery, Columbia University, New York, NY, USA*
[2)]*Department of Biomedical Engineering, Columbia University, New York, NY, USA*

[a)]*Corresponding author: CE Strimbu ces2243@cumc.columbia.edu*
[b)]*Electronic mail: ef2516@columbia.edu*
[c)]*Electronic mail: eao2004@columbia.edu*

Abstract. Cochlear mechanics can be studied by perturbing physiological and mechanical components of the organ of Corti (OC) and observing the outcomes. We have combined OCT-based *in vivo* vibrometry at the base of the gerbil cochlea with pharmacological perturbation of different components of the amplifier, including transiently abolishing the endocochlear potential (EP) with intravenous furosemide and inhibition of somatic electromotility by introducing sodium salicylate into the perilymphatic space. Vibrations in healthy cochleae were measured before and for several hours after the pharmacological perturbations to characterize the loss and recovery of the active process. DPOAEs were monitored and, for the furosemide experiments, EP and local cochlear microphonic observations were available from a previous set of experiments in our lab. For both salicylate and furosemide perturbations, outer hair cell sub-best-frequency (BF) nonlinearity recovered before the BF peak. In the salicylate studies intra-OC changes in the motion occurred as the cochlea recovered. In the furosemide experiments, the recovery of the BF peak occurred many minutes after EP recovery. Normal transduction currents have been shown to be necessary to maintain stereocilia morphology and loss of EP might transiently damage hair cell stereocilia, leading to the delayed recovery of functional amplification of the BF peak. We explored this hypothesis with histological studies of bundle morphology following furosemide. These observations have not yielded clear-cut results – bundle morphology usually appeared normal, although important changes might have occurred at more subtle levels. In sum, the constellation of factors that together give rise to cochlear amplification include EP, electromotility and transducer nonlinearity, and as-yet unidentified factors that must be properly aligned to give rise to a functioning whole.

INTRODUCTION

The endocochlear potential, the $\sim$ +80 mV potential in scala media, is essential for normal cochlear amplification. In previous studies on furosemide, we reversibly eliminated the EP by an intravenous (IV) injection of furosemide in gerbil and measured the basal local cochlear microphonic (LCM) [1], the vibrations of the OC complex (OCC) [2], and distortion product otacoustic emissions (DPOAEs). The EP recovered over $\sim$ 40 minutes while vibrations, LCM and DPOAEs recovered later, $\sim$ 2 hours post injection. Vibration was measured using optical coherence tomography (OCT). The axis of OCT measurement had both longitudinal and transverse components, and the measured motion was a combination of these [3].

In a second set of experiments, we transiently blocked OHC electromotility by introducing millimolar sodium salicylate into the scala tympani using a trans-round window membrane application, and monitored DPOAEs and OCC vibration. In the salicylate experiments we expanded the OCT vibration methodology employed in the furosemide study by constructing areal maps of the motion by taking a series of one-dimensional vibration measurements across the nearly radial field of view. The axis of OCT measurement was similar in the salicylate and furosemide experiments.

Following both furosemide and salicylate treatment, vibration amplitudes were diminished with basilar membrane (BM) vibrations resembling those of a passive cochlea, showing linear growth and a loss of the best frequency (BF) peak. The OHC region, some 60 μm deeper in the cochlear partition also showed decreased amplitudes and a loss of the BF peak, but in contrast to the BM, its vibrations retained broad-band nonlinearity at frequencies below BF following both furosemide and salicylate. Recovery of the BF peak occurred in both studies over hours, taking longer following salicylate.

Vélez-Ortega *et al.* found that normal transduction currents are necessary to maintain stereocilia morphology [4]. We hypothesized that the loss of EP following IV furosemide will thus damage hair cell stereocilia and may in part explain the delayed recovery (after EP recovery) of cochlear amplification. In the new results presented here, we investigate the effect of IV injection of furosemide on the morphology of the OHC bundles by *ex vivo* imaging of the OHC bundles using scanning electron microscopy (SEM). Other new results presented here are *in vivo* areal maps of the vibration of the OCC during recovery from furosemide, to compare to our previously reported vibration maps following salicylate. All experiments were performed in young adult gerbils.

METHODS

Experimental protocols were approved by the Columbia University Institutional Animal Care And Use Committee. Details of the surgery, OCT-based vibrometry, and introduction of furosemide and salicylate have been published [1, 2, 5, 6]. In brief, adult gerbils were anesthetized with ketamine, sodium pentobarbital, and buprenorphine and the left auditory bulla was exposed ventrally and opened. Furosemide, 100 mg/kg, was injected intravenously in the left femoral vein. In the salicylate experiments, 5 μL of 50 mM sodium salicylate in artificial perilymph was deposited on the round window membrane and passively diffused into scala tympani. In the cochlear base, the extracellular salicylate concentration should rapidly reach an equilibrium concentration of $\sim$2.5 – 5 μM, sufficient to block somatic electromotility [7].

A Tucker Davis Technologies D-to-A and A-to-D system running at 97656.25 S/s was used to generate the acoustic signals and to monitor the ear canal pressure. Stimuli were broad-band zwuis complexes in which a number of frequencies, $N = 15$ in these experiments, were presented simultaneously. The frequencies are chosen such that they are linearly independent up to third order; thus there are no harmonics or second- or third-order distortion products in the frequency complex itself, and each sinusoidal component was assigned a random phase so that the total pressure amplitude is $\sim \sqrt{N}$ higher than each primary in the complex. Stimuli were presented from 50 or 60 – 80 dB SPL (where 0 dB SPL $\equiv$ 20 μPa) at each component.

DPOAEs were measured in response to swept two tone, f_1 and f_2, stimuli where f_2 was varied from 1 to 48 kHz, f_1 and f_2 were held in a fixed ratio of $f_2/f_1 = 1.2$, and both frequencies were presented at 50 and 70 dB SPL. The prominent $2f_1 - f_2$ DPOAE served as a real-time measure of general cochlear condition. Baseline DPOAEs were consistent with healthy cochleae; the loss of the DPOAEs at 50 dB and sharp reduction at 70 dB immediately after treatment was used to assess the efficacy of the drug delivery and the recovery of the DPOAEs as a function of time indicated the recovery of cochlear amplification.

For vibrometry, the OCC was imaged through the intact round window membrane near the $\sim$ 25 – 28 kHz location. Measurements were performed with a ThorLabs Telesto III OCT system with a central wavelength of $\sim$ 1300 nm. The lateral resolution is set by the optics and is $\sim$ 10 μm when measuring in water with an index of refraction $n = 1.3$. The axial resolution is set by the wavelength and bandwidth of the light source; the distance between adjacent axial pixels is 2.7 μm in water. The OCT's line camera was locked to the TDT system clock to acquire time-locked 1-dimensional Ascans, or Mscans, which are then analyzed to find the motion of structures along the OCT axis of measurement. The time waveforms of motion at pixels of local maxima in the Ascans, which correspond to anatomical structures in the OCC, were determined by spectral domain phase microscopy and the amplitude and phases of the motion were then determined by Fourier analysis. Before and after each vibration measurement, the system took two dimensional scans, or Bscans, of the OCC and the two images were compared to check for drift in the preparation. Two dimensional, or areal, maps of the vibrations were measured by taking sequential Mscans in 10 μm steps across an approximately radial section of the OCC. Linear interpolation was applied to these coarse maps to construct vibration maps that had the same number of pixels as the structural Bscans. This procedure is explained in more detail in Fig.1.

In each experiment, two sets of baseline recordings (DPOAEs and vibrometry measurements) were taken prior to pharmalogical treatment. In the furosemide experiments, DPOAEs and vibration measurements were taken immediately after the injection was completed and at subsequent times for up to five hours post injection. In the salicylate study, we applied the sodium salicylate solution into the round window space and allowed the solution to passively diffuse into scala tympani for at least 10 minutes before wicking off the solution and washing with fresh artificial perilymph. DPOAE-grams were measured in 10 minute intervals to check for a loss the DPOAEs evoked by the 50 dB primaries. The DPOAE and vibrometry measurements were then repeated approximately every 20 minutes for several hours post treatment.

For SEM histology, gerbils were anesthetized and DPOAEs were measured before and after furosemide injection to confirm the loss of the EP [1]. In order to observe the state of the stereocilia soon after the EP had dropped, the gerbil was euthanized minutes after the confirming DPOAE measurement. Cochleae were extracted, fixed in 2.5% glutaraldehyde supplemented with 2 mM of $CaCl_2$, dehydrated in graded series of ethanol, critical point dried from liquid CO_2, sputter coated with 5 nm platinum, and imaged with a ZEISS VP SEM. Control imaging was done in gerbils without furosemide injection. The electron high tension of the SEM (EHT) was 5 kV.

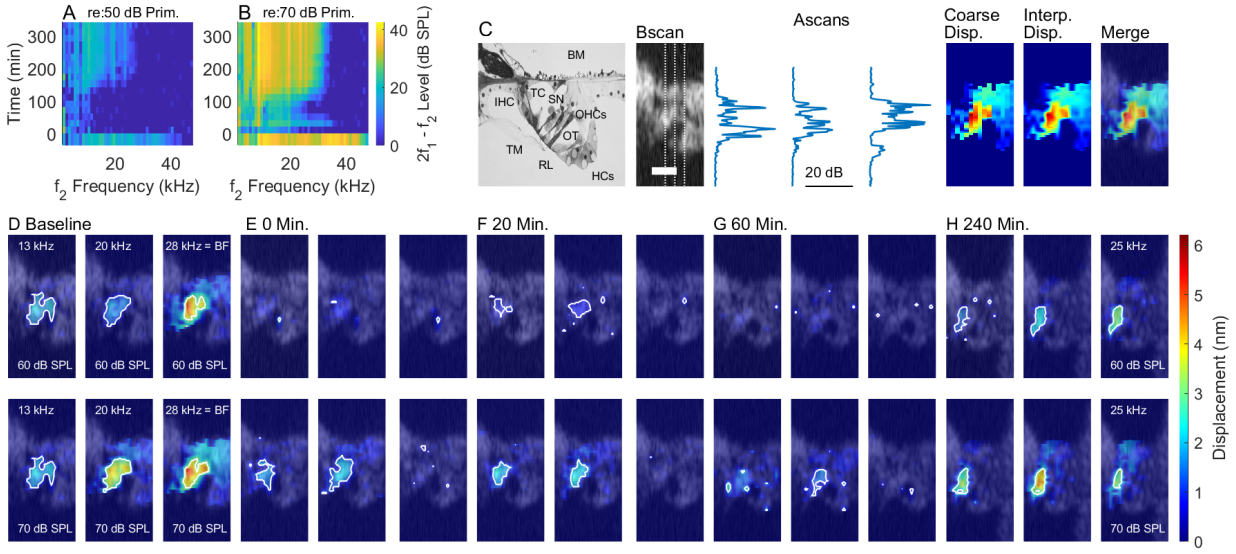


FIGURE 1. (A) and (B) show heat map plots of the DPOAE-grams measured in response to 50 and 70 dB SPL two tone stimuli before and after the administration of furosemide. In these plots the f_2 frequency runs along the horizontal axis, time runs along the vertical axis, and the SPL of the $2f_1 - f_2$ DPOAE level is indicated by the color bar at the right which applies to both panels. Prior the injection, this ear had robust DPOEAs which were abolished immediately after the injection (time $t = 0$). The DPOAEs gradually recovered over the subsequent few hours, but the recovery was not monotonic at 70 dB. (C) Explanation of the method for obtaining the two-dimensional vibration maps. The left image is a micrograph of a guinea pig OCC with selected structures labeled: BM = basilar membrane, TM = tectorial membrane, OHCs = outer hair cells, IHC = inner hair cell, TC = tunnel of Corti, OT = outer tunnel, SN = space of Nuel, RL = reticular lamina, OT = outer tunnel, HCs = Hensen's cells. The image to the right is a Bscan centered around the OHC-region. The OHCs, tunnel of Corti, and outer tunnel are visible in this scan. Scale bar = 50 μm. Time locked Ascans or Mscans were taken across the field of view in 10 μm steps and the vibration amplitudes were determined at selected maxima in the Ascans. Three representative Ascans are shown to the right and were taken along the dashed lines in the Bscan. The vibration amplitudes as a function of position give the coarse displacement map at the right. Linear interpolation was then applied to this to give a heat map of the vibration amplitude with the same resolution, or number of pixels, as the imaging Bscan. Finally the heat maps are overlaid onto the Bscans to match the vibration amplitudes with structures in the OCC. (D–H) show vibration maps at three frequencies and two levels before (D) and at selected times after (E–H) IV furosemide and the loss of the EP. The amplitudes are indicated by the colorbar at the far right which applies to all displacement maps. At each time, pixels with the largest vibrations, within 2 standard deviations of the maximum, are outlined in white. Experiment #925, 10/13/2021.

RESULTS

Fig. 1 shows loss and recovery of the DPOAEs and vibration responses following IV furosemide injection. The DPOAE-grams are shown as heat maps and the two-dimensional vibration maps show the amplitude of the axial vibrations across an approximately radial field of view at selected times before and after treatment. Following the injection and loss of EP, the DPOAEs evoked by the 50 dB primaries were abolished and those evoked by the 70 dB primaries were greatly reduced in amplitude. At 70 dB the DPOAEs showed a nonmonotic recovery with a second stage of recovery beginning at about 120 minutes and recovery largely complete by 240 minutes.

Baseline vibrations showed a stereotyped pattern with the highest amplitudes focused in a narrow region, or "hot spot" using the terminology of Cooper *et al.* [8], extending through the bodies of the OHCs, from the base of the OHCs/Deiters' cell junctions to the RL. In the heat maps, pixels with the largest amplitudes, within two standard deviations of the maximum, are outlined in white to highlight the shape of the vibration patterns. At higher SPLs and at frequencies closer to the BF, the vibration patterns were more diffuse, showing substantial BM motion at points lateral to the intersection of the acuate and pectinate zones (Fig. 1 D). Following the furosemide injection and subsequent loss of the EP, vibration amplitudes fell uniformly at all positions. In spite of the decreased amplitude, the vibrations remained strongest in the OHC region, as shown in Fig. 1 E taken right after the injection. Over the subsequent four hours, this cochlea showed unsteady recovery, with initial recovery at 20 minutes followed by a reduction at 60 minutes. Nonmonotonic recovery was also observed in the DPOAE as noted above, and in our previous furosemide studies, where nonmonotonicity was linked to shifts in MET operating point [1, 2]. At the 240

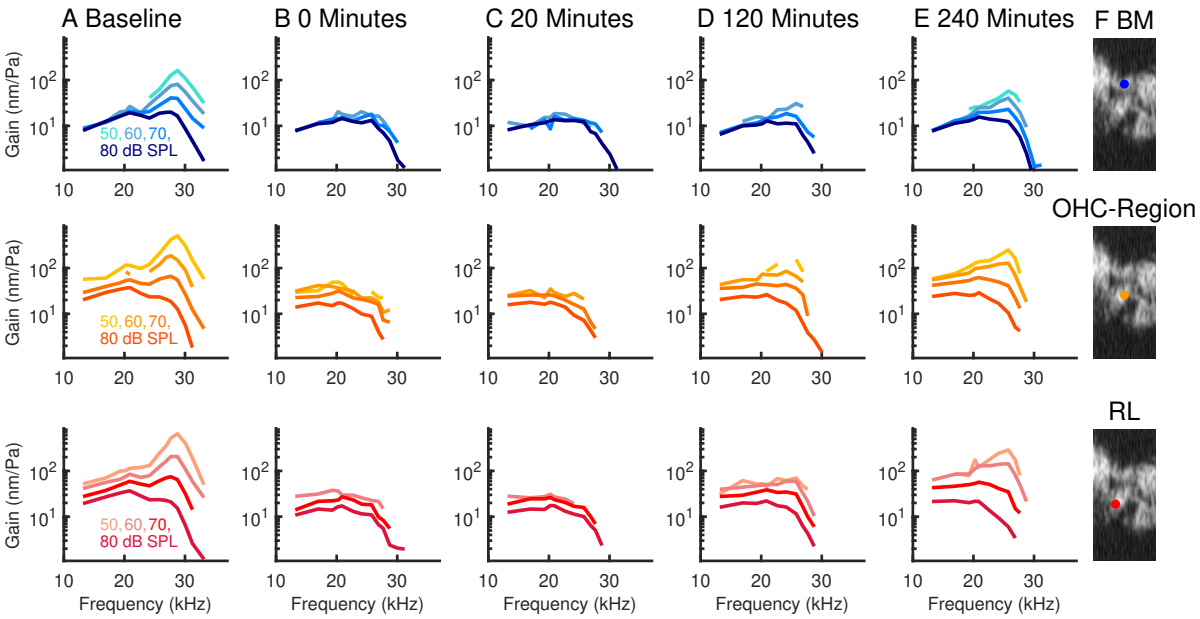


FIGURE 2. (A) – (F) Conventional tuning curves measured at the BM (top row, shades of blue), OHC-region (middle row, shades of orange), and RL, close to the apical surface of the OHCs, (bottom row, shades of red) before (A) and at selected times (B – E) after furosemide. Darker shades indicate higher SPLs and the times at the top of each column are the minutes that elapsed post treatment. Time $t = 0$ in column (B) corresponds to recordings taken immediately after the injection. (F) The Bscans at the far right indicate the points on the OCC where the measurements were taken. Data are from the same experiment as Fig. 1.

minute mark, the vibration patterns had largely recovered to the baseline conditions, with the OHC-region having pre-treatment amplitudes. This preparation showed a somewhat delayed recovery relative to other individual experiments and a slightly reduced BF following recovery.

Fig. 2 shows tuning curves from the same experiment for three locations: the BM, the OHC-region, and the RL measured 20 μm medial to the BM and OHC-region and close to the apical surface of the OHCs (see Bscans at right). The baseline vibrations were consistent with a healthy cochlea: the OHC-region and RL vibrations displayed higher amplitudes than the BM and a broad-band nonlinearity extending across all frequencies tested. Following the loss of the EP, the vibration amplitudes fell at all locations and the BM vibrations resembled those of a passive cochlea while the OHC-region and RL vibrations retained their broad-band nonlinearity. At around the 120 minute mark, the BF peak began to re-emerge and by 4 hours the vibrations appeared close to baseline levels.

The effects of salicylate on the areal vibration patterns were recently reported by our group [9] and are included to compare with the furosemide results that are newly reported here. Blocking somatic electromotility with extracellular salicylate likewise led to a reduction in the amplitudes at all points and a loss of the BF peak. Similar to the furosemide results, the BM responses appeared passive post-salicylate while the OHC-region and RL showed compressive growth for frequencies below BF. (See [9] for individual tuning curves.) In contrast to the treatment with furosemide, however, salicylate led to more qualitative changes in the vibration patterns. Post treatment, the vibrations were more diffuse and frequently showed large vibrations that were present in the lateral compartment of the OC, close to the Hensen's cells, as shown in Fig. 3 C and D. These vibrations appeared around 30 to 40 minutes after the salicylate was introduced into the perilymphatic space; could be greater in amplitude than the BM and OHC-region vibrations at the same time points (while having smaller amplitudes than the baseline vibrations); and were about as well tuned as the BM and OHC-vibrations at these times. As healthy responses recovered, these lateral vibrations faded and at the 200 minute time point, the strong OHC-region to BM axis of motion was again prominent.

Switching gears to the histological study of cochleae from furosemide-treated preparations, Fig. 4 shows scanning electron microscopy (SEM) images of OHC stereocilia bundles from the basal and middle turns. Control SEM images (Fig. 4 A–D) were taken from experiments without furosemide injection. Panels E–F depict the OHC bundles from the cochleae fixed between 20 – 40 minutes post furosemide. To date, results from the SEM images of the OHC

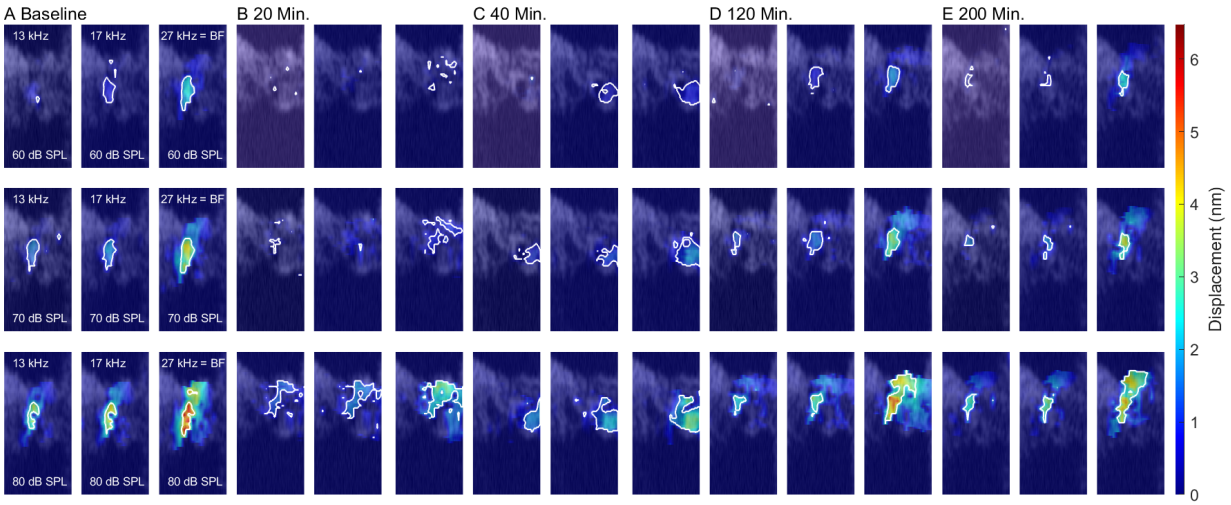


FIGURE 3. Effects of extracellular salicylate on two-dimensional vibration patterns. (A–E) Show vibrations shown at three frequencies and three levels before (A) and at selected times after (B – E) salicylate was introduced into scala typani via diffusion across the round window membrane. Pixels with the largest vibrations, within 2 standard deviations of the maximum, are outlined in white. Similar to the loss of the EP due to furosemide, the vibrations show a loss of tuning and amplitude, but transiently show large vibrations in the lateral region, around the Hensen's cells. These are particularly clear at the highest SPL (bottom row) in panels (B) and (C). These vibrations were present $\sim$ 20 – 40 minutes post treatment and were not apparent at later times. Experiment #855, 12/15/2020. For more details see [9].

bundles after furosemide showed no consistent significant change in the morphology of the OHC stereocilia. In one case, we observed shortening at the first (shortest) row of the OHC stereocilia in the base (Fig. 4 F compared to Fig. 4 B). However, that result was not repeated in additional experiments.

DISCUSSION

Changes in BM vibrations following intravenous furosemide [10] and perilymphatic perfusion of salicyaltes [11, 12] have been known since the 1990s but the effects on other structures within the OCC were unexplored until the advent of OCT and other depth-resolved interferometric techniques [13]. In the furosemide experiments, the complex loss and recovery of the active process following the drop and recovery of the EP suggest changes in the operating point of the hair cell mechano-electrical transduction (MET) complex, but the mechanism remains elusive [1, 2]. Using temporal bone explants, Jacob *et al.* showed that restoring EP led to micromechanical rearrangements in the OC and an increased gain relative to the passive, EP $\sim$ 0 mV, system [14]. Changes in the resting position of the OHCs would alter the open probability of the MET channels and could explain these operating point shifts. The conformational changes observed in [14] were less than a few hundred nanometers, too small to be observed with an OCT in imaging mode where the resolution is on the order of a few microns. SEM imaging could theoretically observe at the 100-nanometer level, but in practice, processing-induced distortion would mask these conformational changes.

Using OCC explants, Vélez-Ortega *et al.* found that blocking normal transduction currents led to abnormal stereocilia lengths in the shorter two, or MET carrying, rows of OHC-bundles [4]. Eliminating the EP with furosemide will reduce the driving force for cation entry into the hair bundles and could lead to similar changes in the bundle morphology *in vivo*. Such changes would, in turn, alter the resting open probability of the MET channels at rest and could provide an explanation for the operating point shift we observed [1, 2]. Because bundle repair is a slow process, this could explain the slow and multiple time scales of the recovery we see post furosemide. However, the SEM studies did not uncover changes in stereocilia morphology following furosemide. Nevertheless, the SEM studies are just a single "snapshot", and it is possible that changes occurred or would have occurred before or after this snapshot (as was hinted in Fig. 4 F). While the findings here do not support the hypothesis that the stereocilia are involved in recovery following furosemide, they also do not disprove it. Continuous in-vivo imaging of stereocilia following furosemide would be useful to probe this hypothesis.

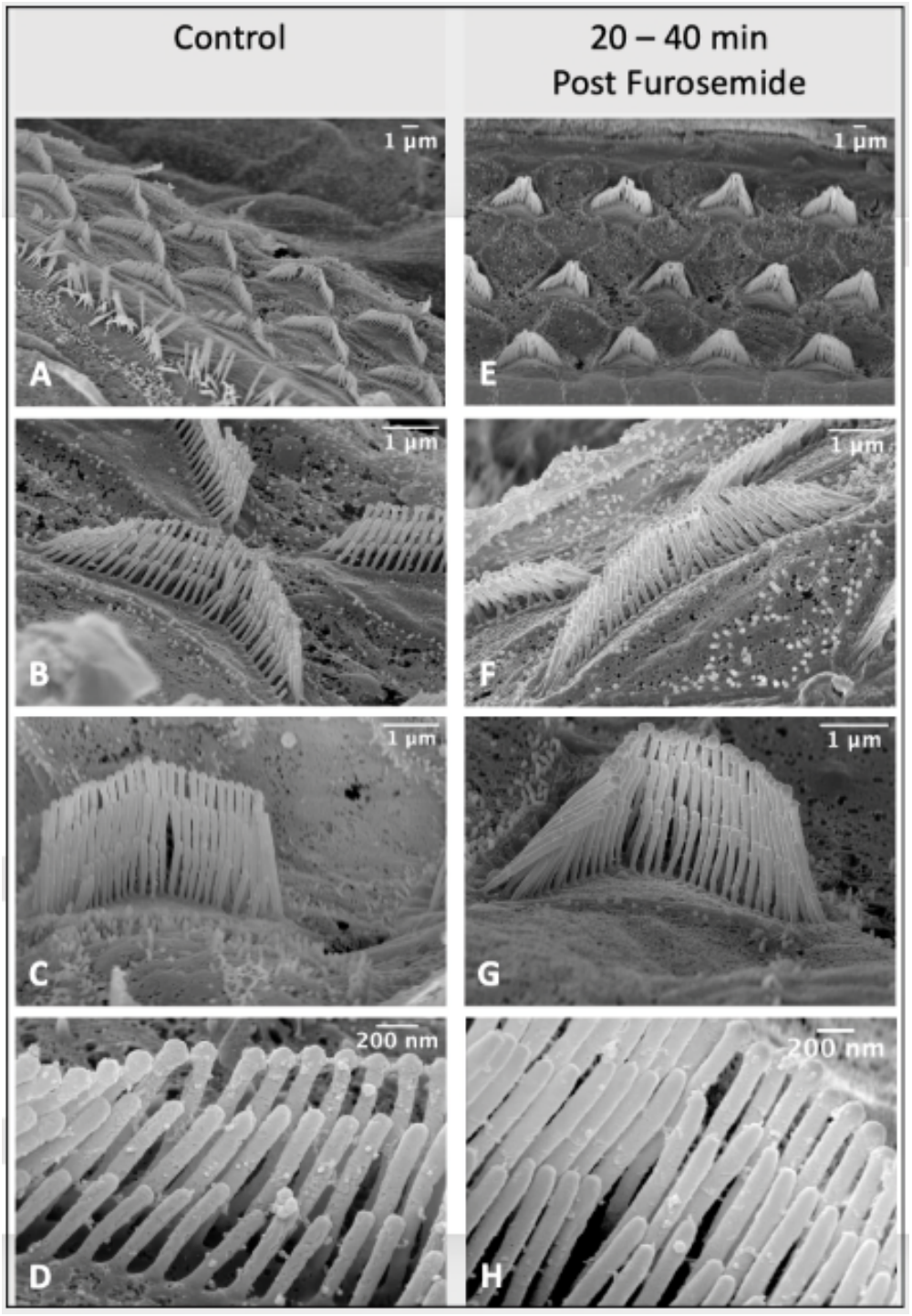


FIGURE 4. Scanning electron microscopy (SEM) images of OHC stereocilia bundles in gerbil. (A – D) from healthy cochlea (g919-LE). (A & B) Basal turn (C) Middle turn (D) Basal turn (apical end). (E -– H) SEM images 20 – 40 minutes after IV injection of furosemide (g916-LE, g923-LE). (E & F) Basal turn. (G & H) Middle turn.

Perfusing salicylate into the perilymphatic space led to reduced amplitudes and a loss of the BF peak, similar to the effects with IV furosemide. Beyond that, salicylate changed the qualitative nature of the areal vibration patterns with the lateral compartment of the OC transiently exhibiting larger vibrations some 20 – 40 minutes post treatment but not at later times. These qualitative changes in the motion patterns would not be captured by single-location uniaxial measurements.

CONCLUSIONS

The vibrations of the OCC are complicated, ultimately leading to the radial shearing motion between the reticular lamina and tectorial membrane that provides the mechanical input to the inner hair bundles' MET machinery, and hearing. Precisely how this shearing motion is generated, and how the active process or cochlear amplifier contributes to it is presently not well understood. The cochlea operates under feedback and determining the contributions of each component of the active process is challenging. Selective pharmacological blocking of individual elements of the cochlear amplifier *in vivo* sheds insight into how the elements of cochlear activity work in synchrony.

ACKNOWLEDGMENTS

This work was supported by the NIDCD grant DC015362 and the Emil Capita Foundation.

REFERENCES

1. Y. Wang, E. Fallah, and E. S. Olson, "Adaptation of cochlear amplification to low endocochlear potential," *Biophysical Journal*, Biophysical Journal **116**, 1769–1786 (2019).
2. C. E. Strimbu, Y. Wang, and E. S. Olson, "Manipulation of the endocochlear potential reveals two distinct types of cochlear nonlinearity," *Biophysical Journal*, Biophysical Journal **119**, 2087–2101 (2020).
3. B. L. Frost, C. E. Strimbu, and E. S. Olson, "Using volumetric optical coherence tomography to achieve spatially resolved organ of Corti vibration measurements," J Acoust Soc Am **151**, 1115 (2022).
4. A. C. Vélez-Ortega, M. J. Freeman, A. A. Indzhykulian, J. M. Grossheim, and G. I. Frolenkov, "Mechanotransduction current is essential for stability of the transducing stereocilia in mammalian auditory hair cells," eLife **6**, e24661 (2017).
5. N. C. Lin, C. E. Strimbu, C. P. Hendon, and E. S. Olson, "Adapting a commercial spectral domain optical coherence tomography system for time-locked displacement and physiological measurements," AIP Conference Proceedings **1965**, 080004 (2018), https://aip.scitation.org/doi/pdf/10.1063/1.5038488.
6. E. Fallah, C. E. Strimbu, and E. S. Olson, "Nonlinearity and amplification in cochlear responses to single and multi-tone stimuli," Hearing Research **377**, 271–281 (2019).
7. I. I. Sadreev, G. W. S. Burwood, S. M. Flaherty, J. Kim, I. J. Russell, T. I. Abdullin, and A. N. Lukashkin, "Drug Diffusion Along an Intact Mammalian Cochlea," Front Cell Neurosci **13**, 161 (2019).
8. N. P. Cooper, A. Vavakou, and M. van der Heijden, "Vibration hotspots reveal longitudinal funneling of sound-evoked motion in the mammalian cochlea," Nat Commun **9**, 3054 (2018).
9. C. E. Strimbu and E. S. Olson, "Salicylate-induced changes in organ of corti vibrations," Hearing Research , 108389 (2021).
10. M. A. Ruggero and N. C. Rich, "Furosemide alters organ of corti mechanics: evidence for feedback of outer hair cells upon the basilar membrane," J Neurosci **11**, 1057–1067 (1991).
11. E. Murugasu and I. J. Russell, "Salicylate ototoxicity: The effects on basilar membrane displacement, cochlear microphonics, and neural responses in the basal turn of the guinea pig cochlea," Auditory Neuroscience **1**, 139–150 (1995).
12. J. Santos-Sacchi, L. Song, J. Zheng, and A. L. Nuttall, "Control of mammalian cochlear amplification by chloride anions," Journal of Neuroscience **26**, 3992–3998 (2006).
13. T. Ren, W. He, and P. G. Barr-Gillespie, "Reverse transduction measured in the living cochlea by low-coherence heterodyne interferometry," Nat Commun **7**, 10282 (2016).
14. S. Jacob, M. Pienkowski, and A. Fridberger, "The endocochlear potential alters cochlear micromechanics," Biophys J **100**, 2586–2594 (2011).

Human Cochlear Partition Anatomy and Motion Using Optical Coherence Tomography

Christopher I. McHugh[1,2], Stefan Raufer[1,2], Nam Hyun Cho[1,2], Aleksandrs Zosuls[3], Andrew A. Tubelli[1,2], Michael E. Ravicz[1,2], Kevin N. O'Connor[1], John J. Guinan[1,2], Sunil Puria[1,2] and Hideko Heidi Nakajima[1,2, a)]

[1]*Eaton-Peabody Laboratory of Auditory Physiology, Massachusetts Eye and Ear, 243 Charles St., Boston MA 02114 USA.*
[2]*Department of Otolaryngology-Head and Neck Surgery, Harvard Medical School, Boston MA 02115 USA*
[3]*Boston University College of Engineering, 44 Cummington Mall Boston MA 02215 USA*

[a)] Corresponding Author: Heidi_Nakajima@meei.harvard.edu

Abstract. It has long been assumed that the motion of the cochlear partition (CP) is similar for all mammals. This "classic view" has been inferred from basal-turn measurements in laboratory animals (e.g., mouse, rat, gerbil, guinea pig, chinchilla, cat). However, our recent findings reveal the human CP anatomy and motion to be substantially different from the classic view, challenging the assumption that cochlear mechanics is universal among mammals. Inconsistent with the classic view, the human CP has a soft tissue structure between the bony osseous spiral lamina and inner pillar cells – the CP "bridge" – that has only recently been anatomically delineated. Laser Doppler vibrometry of the basal human cochlea showed the entire CP to be moving, with the osseous spiral lamina and bridge moving nearly as much as the BM, as opposed to solely the BM region as in the classic view. To better understand the implications of this new finding, we imaged the anatomy and measured the motion of the CP and organ of Corti structures with optical coherence tomography (OCT) in a ~13 h post-mortem fresh human specimen. This approach allows measurements of the intact cochlea through the round-window membrane. We present OCT images that detail the human CP and allow identification of structures within the organ of Corti. OCT vibrometry of CP structures enables us to measure motion of the BM, bridge, and osseous spiral lamina driven by air-conducted sound, as with previous methods; additionally, motion of structures such as the limbus, pillar cells, outer hair cells, supporting cells, reticular lamina, and tectorial membrane can now be measured. These preliminary measurements advance our understanding of the unique architecture and motion of CP structures in the human organ of Corti and how they may shape the sensory input at the inner hair cells.

INTRODUCTION

The cochlea has a complex structure that comprises two fluid-filled compartments separated by the cochlear partition (**CP**), which is responsible for the amplification of mechanical vibrations and subsequent transduction to neural signals. The CP contains the cochlear sensory tissue, the organ of Corti (**OoC**), which is bounded by the basilar membrane (**BM**) and tectorial membrane (**TM**), as well as the bony osseous spiral lamina (**OSL**). The introduction of sound produces a displacement of the BM leading ultimately to the firing of auditory nerve fibers. The sensitivity and frequency response of hearing are dependent upon the anatomy and mechanical properties of structures within the CP, which have been previously thought to be conserved among mammalian species. This generalized "classic view" of cochlear physiology has been widely accepted because radial motion measurements recorded in the cochlear base of several laboratory animals have been shown to be similar (1, 2).

Recent studies, however, have indicated several key differences between the anatomy and motion of the human CP and the classic view at the base of the cochlea (3). In laboratory animals (such as mice, gerbils, guinea pigs, chinchillas, and cats), the BM is attached to the short and relatively static OSL. In the classic view of cochlear mechanics, motion of the CP facing scala tympani is restricted to the BM which moves as a simple beam with maximum motion near the BM center (near the three rows of outer hair cells). However, in the human CP, the OSL is long and flexible, and we have identified a highly mobile soft-tissue structure, termed the "bridge" (absent in the classic view) between the OSL and inner pillar cells. Moreover, in human the TM attaches to the mobile limbus above the bridge instead of the classic view's fixed OSL. Measurements using laser Doppler vibrometry (**LDV**) in human temporal bones have demonstrated that BM motion only accounts for <50% of CP volume velocity and the OSL and bridge have significant mobility, with peak motion near the BM-bridge connection (near the inner hair cells). These divergences in cochlear structure and movement may contribute to differences in cochlear mechanics, such as some of the sharper cochlear tuning which has been observed in humans compared to laboratory animals (4-6).

Here, we describe our efforts utilizing optical coherence tomography (**OCT**) in very fresh (<24 h post-mortem) cadaveric specimens to characterize the anatomy and motion of human CP structures *in situ*. OCT has been shown to be an effective method for measuring movement within the cochlea of animals without disrupting the delicate structures of the CP (7-10). Unlike LDV which has been used extensively to study cochlear mechanics, OCT allows for penetration and visualization of cochlear structures beyond surface structures, *in situ.*

METHODS

OCT System and Vibrometry Software

OCT images and vibrometry were acquired using a commercially available Ganymede-III-HR 900-nm OCT system (Thorlabs, Germany) with a high-speed (up to 240-kHz) camera capable of imaging in 1, 2, or 3 dimensions. Images shown below are 2-D B-scans" of a radial cross-section plane of the human cochlea defined by a radial direction *x* (i.e., OSL to SL) and depth *z* (i.e., round window membrane through CP) along an imaging track. Responses are measured at 1024 depths (pixels) at a single radial location ("A-line") simultaneously, and adjacent A-lines in a given B-scan are measured sequentially. The OCT system penetration depth was ~1.44 mm (in water, with a refractive index of 1.33). The axial resolution was ~2.23 µm (in water) and the lateral resolution was ~8 µm, using a 36-mm, 0.055-NA, 2× objective lens (OCT-LK3-BB, Thorlabs, Germany).

The system was operated using custom LabVIEW-based software (VibOCT version 2.1.4). Camera images of the anatomical approach were obtained and used to determine and record the region of interest for the OCT scans. VibOCT provides real-time depth-resolved 1D A-scan, 2D cross-sectional B-scan, 3D volumetric C-scan imaging, synchronous vibrometry acquisition, and parallel-processing-based nearly-real-time processing of measured whole A-line data to extract displacement waveforms and spectra in response to tones or other stimuli. (For a more detailed description of the custom software used herein, see (11)).

Preparation of Fresh Human Temporal Bone Specimen

Measurements were made on 14 specimens. Here, we report measurements from a non-identified human temporal bone harvested ~13 hours post-mortem to minimize degradation of cochlear structures. The bone was prepared using an expanded facial recess approach allowing for maximal visualization of the round window (**RW**) and first turn of the cochlea (12, 13). The stapes was exposed to allow for stapes velocity measurements using LDV. Specimen was mounted to a magnetic holding stand with an articulated arm attached to a ball-joint and maintained at low temperatures using ice packs and chilled saline throughout preparation and experimentation.

Data Acquisition

Stimulus Generation and Measurement

To measure microphone and LDV outputs simultaneously and synchronize OCT vibrometry, custom stimulus generation and synchronous measurement software (SyncAv, version 0.46) running on a National Instruments PCI-4461 board (Austin, TX) was used. Sound was produced by a Radio Shack 40-1377 speaker and was transmitted into the ear canal via a speaker attached to a coupler consisting of a rubber tube and connecting pieces inserted into a hole in the anterior aspect of the bony external acoustic meatus (**EAC**). A calibrated probe-tube microphone (EK-3103; Knowles, Itasca, IL) was threaded through the probe tip and the assembly placed into the EAC with microphone at approximately 1-2 mm from the umbo.

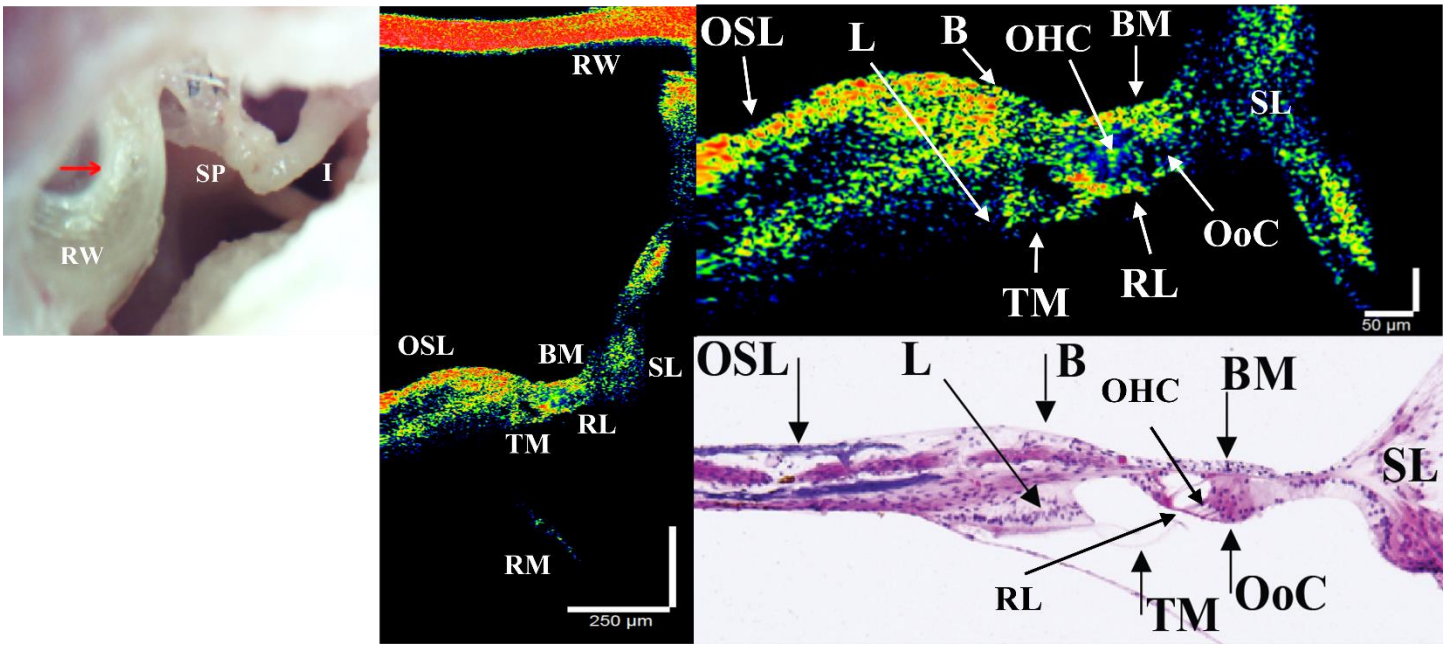
OCT Data Collection

Initial 2D B-scan images of the CP were taken through the intact RW membrane to aid orientation and positioning of the temporal bone such that the optical reflectance and penetration depth were adequate to reveal the structures

of interest. Vibration displacement was then measured in response to applied tones (and with no tone present to estimate noise floor) at multiple radial A-line positions to capture the motion of structures of interest. Tones were presented from 50 to 120 dB SPL in 10-dB steps at 20–40 logarithmically distributed frequencies selected around the approximate best frequency (**BF**) (-3 to +0.5 octaves around BF). For the OCT-vibrometry input reference, an LDV (Polytec, OFV-501/2600) was used to measure velocity of the posterior crus of the stapes simultaneously with the OCT vibrometry.

RESULTS

Images of the Human CP in a Very Fresh Specimen

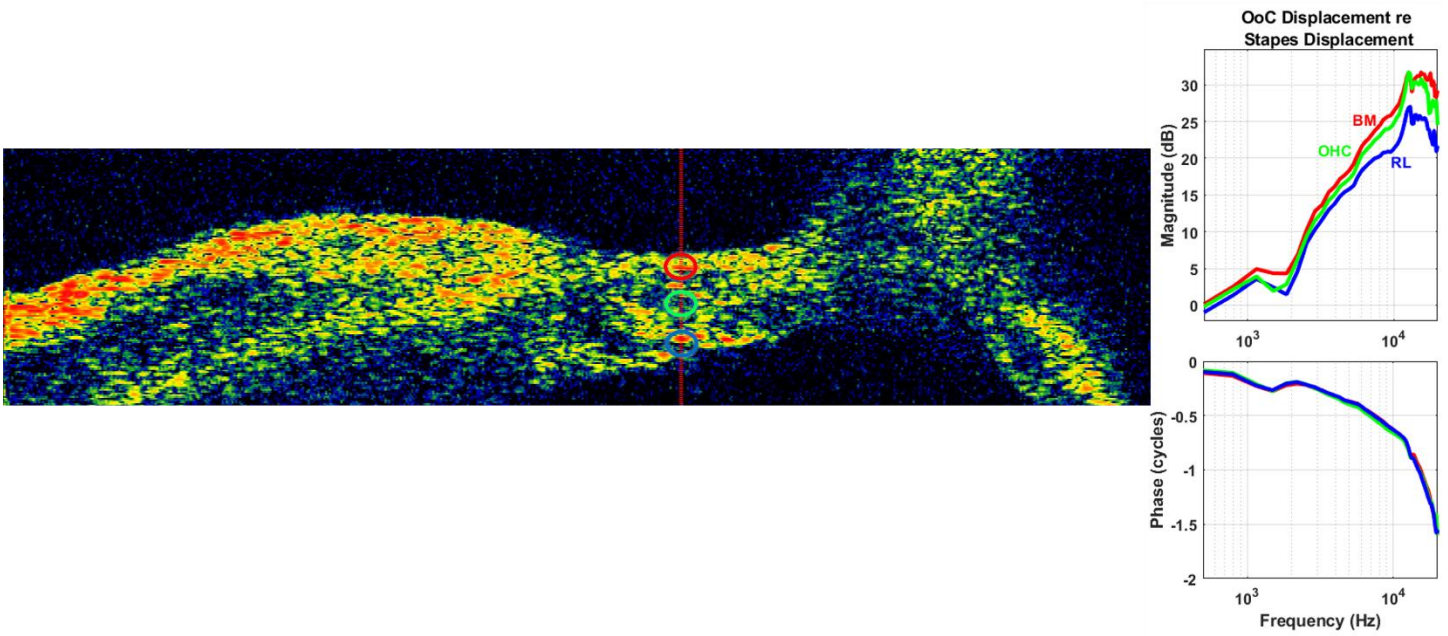


(a) (b) (c) / (d)

Figure 1. Representative images from a fresh specimen (13 hours post mortem). (a) microscopic image of the human middle ear. The arrow over the round window indicates the plane and direction in which the OCT images are taken. (b) B-scan OCT image of the cochlear partition through the RW. (c) Zoomed B-scan image in (b) showing CP structures. (d) Histological image of a fixed human cochlea (different cochlea from (a)-(c)) near the base for reference. B = Bridge region; BM = Basilar membrane; I = Incus; L = Limbus; OoC = Organ of Corti; OHC = Outer hair cell; OSL = Osseous spiral lamina; RL = Reticular lamina; RM = Reissner's membrane; RW = Round window membrane; SL = Spiral ligament; SP = Stapes; TM = Tectorial membrane. [Specimen SR44].

After obtaining adequate exposure to the RW membrane (Figure 1a) we are able to obtain detailed OCT images of the human CP through the RW. Structures in and around the CP can be identified including the BM, OSL, bridge region, TM, spiral ligament (SL), and Reissner's membrane (RM); see Figure 1b and 1c. Structures within the OoC, such as the individual outer hair cells (OHC) are not easy to identify, though their approximate location can be inferred through correlation with a structurally similar histologic image (Figure 1d).

Displacement of Human OoC Structures



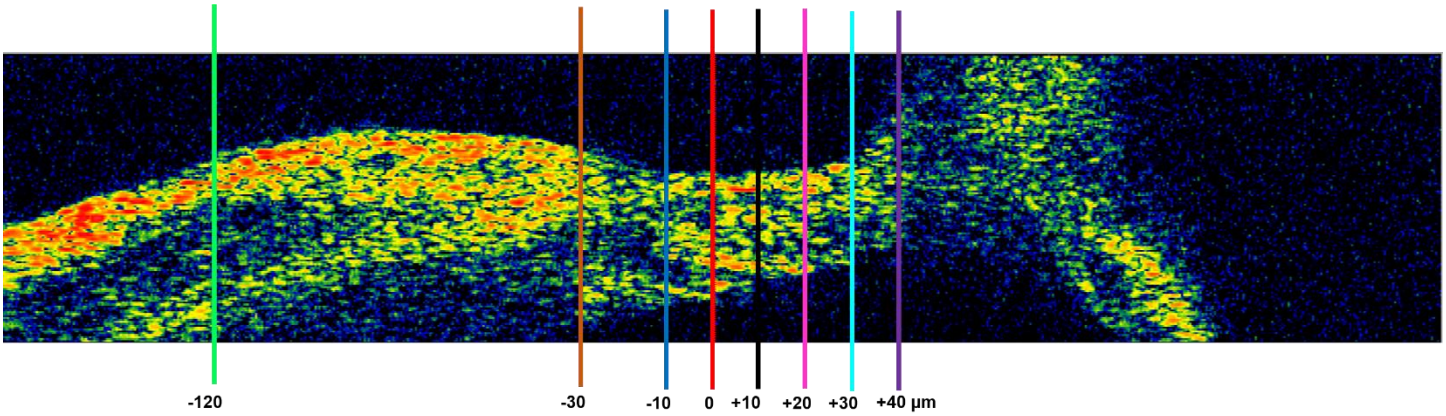
(a) (b)

Figure 2. B-scan image of the cochlear base in a very fresh (13 hours post mortem) human temporal bone, same specimen as Figure 1 (A). (B) Motion displacement of the BM, OHC, and RL obtained using OCT-vibrometry in response to sound, normalized to stapes motion measured simultaneously by LDV. Top: Magnitude (ratio in dB); bottom: phase.

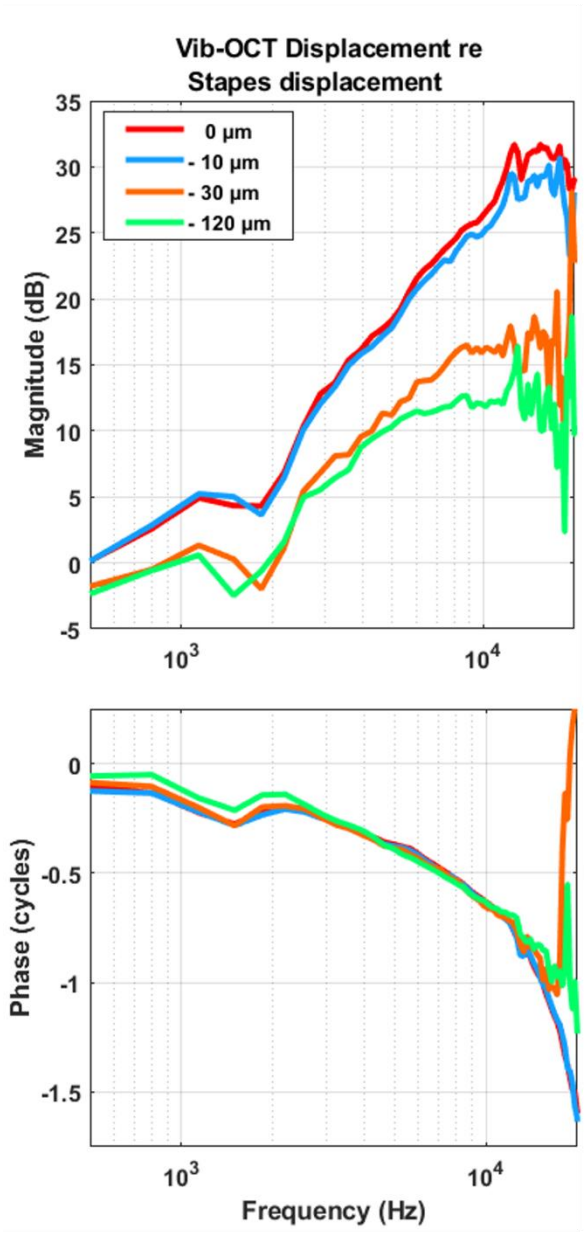
The motions of structures within the human OoC in response to a sequence of applied tones are shown in Figure 2. Figure 2a is a B-scan OCT image of the CP near the cochlear base. The red dashed line is an A-line which passes through the BM (indicated by red circle), an OHC (green circle), and RL (blue circle). Motion is evaluated at the regions indicated by the colored circles and is normalized to the displacement of the stapes. Maximum displacement of all three structures occurred at approximately 12-15 kHz, consistent with the best frequency at the measurement location at the basal aspect of the cochlea (Figure 2b). Displacement magnitude was highest for the BM, and all three structures moved in phase.

Motion Responses Along the Human CP

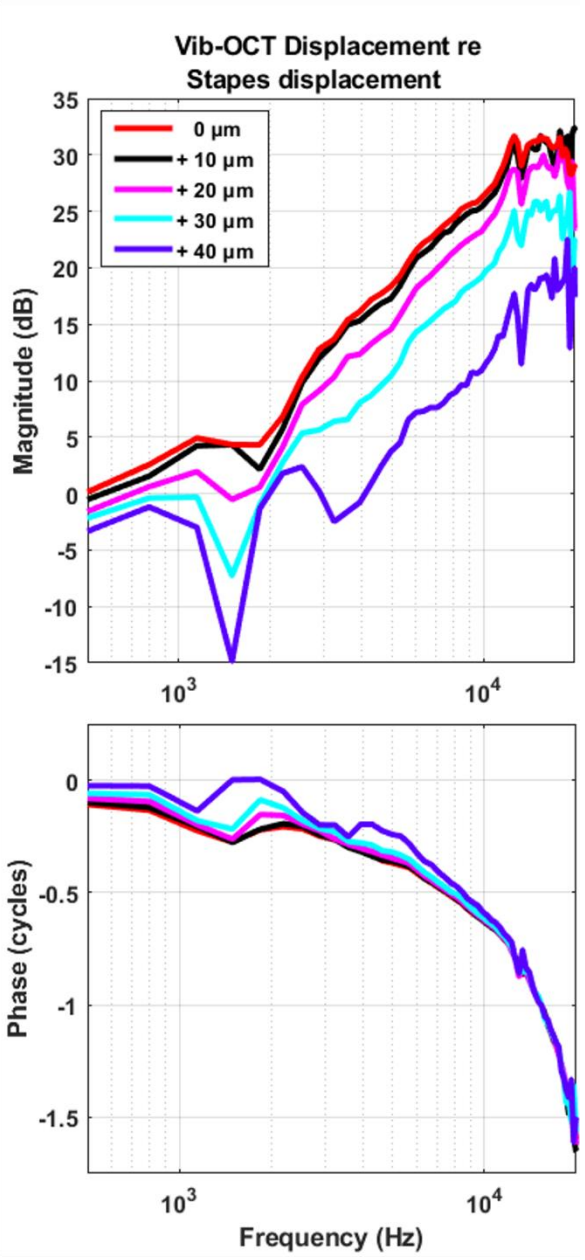
Figure 3 shows vibratory displacements of regions along the length of the CP referenced to stapes displacement (see Fig. 3a for key to locations). Moving laterally from the OSL toward the BM (-120 to 0 μm locations), motion increased with peak displacement at the BM near the junction of the BM-bridge region (Figure 3b). Notably, motion was clearly demonstrated throughout the OSL, even at regions >100 um away from the BM. Moving further laterally towards the SL (0 to +40 μm), displacement then decreased (Figure 3c). All of the indicated regions had maximum displacement at approximately 12-15 kHz. Structures along the BM moved approximately in phase with each other, but there were some phase irregularities at the high frequencies.



(a)



(b)



(c)

Figure 3. (a) B-scan image of the cochlear base in a very fresh (13 hours post mortem) human temporal bone. Colored lines indicate approximate radial locations where vibrometry data was collected. (b) Motion responses of regions along the surface of the CP bordering scala tympani from the OSL

(green) to BM near the BM-bridge region (red), normalized to stapes velocity as measured by LDV. (c) Motion responses along the surface of the CP bordering scala tympani from BM near the BM-bridge region (red) towards the SL (purple).

DISCUSSION

Imaging of the Human CP Using OCT

Previous OCT images of the human cochlea are of lower resolution (due to a longer OCT laser wavelength) where it is difficult to identify many structures within the CP (14, 15). With the shorter wavelength in our system, we are able to clearly identify the BM, OSL, bridge region, and SL. OCT vibrometry also allows for the measurement of motion of specific structures *in situ*. Histologic correlate is a helpful guide to interpreting OCT images, but as Figure 1 demonstrates, the morphology of some structures in histological slides can differ from fresh intact cochlea visualized with OCT.

Motion of the Human CP Measured Using OCT-Vibrometry

To date, these are the first measurements in human investigating the relative motion of structures within the CP with OCT vibrometry. We observe substantial motion of structures over the length of the CP, including the OSL, with peak displacement occurring at approximately the junction of the BM and bridge region. These findings are consistent with published findings in LDV (3). Moreover, different regions of the CP appear to move mostly in phase.

The very fresh specimen utilized for all of the measurements herein exhibited frequency tuning, with maximum displacement between 12-15 kHz, consistent with model calculations with a similar best frequency of about 10 kHz (16). Previously published studies using OCT in laboratory animals have shown sharp frequency tuning in live animals which broadens shortly after death (17, 18). However, studies using LDV in cadaveric human specimens have indicated relatively sharp tuning, even >48 h post-mortem (3). These data suggest that the mechanics responsible for passive tuning may differ across species.

ACKNOWLEDGEMENTS

Supported by NIH/NIDCD R01 DC013303 and NIH/NIDCD - T32DC000020-29A1.

REFERENCES

1. Rhode WS, Recio A. Study of mechanical motions in the basal region of the chinchilla cochlea. J Acoust Soc Am. 2000;107(6):3317-32.
2. Narayan SS, Temchin AN, Recio A, Ruggero MA. Frequency tuning of basilar membrane and auditory nerve fibers in the same cochleae. Science. 1998;282(5395):1882-4.
3. Raufer S, Guinan JJ, Jr., Nakajima HH. Cochlear partition anatomy and motion in humans differ from the classic view of mammals. Proc Natl Acad Sci U S A. 2019;116(28):13977-82.
4. Sumner CJ, Wells TT, Bergevin C, Sollini J, Kreft HA, Palmer AR, et al. Mammalian behavior and physiology converge to confirm sharper cochlear tuning in humans. Proc Natl Acad Sci U S A. 2018;115(44):11322-6.
5. Shera CA, Guinan JJ, Jr., Oxenham AJ. Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. Proc Natl Acad Sci U S A. 2002;99(5):3318-23.
6. Raufer S, Verhulst S. Otoacoustic emission estimates of human basilar membrane impulse response duration and cochlear filter tuning. Hear Res. 2016;342:150-60.
7. Recio-Spinoso A, Oghalai JS. Mechanical tuning and amplification within the apex of the guinea pig cochlea. J Physiol. 2017;595(13):4549-61.
8. Lee HY, Raphael PD, Park J, Ellerbee AK, Applegate BE, Oghalai JS. Noninvasive in vivo imaging reveals differences between tectorial membrane and basilar membrane traveling waves in the mouse cochlea. Proc Natl Acad Sci U S A. 2015;112(10):3128-33.

9. Gao SS, Raphael PD, Wang R, Park J, Xia A, Applegate BE, et al. In vivo vibrometry inside the apex of the mouse cochlea using spectral domain optical coherence tomography. Biomed Opt Express. 2013;4(2):230-40.
10. Dong W, Xia A, Raphael PD, Puria S, Applegate B, Oghalai JS. Organ of Corti vibration within the intact gerbil cochlea measured by volumetric optical coherence tomography and vibrometry. J Neurophysiol. 2018;120(6):2847-57.
11. Cho NH, Wang H, Puria S. Cochlear Fluid Spaces and Structures of the Gerbil High-Frequency Region Measured Using Optical Coherence Tomography (OCT). J Assoc Res Otolaryngol. 2022;23(2):195-211.
12. Nakajima HH, Dong W, Olson ES, Merchant SN, Ravicz ME, Rosowski JJ. Differential intracochlear sound pressure measurements in normal human temporal bones. J Assoc Res Otolaryngol. 2009;10(1):23-36.
13. Nadol JB, Jr. Techniques for human temporal bone removal: information for the scientific community. Otolaryngol Head Neck Surg. 1996;115(4):298-305.
14. Starovoyt A, Putzeys T, Wouters J, Verhaert N. High-resolution Imaging of the Human Cochlea through the Round Window by means of Optical Coherence Tomography. Sci Rep. 2019;9(1):14271.
15. Iyer JS, Yin B, Stankovic KM, Tearney GJ. Endomicroscopy of the human cochlea using a micro-optical coherence tomography catheter. Sci Rep. 2021;11(1):17932.
16. Yoon YJ, Steele CR, Puria S. Feed-forward and feed-backward amplification model from cochlear cytoarchitecture: an interspecies comparison. Biophys J. 2011;100(1):1-10.
17. Fallah E, Strimbu CE, Olson ES. Nonlinearity of intracochlear motion and local cochlear microphonic: Comparison between guinea pig and gerbil. Hear Res. 2021;405:108234.
18. Strimbu CE, Wang Y, Olson ES. Manipulation of the Endocochlear Potential Reveals Two Distinct Types of Cochlear Nonlinearity. Biophys J. 2020;119(10):2087-101.

Differential Transverse Motion of Outer Hair Cells Measured in Gerbil High-Frequency Region

Sunil Puria[1,2,3,a)], Nam Hyun Cho[1,2] and John Guinan[1,2]

[1]*Eaton Peabody Laboratories, Massachusetts Eye and Ear, Boston, MA 02114*
[2]*Harvard Medical School, Boston, MA*
[3]*Speech and Hearing & Bioscience and Technology Graduate Program at Harvard University*

[a)]Corresponding author: sunil_puria@meei.harvard.edu

Abstract. The great sensitivity and frequency selectivity of mammalian hearing originates in the mechanical properties of the cochlea. Cochlear motions in response to sound are amplified using metabolic energy. The motor element of this cochlear amplification is the outer hair cell (OHC), that expands and contracts lengthwise in response to a change in receptor potential at audio frequencies. How thousands of OHCs, acting through the local cytoarchitecture, work to achieve cochlear amplification is not fully understood. The properties of individual OHCs measured in micro-chamber environments are low-pass with a corner frequency of, at most, a few kHz. OHC corner frequencies within the cochlea have been estimated to be ~3kHz. Since in-vivo cochlear motion measurements show amplified frequency responses that extend up to many tens of kHz, but OHCs have corner frequencies of a few kHz, it has been suggested that OHCs are not active at the highest frequencies. How to reconcile these measurements has been a challenge. To address this, we made transverse cochlear motion measurements using a high-resolution optical-coherence-tomography (OCT) system at approximately the two ends of the OHC axial length: at the OHC bottom near the OHC-Deiter-Cell junction and at the OHC top near the reticular lamina. From these in-vivo measurements near the 45 kHz best frequency (BF) region of the gerbil cochlea, we determined the transverse differential response of OHCs. To remove phase from the traveling wave, we normalized the differential response by the motion at the OHC-Deiter-Cell junction. The normalized OHC differential response showed an almost exactly ½-cycle OHC top-to-bottom phase difference at frequencies up to 30 kHz, which is ½-octave below the BF (above this frequency the traveling wave is amplified). As frequency increased above 30 kHz, the phase difference decreased from ½ cycle, until near BF there was very little phase difference. From the lowest frequencies to BF the OHCs had substantial motion. Thus, OHCs continued to be active at frequencies up to BFs of 45 kHz, more than an order of magnitude above their low-pass corner frequency.

INTRODUCTION

It well-established that outer-hair cells (OHC) have piezo-electric like electromotile properties[1] thought to generate forces up to 80 kHz[2] and pump power into cochlea over the wide frequency range found in high-frequency mammals such as mice[3]. Most measurements demonstrating OHC electro-motility have been from isolated OHCs in micro-chamber like environments. These measurements indicate a resistance-capacitance (RC) time constant "problem" that would limit electromotility to a few kHz[4, 5]. Theoretical work indicates that OHCs must be analyzed with appropriate loading for them to do work at high frequencies[6, 7]. Recent in vivo measurements from the gerbil ~25 kHz best frequency (BF) region suggests that the OHC current-to-voltage conversion is low-pass with a corner frequency of about 3 kHz[8, 9]. This raises questions about the ability of OHCs to do work on the basilar membrane (BM) and reticular lamina (RL) at the highest frequencies in vivo. These gerbil studies did not report measurements from which OHC top-to-bottom differential motion could be determined. Measurements in the 9 kHz region of the mouse found nearly out-of-phase OHC top-to-bottom differential motion that, based on motion from harmonic distortion, extended to 20 kHz[10].

We made measurement from the gerbil ~45 kHz BF region, with an approach that is more transverse to the organ of Corti (OoC) than previously reported measurements. This approach allowed transverse measurements at

approximately the two ends of OHC axial length, i.e. (1) at the bottom end near the OHC-Deiter-cell (DC) junction (OHC-DC junction) and (2) at the top end near the reticular lamina (RL). The high-resolution capability of our optical-coherence-tomography (OCT) system allowed measurements in the same preparation across individual rows of OHCs.

METHODS

Displacement measurements were made through the gerbil round-window membrane (RMW) at several OoC locations including the RL and the OHC-DC junction (Fig. 1) at approximately constant ear-canal input pressure levels. The BFs were in the 40-48 kHz range. The OCT system, sound stimulus system, animal preparation and other measurement methods are described elsewhere[11]. Here we concentrate on analyzing the differential displacement between the ends of OHCs approximately along their axial length from an intact healthy cochlea in vivo, and after death. The animal care and experimental methods were approved by the MEE institutional review board.

The measured RL displacement (D_{RL}), and the OHC-DC junction displacement ($D_{OHC\text{-}DC}$), were normalized by ear canal sound pressure (Pec) to obtain the gain in mm/Pa. From these we calculated the local OHC-differential gain (D_{RL} - $D_{OHC\text{-}DC}$)/Pec for each input level. Measurements of BM motion showed a traveling wave with a phase that accumulated by several cycles with frequency. Since the OoC motion follows that of the BM, all of our OoC measurements have a traveling-wave phase accumulation. To remove the traveling-wave phase, we normalized the differential gain by the $D_{OHC\text{-}DC}$ motion, which yielded the normalized OHC-differential response (D_{RL} - $D_{OHC\text{-}DC}$) / $D_{OHC\text{-}DC}$, which is dimensionless. We report measurements from individual OHC rows 3, 2 and 1 (Fig. 1D), color coded in Figs. 2-4 as red, green, and blue, respectively.

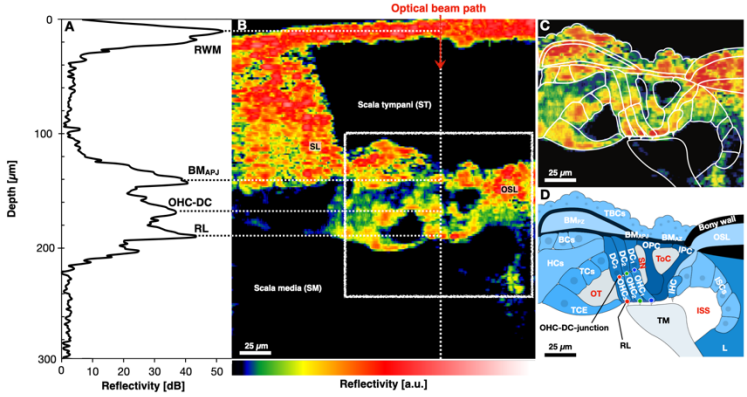


FIGURE 1. The anatomy of the cochlea from gerbil G614 shown by in vivo OCT imaging. **A**) An example of the depth profile (backscattered-light reflectivity) of a single 1D "A-scan" (white dotted line in B) shows several peaks corresponding to: the round-window membrane (RWM); the arcuate-pectinate junction of the BM (BMAPJ); the outer hair cell (OHC) Deiters' cell (DC) junction (OHC-DC); and the reticular lamina (RL). **B**) An in vivo cross-sectional OCT B-scan image, obtained through the RWM. Color indicates strength of reflectivity (in arbitrary units). **C**) Enlarged view of white box in (B) detailing organ of Corti (OoC) structures with overlayed outline drawings. **D**) A labeled cross-sectional drawing (from C) of representative OoC structures with the three rows of OHCs. Other abbreviations: TBCs: tympanic border cells; BCs: Boettcher cells; OSL: osseous spiral lamina; HCs: Hensen's cells; TCs: tectal cells; TCE: tectal-cell extension; ISCs: inner-sulcus cells; TM: tectorial membrane; L: limbus; OT: outer tunnel; SN: Space of Nuel; ToC: tunnel of Corti; ISS: inner spiral sulcus.

RESULTS

Figure 2 shows the motion gains of the three OHC rows near the RL (left column) and the OHC-DC junction (right column) measured for sound pressure levels of ~70, ~80, and ~90 dB SPL. Shown are both in vivo (darker colors) and postmortem (faded colors) measurements. As reported previously[11], near BF there was significant gain (relative to postmortem) for RL motion but not for OHC-DC motion. At low frequencies (10-30 kHz) there was some gain in RL motion but not nearly as much as in OHC-DC motion (Fig. 2A, C). In addition, RL3 and RL2 gain magnitudes were

a few dB greater than those of RL1 (Fig. 2A). All of the motion measurements showed the classic traveling wave phase accumulation of 2-3 cycles at the BF of 41.2 kHz, similar to previous reports[12-14]. There was a peak near BF in the OHC-DC response (Fig. 2C).

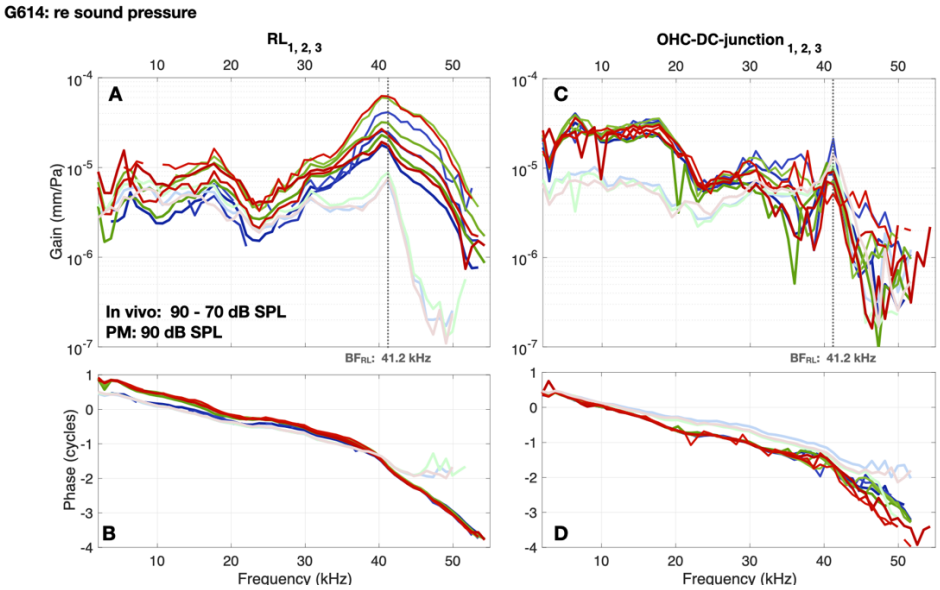


FIGURE 2. The in vivo and postmortem (PM) gains at the two OHC top and bottom ends and across the three OHC rows. Magnitudes (**A, C**) and corresponding phases (**B, D**) of the motions of the RL at the OHC tops (**A, B**) and of the OHC-DC-junction at the OHC bottoms (**C, D**). Data from OHCs 3, 2 and 1 are red, green and blue, respectively, with in vivo as dark colors and PM as faded colors. Thicker lines correspond to higher SPLs. Animal G614.

Using the gains in Fig. 2, the differences in the motions from the OHC tops to the OHC bottoms were calculated (using complex-number differences) for each OHC row (Fig. 3). Generally, there was substantial OHC-top-to-bottom-differential gain from ~5 kHz all the way to BF and even beyond. Near BF, the lower levels (thinner, lighter lines) had more gain than higher levels (thicker, darker lines) indicating gain compression. In the low-frequency region (10-30 kHz), there was much less compression. There was a slight dip the OHC-differential gain at about 25 kHz. Postmortem, the OHC-differential gain was typically reduced, as expected. The phases indicated that a traveling wave remained in the OHC-top-to-bottom-differential motions.

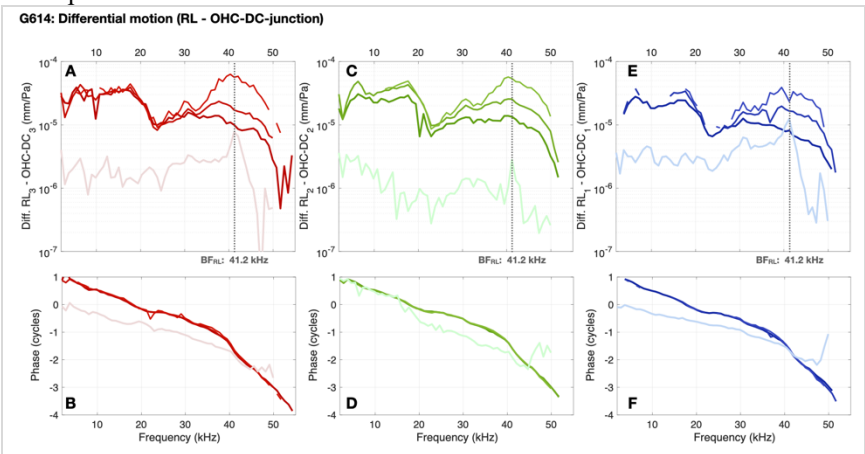


FIGURE 3. OHC top-to-bottom differential gain for OHC row 3 (A, B), row 2 (C, D), and row 1 (E, F). The OHC differential gain is the gain of the RL at the top of the OHC minus the gain of the OHC-DC junction at the bottom of the OHC.

To remove the travelling-wave phase, the differential motions of Fig. 3 were normalized by the corresponding $D_{OHC\text{-}DC}$ motions, done separately for each input level (Fig. 4). The normalized OHC top-to-bottom-differential responses increased gradually from ~0 dB below 10 kHz to ~6 dB near 30 kHz. Above 30 kHz, which is about ½ octave below BF, the normalized-OHC-differential response was more variable but increased to nearly 20-30 dB. There was a dip near BF, perhaps related to the peak in OHC-DC junction response.

Below 30 kHz, the normalized-OHC-differential-response phase was very close to ½ cycle (Fig. 4B, D, F). Around BF, the phase was generally near zero and above BF there was a variable phase lag. For the postmortem condition the magnitude was typically less than 0 dB and the phase tended to be about -½-cycle for RL3 and RL1, but for RL2 was a bit messier.

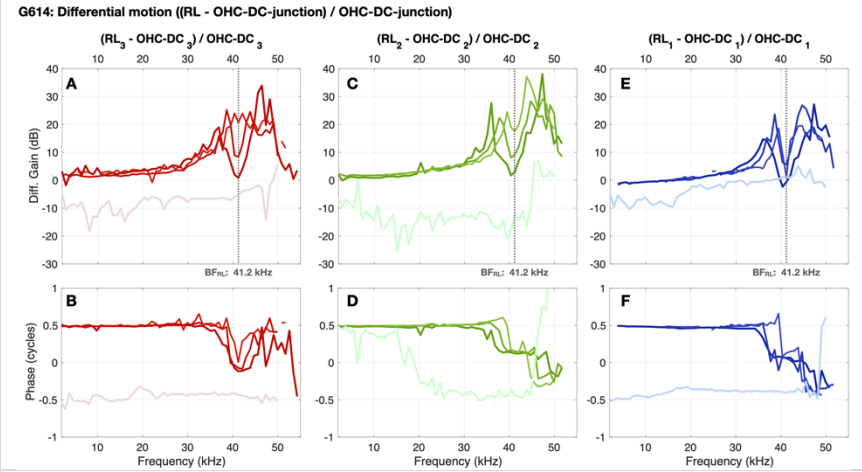


FIGURE 4. Same as Fig. 3 but for the normalized OHC-top-to-bottom differential motion.

DISCUSSION

The role of OHCs in endowing mammals with high sensitivity and fine frequency resolution has been debated since the inception of ideas about cochlear amplification[15, 16]. In the early days, it was not possible to make in vivo measurements with the spatial resolution needed to characterize motion of individual hair cells. Because of this, the idea of OHC motility was born in a micro-chamber environment and the concentration was on characterizing OHCs in this petri-dish-like milieu. Great progress has been made to understand the molecular and genetic aspects of the prestin motor molecules responsible for OHC motility[17]. One vexing and recurring "problem" resulting from these in vitro measurements and analysis has been the capacitance and resistance of the outer wall of OHCs that resulted in the so-called "RC time-constant problem". There have been numerous studies focused on this problem that suggested that the RC time constant would limit the effect of OHC force production to just a few kHz[8, 18]. Others argued that the isolated-OHC low-pass responses were not wrong but were fundamentally different from in vivo responses where OHCs are bathed in the fluid environment of the cochlea and act on the internal impedance of the OoC to produce traveling wave amplification[19]. Modern in vivo measurements using OCT allow us to look at the problem with a different perspective. We are now able measure OHC electromotility while keeping the cochlea intact and healthy and the OHCs working in their normal environment.

Recent gerbil measurements from 13-25 kHz BFs at the OHC-DC junction[14] used non-linear responses evoked by a multi-tone (zwuis-tone) stimulus to indicate that the OHC current-to-voltage conversion is low-pass with a corner frequency ~3 kHz. Although this corner frequency is several octaves lower than the BF at which it was measured, this study addressed the question: What is the OHC corner frequency when OHCs are working normally in the cochlea? It did not settle the question of whether OHC motility is evident above the corner frequency and acts to produce cochlear amplification at high frequencies. The question of whether OHC motility is still active above the OHC corner frequency was addressed by Dewey et al.[10] in the mouse 9 kHz region. They found out-of-phase differential OHC

top-to-bottom motion at frequencies ½ octave or more below BF, and at and above BF they found nearly out-of-phase motions. Here, we addressed this question in the 40-48 kHz region of the gerbil. We found almost exactly out-of-phase motion up to 30 kHz (Fig. 4), which is an order of magnitude above the OHC corner frequency reported for OHCs in the gerbil base[8,9]. As frequency increased above 30 kHz, there continued to be differential motion between the top and bottom of the OHCs (Figs. 3-4), but the phase difference between the OHC top and bottom decreased from ½ cycle, until near BF there was very little phase difference (Fig. 4, bottom). 30 kHz is ~½ octave below BF, which is the transition frequency between OHCs only amplifying local motions (such as the RL) and OHCs amplifying motion that adds to the traveling wave[13]. Although in the above-30-kHz region, the OHC top-to-bottom phase difference was no longer exactly ½ cycle, the OHCs continued to show substantial top-to-bottom motion differences (Figs. 3-4).

At frequencies where the traveling wave was amplified (30 kHz to BF), the OHC top-to-bottom differential motion was no longer exactly out-of-phase (Fig. 4), and it is interesting to consider why it was not exactly out-of-phase. Presumably, at these frequencies, the tissues that the OHCs push against no longer have impedances that are much less than the effective source impedance of the OHC, i.e., the OHCs are driving substantial loads. There is also more overall motion near the BF frequency, but it is not overall motion that is causing the OHC phase to be near zero at BF because the phase patterns seem to be independent of input level and are similar across the three OHC rows even though they have different amplitudes (Figs. 3-4). While there are no similar measurements of in vivo OHC forces, there are measurements of the extracellular voltage (a surrogate for the local OHC current), scala-tympani pressure, and BM displacement from the gerbil ~24 kHz BF region[13]. From these measurements it was surmised that the extracellular voltage and displacements are in phase at low frequencies (< ½-octave below BF), but ½-octave below BF there was a rapid phase change so that from ½-octave below BF up to BF the phase of the local OHC current led BM displacement by ¼ to ½ cycle[13]. Our measurements of the normalized OHC-top-to-bottom-differential motion (Fig. 4) have a somewhat similar pattern of constancy at low frequencies and a transition at ½-octave below BF that led to a phase change of ½ cycle or more as frequency increased. These patterns seem likely to be highly related, but we don't understand exactly how. One thing that stands out is their constancy at frequencies more than a ½ octave below BF and an abrupt change at ½ octave below BF.

We reported[11] significant differences in RL motion across the three OHC rows and wondered whether these differences were due to differences in OHCs properties across the three rows or were due to differences in their surrounding cytoarchitecture. The uniformity in the responses of the OHC-top-to-bottom-differential motions across the three rows (Fig. 4) suggests that the OHC properties across the rows are similar. This suggests that differences in the cytoarchitecture[20] produce the differences in the RL motion observed across the three rows of OHCs.

ACKNOWLEDGMENTS

We thank Andrew Tubelli for help with Figure 1. Work supported by grant R01 DC07910 from the NIDCD of NIH and the Amelia Peabody Charitable fund (both to SP).

REFERENCES

1. W. E. Brownell, C. R. Bader, D. Bertrand and Y. de Ribaupierre, Science **227** (4683), 194-196 (1985).
2. G. Frank, W. Hemmert and A. W. Gummer, Proceedings of the National Academy of Sciences of the United States of America **96** (8), 4420-4425 (1999).
3. Y. Wang, C. R. Steele, S. Puria and A. J. Ricci, Commun Biol **4** (1), 958 (2021).
4. J. Santos-Sacchi, J Neurosci **12** (5), 1906-1916 (1992).
5. J. Santos-Sacchi and W. Tan, J Neurosci **38** (24), 5495-5506 (2018).
6. R. D. Rabbitt, Proceedings of the National Academy of Sciences of the United States of America **117** (36), 21880-21888 (2020).
7. K. H. Iwasa, Sci Rep **7** (1), 12118 (2017).
8. A. Vavakou, N. P. Cooper and M. van der Heijden, Elife **8** (2019).
9. M. van der Heijden and A. Vavakou, Hear Res, 108367 (2021).
10. J. B. Dewey, A. Altoe, C. A. Shera, B. E. Applegate and J. S. Oghalai, Proc Natl Acad Sci U S A **118** (43) (2021).
11. N. H. Cho and S. Puria, bioRxiv, 2022.2003.2001.482580 (2022).

12. W. He, D. Kemp and T. Ren, Elife **7** (2018).
13. W. Dong and E. S. Olson, Biophys J **105** (4), 1067-1078 (2013).
14. N. P. Cooper, A. Vavakou and M. van der Heijden, Nat Commun **9** (1), 3054 (2018).
15. D. T. Kemp, J Acoust Soc Am **64** (5), 1386-1391 (1978).
16. D. T. Kemp, Hearing Res. **2**, 533-548 (1980).
17. M. C. Liberman, J. Gao, D. Z. He, X. Wu, S. Jia and J. Zuo, Nature **419** (6904), 300-304 (2002).
18. G. D. Housley and J. F. Ashmore, The Journal of physiology **448**, 73-98 (1992).
19. A. Altoè and C. A. Shera, bioRxiv, 2022.2002.2002.478769 (2022).
20. J. A. Soons, A. J. Ricci, C. R. Steele and S. Puria, J Assoc Res Otolaryngol **16** (1), 47-66 (2015).

Advective mass transport along the cochlear coil

Mohammad Shokrian[1, a)] Douglas Kelley[1, b)] and Jong-Hoon Nam[1, 2, c)]

[1]*Department of Mechanical Engineering, University of Rochester*
[2] *Department of Biomedical Engineering, University of Rochester*

[a)]mshokria@ur.rochester.edu
[c)]d.h.kelley@rochester.edu
[c)] Corresponding author: jong-hoon.nam@rochester.edu

Abstract. Mammalian auditory epithelium (the organ of Corti) stands out among different inner-ear epithelia in that it has large extracellular fluid spaces such as the tunnel of Corti, the Nuel's space, the outer tunnel, and the spacing between the outer hair cells. We tested the hypothesis that advective flow facilitates mass transport in the cochlear fluids, using computational simulations of cochlear fluid dynamics and *ex vivo* experiments to investigate mass transport in extracellular fluid spaces of the cochlea. Three model simulations were performed in series—cochlear mechanics, nonlinear fluid dynamics, and mass transport. In nonlinear fluid dynamics, we incorporated the convection terms for more accurate computation of drift flow. For mass transport, both diffusion and advection were considered. For experiments, we measured vibrations of excised cochlear turns using optical coherence tomography. The excised OoC was subjected to acoustic and electrical stimulations.

INTRODUCTION

The organ of Corti (OoC), the sensory epithelium in the mammalian cochlea, includes substantial extracellular spaces such as the tunnel of Corti, outer tunnel, and Nuel's space. Fluid in these spaces is interconnected by spacings between the columns of outer pillar cells and outer hair cells. Only mammalian species has such a large fraction of extracellular fluid volume within the auditory epithelium. Chemical and electrical properties of this Corti fluid are similar to yet distinct from the lymph in the scala tympani [1, 2]. The role of Corti fluid has hardly been studied perhaps because it was difficult to access in the past.

The OoC transmits basilar membrane vibrations to the hair bundles. For such kinematics, a rigid OoC is beneficial to convey vibration signals without compromising vibrations. On the other hand, the OoC is deformed by active outer hair cells. Outer hair cells contract or elongate in response to change in their transmembrane potential [3]. This electro-motility is required for cochlear amplification [4-6]. Investigations have been performed to explain how electro-motility amplifies OoC vibrations, both experimentally [7-13] and theoretically [14-24]. Recent observations showed apparent outer-hair-cell motility well below the characteristic frequency of measurement location [13, 25-29]. This led to diverse interpretations. Some believe that cochlear amplification is not local—outer-hair-cell energy accumulates over the length of the cochlea for amplification [25, 26], while others maintained that cochlear amplification is locally achieved [12, 28]. Another explanation was that outer-hair-cell motility is relevant to encoding the envelope of sound signals [30].

One way to summarize recent observations is to describe them as a substantial change in OoC cross-sectional area well below CF. Interestingly, these observations when combined with classical observations on cochlear traveling waves (wavelength longer than transvers dimensions and propagation speed on the order of m/s) form an

ideal condition for longitudinal fluid motion [31]. Indeed, longitudinal flux of Corti fluid has been observed *ex vivo*. Karavitaki and Mountain reported longitudinal fluid flow along the tunnel of Corti (based on the motion of nerve fibers crossing the tunnel) caused by outer-hair-cell motility [32]. Although the Corti fluid has hardly been considered in cochlear physics studies, a couple of theoretical studies suggested a role of the Corti fluid tube in longitudinal power transmission along the cochlea [33, 34]

Using computational models, this study examines whether fluid advection in the Corti fluid can contribute to mass transport in the cochlea. According to current understanding, diffusion is the only mechanism for mass transport in the inner ear fluids [35]. Diffusion is highly efficient mechanism to spread molecules in sub-micrometer space such as the gap between cells and neurons. However, diffusion is highly inefficient for mass transport along millimeter-long tube. For a long tube-bundle filled with fluid, like the mammalian cochlea, longitudinal mass transport can be a viable mechanism for maintaining fluid homeostasis.

METHODS

Modeling

Three dynamics were solved in sequence. First, cochlear mechanics, interactions between OoC complex, scala fluids, and hair cells, are solved. Second, nonlinear fluid dynamics to solve for steady-streaming. Third, mass transport (advection and diffusion) in the scala tympani and the Corti tube was computed. The results of the cochlear mechanics analysis (OoC vibrations) were used to compute steady-streaming. The velocity field of the stream was used to compute spatiotemporal distribution of a substance in the fluid. Model components were formulated using finite element (FE) approach, where variables such as the displacement/velocity, pressure, voltage, and

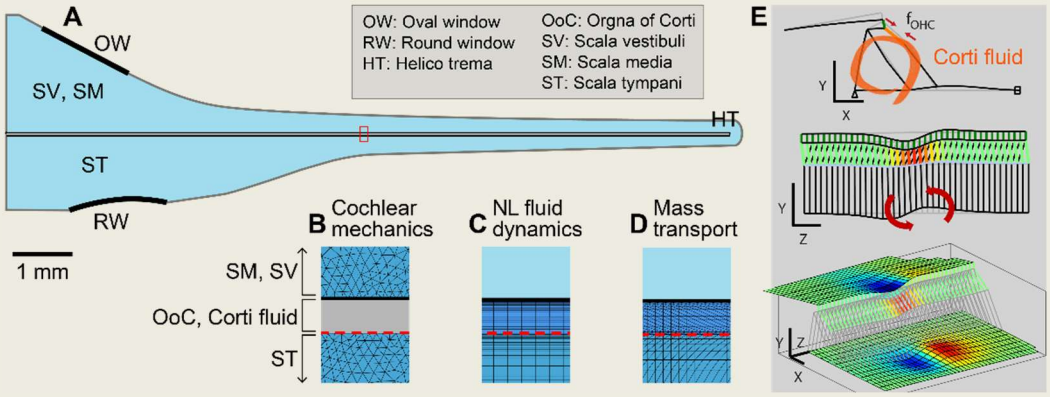


FIGURE 1. Model to analyze mass transport along the cochlea. (**A**) Scala fluid spaces of the gerbil cochlea. For the purpose of cochlear fluid mechanical analysis, the cochlea was 'uncoiled', and the SV and SM were merged. While 3-D fluid spaces were reduced into 2-D, the cross-sectional area change along the cochlear length was represented by the change of the scala height. (**B**, **C**, **D**) Mass transport along the scala tympani was simulated in three steps. As each step solved different physics problem, relevant mesh grids were employed. These 100 µm-wide domains in (B-D) correspond to the red rectangle in (A). (**B**) In cochlear mechanics model, the interaction between inviscid fluid and elastic partition (OoC complex) is solved. The vibrations of the basilar membrane and the reticular lamina were used for the next step. (**C**) Using nonlinear fluid dynamics model, the motions of the Corti fluid and the ST fluid were solved. The fluids were considered viscous. The permeability of the basilar membrane (red broken line) was incorporated. The nonlinear (convection) term of the Navier-Stokes equation was considered to solve for steady-streaming (drift flow). The drift velocity field was used for the next step. (**D**) Advection-diffusion equation was used to simulate mass transport from the RW along the length of the ST. Note that the Corti fluid was not explicitly considered in cochlear mechanics (B), while the SV/SM fluid and OoC structure were not considered during the analyses of nonlinear fluid dynamics (C) and mass transport (D). (**E**) Detailed structural model used for cochlear mechanics. Three-dimensional FE model incorporated the detailed geometry, and hair cell physiology of the gerbil cochlea. Illustrated is the 3-D deformation of OoC complex when outer hair cells in the middle of the longitudinal span were contracted (adapted from [21]).

concentration were defined at nodes and the physical connectivity between the nodes were defined as elements. Model components related to cochlear mechanics have been published such as the OoC complex structure and electric circuit [21], and fluid-structure interaction [24, 36]. In the following, three new developments are described. One is the permeability of the basilar membrane. The other is nonlinear fluid dynamics including convection. Another is regarding mass transport (advection-diffusion) in the Corti fluid and the scalar fluid.

Nonlinear fluid dynamics including convection: We focused on the steady-streaming in the Corti fluid. As the Corti fluid is connected to the fluid in scala tympani, both fluid-spaces were solved while the other fluid spaces (scala media and vestibuli) were not considered. For simplicity, the analyzed fluid-spaces (scala tympani and Corti fluid) were modeled in 2-D. The change in the OoC cross-sectional area was represented by the relative motions between the reticular lamina and the basilar membrane. The height of the scala tympany changed along its length to reflect the change in its cross-sectional area. The velocity ($\mathbf{u}$), vorticity (ω) and the pressure (p) of the fluid were obtained by solving the Navier-Stokes and the continuity equations,

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\frac{1}{\rho}\nabla p + \nu \nabla \times \omega, \quad (1)$$

$$\omega - \nabla \times \mathbf{u} = 0, \text{ and} \quad (2)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (3)$$

where ν is the kinematic viscosity (0.7×10^{-6} m^2/s), and ρ is mass density of fluid (1000 kg/m^3). As we were interested in the accuracy of longitudinal drift flow, both the convective and viscous terms were considered. The Reynolds number ($Re = ch/\nu, c = f\lambda$) of the Corti fluid motion supported the inclusion of these terms. For instance, traveling wave at 1 kHz with an amplitude of 25 nm and a wavelength of 4 mm would result in Re > 100 [31]. Nonlinear fluid dynamics was solved using the Newton's method.

For the scala tympani fluid, the bottom boundary represented the scala wall and the top boundary represented the permeable basilar membrane shared with the Corti fluid. The kinematic boundary condition of these boundaries were from the fluid-structure interaction simulation of the cochlear mechanics model [24, 36]. The mesh for the fluid dynamics model consists of approximately 40,000 2nd order quadrilateral elements with the same degrees of freedom for both velocity and pressure. The boundary layer thickness was chosen based on an educated guess from the Stokes' problem ($\delta \approx \sqrt{\nu/\omega}$). The time integration was carried out using the 2nd order Crank-Nicolson method. The least-squares finite element approach was used with a reduced-order integration to ensure stability. The time step size of 2 μs was used after considering computational time and numerical accuracy to compute drift motion of fluid particles. The time-span for nonlinear dynamic analysis was chosen to simulate five to ten repetitions of a stimulation peep (10-20 ms, Fig. 2A). After this time-span, the flow field formed a stable pattern of steady-streaming. The trajectory of fluid particles c, $\mathbf{x}_c(t)$, were obtained by integrating Eulerian velocity $\mathbf{u}_\mathrm{E}$ over an oscillatory period *T*.

$$\mathbf{x}_c(t+T) = \mathbf{x}_c(t) + \int_t^{t+T} \mathbf{u}_\mathrm{E}(\mathbf{x}_c, \tau) d\tau. \quad (4)$$

The drift of fluid particles was represented by the velocity $\mathbf{u}_\mathrm{D}$.

$$\mathbf{u}_\mathrm{D}\big(\mathbf{x}_c(t)\big) = \big(\mathbf{x}_c(t+T) - \mathbf{x}_c(t-T)\big)/2T. \quad (5)$$

Permeable basilar membrane: The basilar membrane is modeled using the Darcy's law after considering it as a fully saturated porous medium. According to the Darcy's law, the flux per unit area of porous medium (q) is proportional to the pressure gradient.

$$q = -K\nabla p, \quad (6)$$

where q is the flux in $\mathrm{m/s}$ and K has the unit of $\mathrm{m/Pa.s}$. The value K of 0.01 was chosen empirically. The nodal velocity values obtained from the cochlear mechanics model were prescribed at the top and bottom of the Corti fluid space, represented by the reticular lamina and basilar membrane respectively. Although there is an assigned value of velocity on the basilar membrane, the fluid velocity at this boundary remains free. The fluid velocity (u_f) at the basilar membrane was evaluated from

$$u_f = u_p + u_m, \quad (7)$$

where u_p is the prescribed velocity of the membrane and u_m is the velocity of the fluid that is permeating through the membrane as a result of pressure difference.

Mass transport was computed using the advection-diffusion equation, where concentration C was determined by diffusion coefficient D and drift velocity $\mathbf{u}_\mathrm{D}$.

$$\partial C/\partial t = D\nabla^2 C - \mathbf{u}_\mathrm{D} \cdot \nabla C \quad (8)$$

The time scale of mass transport is different from the fluid dynamics. The time-step of the problem varied adaptively, considering that the concentration gradient decreased with time. The initial time-step was chosen equal to the stimulation period. The element length was 1-2μm which resulted in approximately 340,000 first-order triangular elements. The model parameters are summarized in Table 1.

Experiment

The cochleae were isolated from young gerbil (15-30 day old). After being reduced to a single turn, the excised cochlea, between 50 and 80-percentile location from the basal end, was placed in a micro-chamber. The mechanical stimulation was delivered through an opening of the chamber sealed with elastomer, which was in contact with a piezoelectric actuator. The other opening of the bottom channel served as a pressure release (analogous to the function of the round window in natural cochlea). A pair of inlet-outlet ports were used to refresh the fluid in the bottom chamber. After placing and sealing of the reduced cochlear turn on the chamber slit, the top fluid was replaced with an endolymph-like (145 KCl, 0.1 CaCl2, 4 HEDTA, 10 K-HEPES, 8 Glucose, 2 Na-Pyruvate, in mM, pH 7.35±3 and 300±3 mOsm) solution. A pair of electrodes were placed below and above the excised cochlear turn. Mechanical and electrical stimulations were applied to the tissue at different frequencies to evoke the passive (mechanically agitated) and active (electrically induced) vibrations. Resulting vibrations were measured using an optical coherence tomography (OCT) system [38]. Typical level of stimulation was 1 Pa and 0.1 mA for respective stimulation methods. Typical vibration amplitude gain at the reticular lamina was 50-200 nm/Pa and 100-300 nm/mA, respectively. Vibrations were measured by running 40-60 M-scans across the OoC span.

RESULTS

Mass transport by active streaming along the cochlea

A hypothetical case of inner ear drug delivery through the round window was simulated (Fig. 2). A high concentration solution of substance (drug) was applied to the round window. Effect concentration at the target cite (OoC) was assumed 1 percent of the applied concentration. For the cochlear mechanics part, the stapes velocity was the input and the vibrations of OoC complex were considered the output (Fig. 2A and B). To simulate acoustic agitations, a random signal was band-pass filtered between 2 and 10 kHz. The filtered signal was enveloped in a 2-ms pulse with a rise- and fall-time of 0.1 ms. The level of the agitation was chosen to be equivalent to 70 dB SPL (sound pressure level). Steady streaming (non-oscillatory drift flow) of the Corti fluid and scala fluid was computed by solving the Navier-Stokes equation including the convection term (Eqs. 1-3). For this part of simulations, the kinematic boundary motions (the vibrations at the basilar membrane and at the reticular lamina) were the input and the velocity field of fluid was the output (Fig. 2C). Finally, the advection-diffusion equation (Eq. 8) was solved. The fluid velocity field (only the steady-streaming component) and the concentration at the round window were the input and the spatiotemporal distribution of the diffusate was the output (Fig. 2D).

Substantial drift flow on the order of 0.1 mm/s was generated by the sound stimulation equivalent to 70 dB SPL (Fig. 2C). The Péclet number (Pe), representing the relative contribution of advection versus diffusion in mass transport, might provide the sense of how substantial this velocity is. The drift velocity of u_D = 0.1 mm/s corresponds to Pe > 10 (Pe = $u_D L/D$), after considering a diffusion coefficient of D = 0.7 μm^2/ms and a reasonable

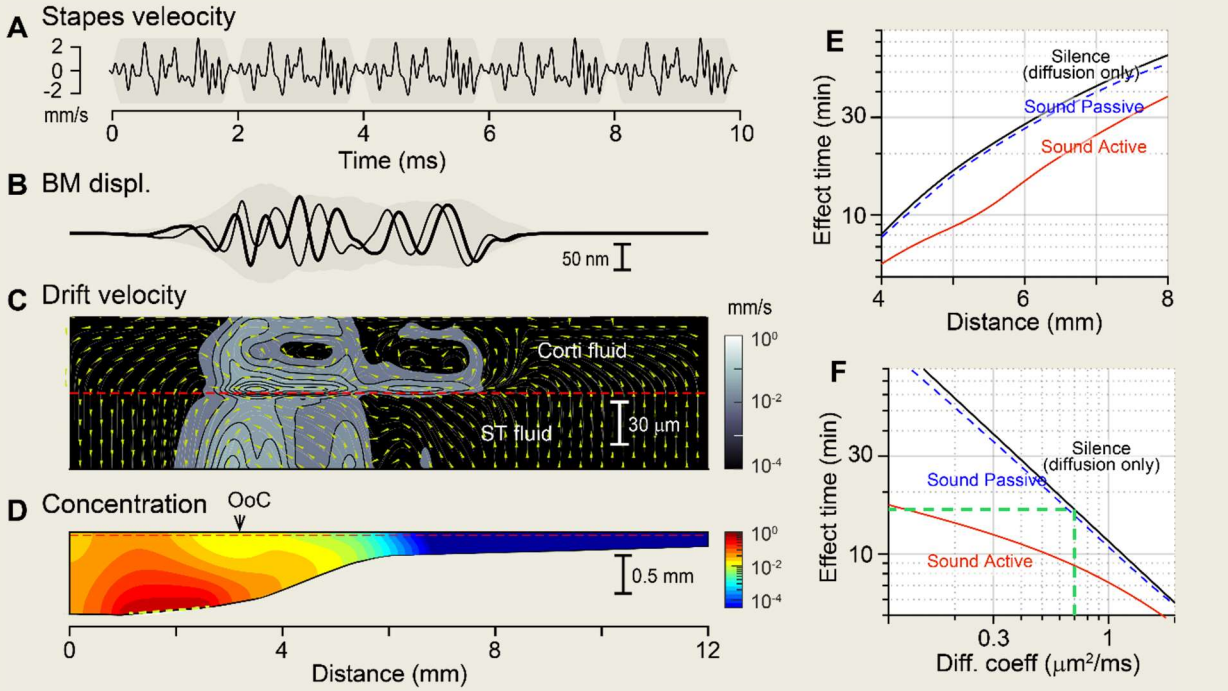


FIGURE 2. Simulated drug delivery. (**A**) Acoustic stimulation at the stapes. Two-ms-long noise patterns were repeated. The simulation is equivalent to 70 dB SPL sounds of which frequency range is banded between 2 and 10 kHz. (**B**) Cochlear mechanics model simulated for the vibrations of OoC complex. Two curves represent basilar membrane displacement pattern along the cochlear length at two moments of time. (**C**) Nonlinear fluid dynamics model computed the drift velocity field of the fluid over the entire cochlear length. Shown is the Corti fluid and the scalar tympani fluid near the basilar membrane. Permeable basilar membrane (red horizontal broken line) divides the two fluid-spaces. (**D**) Advection-diffusion equation was solved using the velocity field shown in (C). A drug with 100 times the effective concentration was applied to the round window. Shown contour plot represents the drug concentration after 10 min of delivery under sound. (**E**) Effect time represents the time when the drug concentration in the OoC reaches the effective concentration. When sounds were delivered but there was no outer hair cell motility, the effect time became very close to the case of pure diffusion (silence). (**F**) Effect time at the distance of 5 mm varies depending on the diffusion coefficient of the substance. For the simulation in (E), the diffusion coefficient of 0.7 μm^2/ms (broken line) was used.

characteristic dimension of L = 100 μm (the length scale of either the basilar membrane or the wavelength of traveling waves). On the other hand, the region with substantial drift was local—between 2 and 5 mm from the basal end and close vicinity to the basilar membrane. Notably, this region was more basal than the envelope of basilar membrane traveling waves (3-8 mm, Fig. 4B). Consistent with this spatial pattern of drift velocity, advection reduced the effect time in the location < 5 mm (the red curve has shallower slope than the black curve when the distance < 5 mm, Fig. 2E), but advection had modest effect for drug delivery in more apical locations (the slopes of the red and black curves are similar when the distance > 5 mm).

The effect time of drug delivery was defined as the time for substance concentration at a site of interest (the OoC) to reach 1% of the application concentration. Fig. 2E shows the effect time along the distance under three conditions: when prescribed acoustic stimuli were delivered to sensitive cochlea ('Sound Active'); when acoustic stimuli were delivered to insensitive cochlea ('Sound Passive'); when no acoustic stimuli were applied ('Silence'). The sensitive and insensitive cochlea correspond to the OoC complex model with and without active outer hair cell motility, respectively. The silence condition corresponds to mass transport by pure diffusion. It took more than an hour to reach 8 mm location by pure diffusion. Acoustic stimulation accelerated the drug delivery by a factor of 1.3-2 depending on locations. Fig. 2F shows how the effect time changes as the diffusion coefficient varies. The green broken lines represent the effect time as the distance of 5 mm when D = 0.7 $\mu m^2/ms$. As the value of diffusion coefficient increases, the effect time decreases and the difference in effect times between the advection-diffusion case and the diffusion-only case decreases, also.

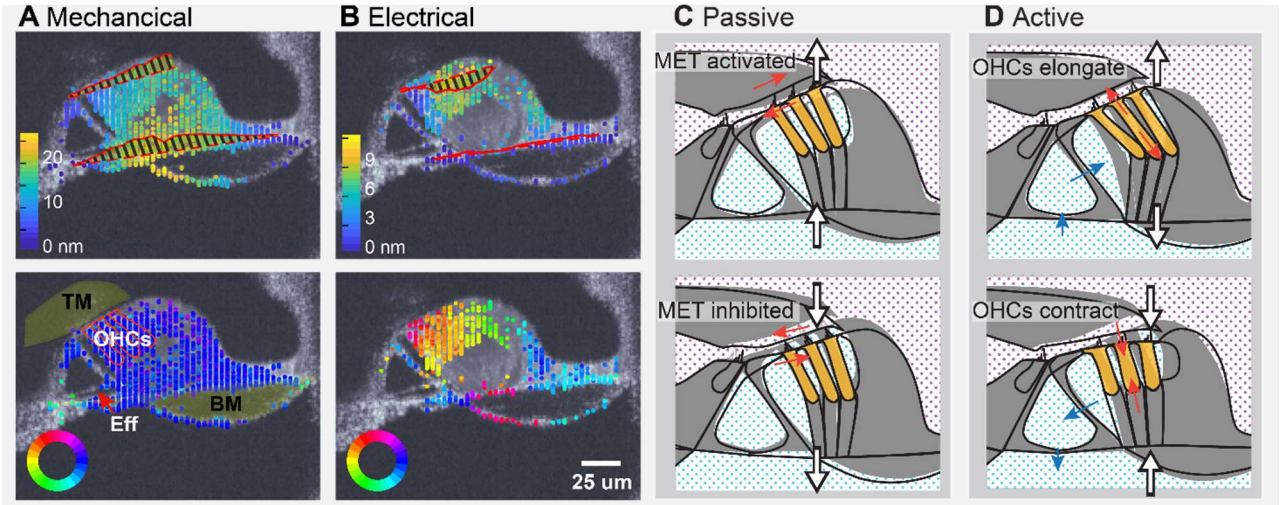


FIGURE 3. Motion of the organ of Corti and the Corti fluid. (**A, B**) The OoC in excised cochlea was subjected to sinusoidal stimulations either as alternating pressures ('Mechanical') or alternating trans-epithelial currents ('Electrical'). Top and bottom panels present vibration amplitude and phase, respectively. The background gray-scale images are from OCT B-scans across OoC cross-section at 8.5 mm from the basal end of the gerbil cochlea. Color dots indicate the data points acquired from OCT velocimetry. (**A**) When mechanically agitated, the entire OoC complex vibrates roughly in-phase. The maximum vibration took place between the roots of the outer pillar cell and the Deiters cells. (**B**) When electrically agitated, there was a phase variance across the OoC complex as large as 180 degrees. The greatest vibrations took place at and around the outer hair cells, reflecting their electro-motility. (**C, D**) The substantial volume fraction of the OoC is filled with extra-cellular fluid, called the Corti fluid. These cartoons exaggerate the deformation of the OoC under two conditions—when passive (insensitive or dead) cochlea is subjected to inter-scala pressures; and when active (sensitive) cochlea is subjected *only* to outer hair cell (OHC) motility. These conditions are comparable to the conditions of 'mechanical' and 'electrical' stimulation of our experiment, respectively. Gray areas and black lines represent non-deformed and deformed configurations, respectively. (**C**) Passive deformation due to hydrodynamic pressure. As the basilar membrane deflects upward/downward (thick white arrows), the stereocilia are deflected radially (red arrow pairs). (**D**) Deformation due to active outer-hair-cell motility. As the outer hair cells elongate/contract, the fluid is sucked into/extruded out of the Nuel's space (blue arrows). As a result, fluid will move longitudinally along the tunnel of Corti (in the x-direction). To measure these deforming patterns, we use an OCT system with high resolution (2 μm, Aim 2).

The origin of active streaming

Vibrating patterns of the OoC were clearly distinguished depending on stimulation method (Fig. 3). When mechanically stimulated (deformed by transepithelial pressure), the entire OoC complex vibrated roughly in phase, and the peak displacement occurred near the roots of Deiters' cells. When electrically stimulated, there was about 180 degree phase variance across the cross-section, and the displacement was greater in the top (reticular lamina) side than in the bottom (basilar membrane) side. Notably, the phase gradient was greater in the radial direction as compared to the transverse direction. The vibration pattern remained about the same despite different frequencies when mechanically stimulated, but the pattern varied depending on frequencies when electrically stimulated. When mechanically stimulated, the basilar membrane was deformed like a pulled bowstring, reflecting its medial and lateral attachment to the bony edges via ligaments. The reticular lamina was displaced as if a rigid plate (red curves in Fig. 3A). When electrically stimulated, however, the reticular lamina bended apparently.

The vibrating pattern due to outer-hair-cell motility showed the change of OoC cross-sectional area. In contrast, under passive mechanical deformation, the OoC vibrated like a rigid block sitting on top of the basilar membrane. Cartoons in Fig. 3C and D illustrate possible roles of different vibration patterns. Panel C is in line with the classical notion of OoC kinematics that the upward/downward motion of the OoC results in excitatory/inhibitory deflection of stereocilia bundles. In this model, deforming OoC will compromise the kinematic gain between the basilar membrane and the stereocilia bundle. It is not immediately clear whether the contraction of outer hair cells results in excitation or inhibition of hair cell mechanotransduction (panel D). On the other hand, it is apparent that outer-hair-cell motility will be an efficient means to change the cross-sectional area of the OoC. The displaced area will squeeze or aspirate the Corti fluid in and out of the section.

DISCUSSION

The inner ear labyrinth is subjected to adverse conditions for mass transport: stationary fluids are contained in long narrow tube. Cochlear fluid system is virtually isolated from other body fluid systems. Although the inner ear labyrinth is connected to the brain lymphatic space through narrow aqueducts, there exists negligible flux between the two fluid systems [39]. The blood-labyrinth barrier is equal or less permeable than the blood-brain barrier [40]. When it is admitted that the fluids in the cochlea are considered stationary, mass transport in the cochlear fluid must be limited by diffusion. This implies that for small molecules to travel a few millimeters along the cochlear it would take hours.

This study is not the first one that reports steady streaming within the inner ear labyrinth. Lighthill [41] predicted the existence of the steady streaming of the scala fluid through an analytical approach. Edom *et al*. [42] provided further quantitative details of the steady streaming, confirming Lighthill's theoretical prediction. More recently, Sumner *et al*. [43] investigated the potential of steady streaming in the cochlea as means of drug delivery mechanism. We have reported conditions for steady streaming within the Corti fluid (high phase velocity and long wavelength), and the physical basis of how the interaction between advection and diffusion results in efficient fluid mixing [31]. This study reflected more realistic physiological conditions as compared to our previous study in that: cochlear mechanics and scala fluids were explicitly incorporated; the fluid transport across porous basilar membrane was considered. Our study is different from others' previous studies in three aspects: the Corti fluid was considered as well as the scala fluid; ours is mass transport analysis considering both diffusion and advection; we suggest outer hair cells as an actuator for longitudinal mass transport. Our nonlinear fluid dynamics analysis obtained similar level of steady streaming as the previous studies. However, when considered together with diffusion, passive streaming due to basilar membrane vibrations hardly facilitated cochlear mass transport as it was slower than the transport by diffusion (Fig. 2E, F). Our results suggest that peristatic vibration of the Corti fluid tube can serve a mechanism that accelerates mass transport along the length of the cochlea.

To test our theory of active streaming in the cochlea, explicit experimental evidence may be required. A series of experiments regarding drug delivery where sounds and outer-hair-cell motility are controlled delicately may examine the proposed theory. There have been numerous experiments regarding drug or gene delivery into the cochlea. To our knowledge, few studies considered sounds and outer-hair-cell motility as experimental control factors.

ACKNOWLEDGMENTS

Wei-Ching Lin and Jonathan Becker participated in experiments. Drs. Kenneth Henry and Laurel Carney provided constructive thoughts regarding neurophysiological aspect of this study. The project was supported by NSF CMMI-1661413, NIH R01 DC014685, and SPIN award of Neuroscience Program at University Rochester Medical Center.

REFERENCES

1. Zidanic, M. and W.E. Brownell, *Fine structure of the intracochlear potential field. I. The silent current.* Biophys J, 1990. **57**(6): p. 1253-68.
2. Johnstone, B.M., R. Patuzzi, J. Syka, and E. Sykova, *Stimulus-related potassium changes in the organ of Corti of guinea-pig.* J Physiol, 1989. **408**: p. 77-92.
3. Brownell, W.E., C.R. Bader, D. Bertrand, and Y. de Ribaupierre, *Evoked mechanical responses of isolated cochlear outer hair cells.* Science, 1985. **227**(4683): p. 194-6.
4. Zheng, J., W. Shen, D.Z. He, K.B. Long, L.D. Madison, and P. Dallos, *Prestin is the motor protein of cochlear outer hair cells.* Nature, 2000. **405**(6783): p. 149-55.
5. Ashmore, J., *Cochlear outer hair cell motility.* Physiol Rev, 2008. **88**(1): p. 173-210.
6. Fettiplace, R. and K.X. Kim, *The physiology of mechanoelectrical transduction channels in hearing.* Physiol Rev, 2014. **94**(3): p. 951-86.
7. Chen, F., D. Zha, A. Fridberger, J. Zheng, N. Choudhury, S.L. Jacques, R.K. Wang, X. Shi, and A.L. Nuttall, *A differentially amplified motion in the ear for near-threshold sound detection.* Nat Neurosci, 2011. **14**(6): p. 770-4.
8. Ren, T., W. He, Y. Li, K. Grosh, and A. Fridberger, *Light-induced vibration in the hearing organ.* Sci Rep, 2014. **4**: p. 5941.
9. Ren, T., W. He, and P.G. Barr-Gillespie, *Reverse transduction measured in the living cochlea by low-coherence heterodyne interferometry.* Nat Commun, 2016. **7**: p. 10282.
10. Lee, H.Y., P.D. Raphael, J. Park, A.K. Ellerbee, B.E. Applegate, and J.S. Oghalai, *Noninvasive in vivo imaging reveals differences between tectorial membrane and basilar membrane traveling waves in the mouse cochlea.* Proc Natl Acad Sci U S A, 2015. **112**(10): p. 3128-33.
11. Lee, H.Y., P.D. Raphael, A. Xia, J. Kim, N. Grillet, B.E. Applegate, A.K. Ellerbee Bowden, and J.S. Oghalai, *Two-Dimensional Cochlear Micromechanics Measured In Vivo Demonstrate Radial Tuning within the Mouse Organ of Corti.* J Neurosci, 2016. **36**(31): p. 8160-73.
12. Dong, W. and E.S. Olson, *Detection of cochlear amplification and its activation.* Biophys J, 2013. **105**(4): p. 1067-78.
13. Vavakou, A., N.P. Cooper, and M. van der Heijden, *The frequency limit of outer hair cell motility measured in vivo.* Elife, 2019. **8**.
14. Meaud, J. and K. Grosh, *The effect of tectorial membrane and basilar membrane longitudinal coupling in cochlear mechanics.* J Acoust Soc Am, 2010. **127**(3): p. 1411-21.
15. Meaud, J. and K. Grosh, *Effect of the attachment of the tectorial membrane on cochlear micromechanics and two-tone suppression.* Biophys J, 2014. **106**(6): p. 1398-405.
16. Yoon, Y.J., C.R. Steele, and S. Puria, *Feed-forward and feed-backward amplification model from cochlear cytoarchitecture: an interspecies comparison.* Biophys J, 2011. **100**(1): p. 1-10.
17. Wang, Y., C.R. Steele, and S. Puria, *Cochlear Outer-Hair-Cell Power Generation and Viscous Fluid Loss.* Scientific reports, 2016. **6**: p. 19475.
18. Liu, Y.W. and S.T. Neely, *Outer hair cell electromechanical properties in a nonlinear piezoelectric model.* J Acoust Soc Am, 2009. **126**(2): p. 751-61.
19. Liu, Y.W. and S.T. Neely, *Suppression tuning of distortion-product otoacoustic emissions: results from cochlear mechanics simulation.* J Acoust Soc Am, 2013. **133**(2): p. 951-61.
20. Nam, J.H. and R. Fettiplace, *Optimal electrical properties of outer hair cells ensure cochlear amplification.* PLoS One, 2012. **7**(11): p. e50572.
21. Nam, J.H., *Microstructures in the organ of Corti help outer hair cells form traveling waves along the cochlear coil.* Biophys J, 2014. **106**(11): p. 2426-33.

22. Liu, Y., S.M. Gracewski, and J.H. Nam, *Consequences of Location-Dependent Organ of Corti Micro-Mechanics.* PLoS One, 2015. **10**(8): p. e0133284.
23. Cormack, J., Y. Liu, J.H. Nam, and S.M. Gracewski, *Two-compartment passive frequency domain cochlea model allowing independent fluid coupling to the tectorial and basilar membranes.* The Journal of the Acoustical Society of America, 2015(To appear).
24. Prodanovic, S., S.M. Gracewski, and J.H. Nam, *Power Dissipation in the Cochlea Can Enhance Frequency Selectivity.* Biophys J, 2019. **116**(7): p. 1362-1375.
25. Ren, T., W. He, and D. Kemp, *Reticular lamina and basilar membrane vibrations in living mouse cochleae.* Proc Natl Acad Sci U S A, 2016.
26. He, W., D. Kemp, and T. Ren, *Timing of the reticular lamina and basilar membrane vibration in living gerbil cochleae.* Elife, 2018. **7**.
27. Cooper, N.P., A. Vavakou, and M. van der Heijden, *Vibration hotspots reveal longitudinal funneling of sound-evoked motion in the mammalian cochlea.* Nat Commun, 2018. **9**(1): p. 3054.
28. Dewey, J.B., B.E. Applegate, and J.S. Oghalai, *Amplification and Suppression of Traveling Waves along the Mouse Organ of Corti: Evidence for Spatial Variation in the Longitudinal Coupling of Outer Hair Cell-Generated Forces.* J Neurosci, 2019. **39**(10): p. 1805-1816.
29. Strimbu, C.E., Y. Wang, and E.S. Olson, *Manipulation of the Endocochlear Potential Reveals Two Distinct Types of Cochlear Nonlinearity.* Biophys J, 2020. **119**(10): p. 2087-2101.
30. van der Heijden, M. and A. Vavakou, *Rectifying and sluggish: Outer hair cells as regulators rather than amplifiers.* Hear Res, 2021: p. 108367.
31. Shokrian, M., C. Knox, D.H. Kelley, and J.H. Nam, *Mechanically facilitated micro-fluid mixing in the organ of Corti.* Sci Rep, 2020. **10**(1): p. 14847.
32. Karavitaki, K.D. and D.C. Mountain, *Evidence for outer hair cell driven oscillatory fluid flow in the tunnel of corti.* Biophys J, 2007. **92**(9): p. 3284-93.
33. Zagadou, B.F. and D.C. Mountain, *Analysis of the cochlear amplifier fluid pump hypothesis.* J Assoc Res Otolaryngol, 2012. **13**(2): p. 185-97.
34. van der Heijden, M., *Frequency selectivity without resonance in a fluid waveguide.* Proc Natl Acad Sci U S A, 2014.
35. Salt, A.N. and K. Hirose, *Communication pathways to and from the inner ear and their contributions to drug delivery.* Hear Res, 2018. **362**: p. 25-37.
36. Liu, Y., S.M. Gracewski, and J.H. Nam, *Two passive mechanical conditions modulate power generation by the outer hair cells.* PLoS Comput Biol, 2017. **13**(9): p. e1005701.
37. Johnson, D.L., J. Koplik, and R. Dashen, *Theory of dynamic permeability and tortuosity in fluid-saturated porous media.* Journal of fluid mechanics, 1987. **176**: p. 379-402.
38. Jabeen, T., J.C. Holt, J. Becker, and J.H. Nam, *Interactions between passive and active vibrations in the organ of Corti In vitro.* Biophysical Journal, 2020. **119**(2): p. 314-325.
39. Ohyama, K., A.N. Salt, and R. Thalmann, *Volume flow rate of perilymph in the guinea-pig cochlea.* Hear Res, 1988. **35**(2-3): p. 119-29.
40. Juhn, S.K., L.P. Rybak, and S. Prado, *Nature of blood-labyrinth barrier in experimental conditions.* Ann Otol Rhinol Laryngol, 1981. **90**(2 Pt 1): p. 135-41.
41. Lighthill, J., *Acoustic Streaming in the Ear Itself.* Journal of Fluid Mechanics, 1992. **239**: p. 551-606.
42. Edom, E., D. Obrist, and L. Kleiser, *Steady streaming in a two-dimensional box model of a passive cochlea.* Journal of Fluid Mechanics, 2014. **753**: p. 254-278.
43. Sumner, L., J. Mestel, and T. Reichenbach, *Steady streaming as a method for drug delivery to the inner ear.* Sci Rep, 2021. **11**(1): p. 57.

Nonlinearity and Energetics of Active Cochlear Models

Karl Grosh[1, 2, 3, a)] and Wen Cai[1, b)]

[1)] *Department of Mechanical Engineering, University of Michigan, G.G. Brown Building, 2350 Hayward St., Ann Arbor, MI 48109-2125, U.S.A.*

[2)] *Department of Biomedical Engineering, University of Michigan, 1107 Carl A. Gerstacker Building, 2200 Bonisteel, Blvd., Ann Arbor, MI 48109-2099, U.S.A.*

[3)] *Kresge Hearing Research Institute, 4605 Medical Science Unit II, Ann Arbor, MI 48109-5616, U.S.A.*

[a)] *Corresponding author: grosh@umich.edu*

[b)] *Electronic mail: caiwen@umich.edu*

Abstract. The outer hair cell (OHC) of the mammalian cochlea is the nexus of the active processes giving rise to the nonlinear, biologically vulnerable, acoustic response. We present a model for the behavior of the OHC in view of its mechanical and electrical properties, and the external loading of the cell. Because of the low-pass electrical membrane impedance and rate dependent processes, there is a continuing debate on the mechanism of the amplification process at high frequencies. We will focus on the electrical-to-mechanical energy conversion at the cellular level, and show how we must consider the external mechanical loading of the cell to interpret the power transfer. In addition, we show that simple models can be used to fit *in vitro* data from experiments, but subtle model changes in the parameters change the predictions of power deposition by the OHCs.

INTRODUCTION

The outer hair cell (OHC) of the mammalian cochlea is the mediator of active and nonlinear processes necessary for healthy hearing. The mechano-electric-transducer (MET) channels of the OHC stereocilia are mechanically gated ion channels which present a nonlinear deformation dependent conductance to the resting voltage difference between the scala media and OHC intracellular voltage. Prestin-related somatic motility of the OHC is thought to either power a cycle-by-cycle conversion of electrical to mechanical power or be implicated in an automatic gain control algorithm. In the present paper, we will investigate the constitutive behavior of the OHC and the implications of this behavior for cycle-by-cycle power conversion.

The discovery of the electromotile response of outer hair cells (OHCs) by Brownell [5] has spawned research into determining the precise nature by which the OHCs participate in the hearing process. To this end, many experimental and theoretical studies have been undertaken to quantify the transduction process. OHCs exhibit voltage and stress dependent capacitance [4, 13, 17, 20], voltage dependent stiffness [12], and voltage dependent length changes [3, 5, 21]. There is contradicting evidence as to the ability of OHCs to contribute effective cycle-by-cycle forcing at high frequencies. Some *in vitro* (e.g., [8]) and *in vivo* [10] experiments indicate that OHC-mediated electromotility extends to 100 kHz, while *in vitro* experiments indicate that rate dependence could limit high frequency forcing (e.g., [9, 22]). In the present study, we focus on fitting models to the *in vitro* experiments and testing the predictions of the mathematical models along the lines of recent work [19, 22].

OHC MECHANICS

The OHC is a fluid-filled cell of nearly cylindrical shape *in situ* and when carefully extracted from the organ of Corti [2, 15]. Its outer cell wall consists of highly dense prestin molecules packed into the plasma membrane mechanically connected in parallel to a cortical lattice. As has been documented for isotropic and anisotropic nonlinear constitutive modeling of OHCs [6, 14, 15], assuming isovolumetric expansion and contraction of the cell due to the enclosed fluid, the radial strain can be related to the axial strain, and the stress-strain relations can be written in terms of the axial force on the cell and axial displacement of the cell [1] along with a dependence on the transmembrane potential. Under equilibrium conditions at constant temperature, the Gibb's free energy function of the coupled electro-mechanical system, which depends on the load applied to the electromotile element and the transmembrane potential, can be used to represent the thermodynamics state of the OHC [6, 7, 19]. In Fig. 1, a schematic of the various components of a

one-dimensional model of the OHC are shown, and we next discuss the integration of these elements into a model for OHC response.

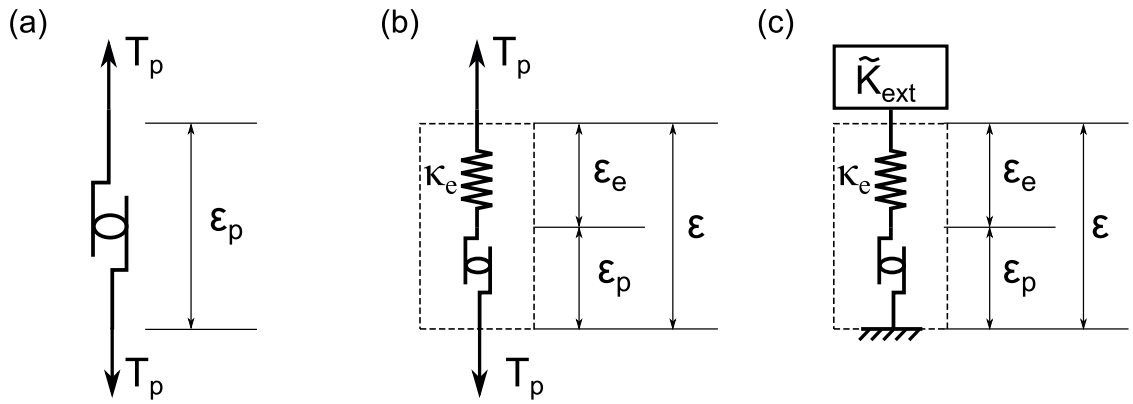


FIGURE 1. (a) Electromotile element. (b) Electromotile element with a series elastic element (with compliance κ_e). (c) OHC loaded by an external impedance ($\tilde{K}_{ext}(\omega)$.)

Electromotile OHC Mechanics

Equilibrium Thermodynamics

In Fig. 1a, the nonlinear electromotile element of the OHC is shown. The Gibb's free energy function of the coupled electro-mechanical system, $G(T_p, V)$, where T_p is the load on the electromotile element and V is the transmembrane potential. At equilibrium, the displacement (or strain) of the electromotile element, ε_p, and charge transfer, Q_{nl}, can be related to the Gibb's free energy by

$$\varepsilon_p = -\frac{\partial G}{\partial T_p} \quad \text{and} \quad Q_{nl} = -\frac{\partial G}{\partial V} \tag{1}$$

where G is the coupled thermodynamic potential function for the OHC. This potential may be derived for various assumption of the conformal statistics of the electromotile protein (e.g., two-state Boltzmann model [6, 15] or using Langevin statistics [19]). Our exposition here is less focused on the precise details of the constitutive model than linearized outcome of this nonlinear constitutive behavior about an equilibrium point. For instance, the linearized capacitance varies like a bell shaped function with respect to large changes in the transmembrane voltage [4, 20], other properties similarly vary and are may be relatively easier (like free electromotility) or more difficult (like the cell stiffness or displacement-clamped capacitance) to measure, but are likely to assume a similar nonlinear dependence on resting electrical and mechanical conditions [2]. Deviations from equilibrium at rates slower than the speed of conformal changes in the protein can be described using a quasistatic perturbation about from equilibrium as:

$$\delta\varepsilon_p = -\left.\frac{\partial^2 G}{\partial T_p^2}\right|_{eq} \delta T_p - \left.\frac{\partial^2 G}{\partial T_p \partial V}\right|_{eq} \delta V = \kappa_{nl}\delta T_p + d_{nl}\delta V \quad , \tag{2}$$

$$\delta Q_{nl} = -\left.\frac{\partial^2 G}{\partial V^2}\right|_{eq} \delta V - \left.\frac{\partial^2 G}{\partial V \partial T_p}\right|_{eq} \delta T_p = d_{nl}\delta T_p + C_{nl}\delta V \tag{3}$$

where the notation $|_{eq}$ indicates the derivatives are taken at the equilibrium values of T_p and V, $\kappa_{nl} = -\frac{\partial^2 G}{\partial T_p^2}|_{eq}$ is the compliance, $C_{nl} = -\frac{\partial^2 G}{\partial V^2}|_{eq}$ is the capacitance, and $d_{nl} = -\frac{\partial^2 G}{\partial V \partial T_p}|_{eq}$ is the piezoelectric strain coefficient. The δ notation indicates the perturbation of the quantity about its equilibrium value (e..g, $\delta T_p = T_p - T_p|_{eq}$).

Rate dependent state transitions

The equations above provide the coupled force-displacement-voltage-charge relations of the motile elements in the OHC lateral wall if the state dependent transitions happen nearly instantaneously. We can model the non-equilibrium probability function ($P(t)$) to include rate dependence by using an absolute Eyring absolute rate equation [16, 22, 23]

$$\frac{\partial P(t)}{\partial t} = \mathcal{K}_+ - (\mathcal{K}_+ + \mathcal{K}_-)P(t) \quad \text{where the rate coefficients are} \tag{4}$$

$$\mathcal{K}_+ = \mathcal{K} e^{-\beta\alpha(T_p a + (V-V_0)q)} \quad , \tag{5}$$

$$\mathcal{K}_- = \mathcal{K} e^{\beta(1-\alpha)(T_p a + (V-V_0)q)} \quad , \tag{6}$$

$\mathcal{K}$ is related to the energy barrier between the two states at equilibrium (when $T_p = 0$ and $V = V_0$), $\beta = k\Theta$, where k is Boltzmann's constant, and Θ is the absolute temperature, $T_p a$ and Vq represent the difference in Gibb's free energy between the short and long states of a two-state Boltzmann model for the OHC electromotile mechanics, and deviations of the constant α from 1/2 represent the asymmetry of the energy barrier ($0 \leq \alpha \leq 1$). The constants a and q represent whole cell length change and charge change between the two states. The equilibrium probability for such a model is given by

$$P_e = 1/(1 + e^{-\beta(T_p|_{eq} a + (V|_{eq} - V_0)q)}) \quad . \tag{7}$$

Linearizing Eq. 4 about equilibrium yields

$$(1 + \tau_{nl}\frac{\partial}{\partial t})\delta\varepsilon_p = \kappa_{nl}\delta T_p + d_{nl}\delta V \quad , \tag{8}$$

$$(1 + \tau_{nl}\frac{\partial}{\partial t})\delta Q_{nl} = d_{nl}\delta T_p + C_{nl}\delta V \tag{9}$$

where $\tau_{nl} = (\mathcal{K}_+ + \mathcal{K}_-)|_{eq}$, and the other constants are defined in Eqs. 2 - 3. Overall, the linearized effect of these rate transformations manifest as viscous–like damping (or loss) contributions. However, these loss factors depend on the intrinsic resting properties of the cell differently than external viscous loading from fluidic or viscoelastic effects.

Using a $e^{j\omega t}$ time dependence and letting $\delta\tilde{\varepsilon}$, $\delta\tilde{T}_p$, $\delta\tilde{Q}_{nl}$, and $\delta\tilde{V}$ represent the complex amplitudes of the corresponding variables, we can solve for the complex, frequency dependent quantities as

$$\delta\tilde{\varepsilon} = \frac{1}{1 + j\omega\tau_{nl}}\left(\kappa_{nl}\delta\tilde{T}_p + d_{nl}\delta\tilde{V}\right) = \tilde{\kappa}_{nl}\delta\tilde{T}_p + \tilde{d}_{nl}\delta\tilde{V} \quad , \tag{10}$$

$$\delta\tilde{Q}_{nl} = \frac{1}{1 + j\omega\tau_{nl}}\left(d_{nl}\delta\tilde{T}_p + C_{nl}\delta V\right) = \tilde{d}_{nl}\delta\tilde{T}_p + \tilde{C}_{nl}\delta\tilde{V} \quad . \tag{11}$$

Here definitions of the frequency and rate dependent constants are designated by a tilde and defined in Eqs. 10-11 (e.g., $\tilde{\kappa}_{nl} = \kappa_{nl}/(1 + j\omega\tau_{nl})$.)

Series elastic element

In addition to the nonlinear prestin-related gating mechanics, the linear elastic properties of the plasma membrane or the electromotile element itself [6, 15, 19] will likely contribute to the response. Such a model is shown in Fig. 1b. The total displacement of the cell becomes

$$\varepsilon = \varepsilon_p + \varepsilon_e = \varepsilon_p + \kappa_e T_p \tag{12}$$

where κ_e is the elastic compliance of the OHC lateral wall. In the frequency domain, we include a viscous element mechanically in parallel with this elastic spring by modifying the elastic compliance as $\tilde{\kappa}_e = \kappa_e/(1 + j\omega\tau_e)$, where the viscous force is equal to $\tau_e\dot{\varepsilon}_e/\kappa_e$. Combining Eq. 12 with Eqs. 10-11 we can write the linearized relations of the OHC as

$$\delta\tilde{\varepsilon} = \tilde{\kappa}_{ohc}\delta\tilde{T}_p + \tilde{d}_{nl}\delta\tilde{V} \quad , \tag{13}$$

$$\delta\tilde{Q}_{nl} = \tilde{d}_{nl}\delta\tilde{T}_p + \tilde{C}_{nl}\delta\tilde{V} \tag{14}$$

where $\tilde{\kappa}_{ohc} = \tilde{\kappa}_e + \tilde{\kappa}_{nl}$. These relations may be converted for force-charge form as

$$\delta\tilde{T}_p = \frac{1}{\tilde{\kappa}_{ohc}}\tilde{\delta\varepsilon} - \frac{\tilde{d}_{nl}}{\tilde{\kappa}_{ohc}}\delta\tilde{V} \quad , \tag{15}$$

$$\delta\tilde{Q}_{nl} = \frac{\tilde{d}_{nl}}{\tilde{\kappa}_{ohc}}\tilde{\delta\varepsilon} + (\tilde{C}_{nl} - \frac{\tilde{d}_{nl}^2}{\tilde{\kappa}_{ohc}})\delta\tilde{V} \tag{16}$$

From these relations it can be seen that C_{nl} is the free capacitance of the cell at constant force [22] while $C_{nl} - \frac{d_{nl}^2}{\kappa_{ohc}}$ is the blocked capacitance (at low frequency). The compliance of the OHC is κ_{ohc} and its value is both force and voltage dependent. However, there has been considerable variation in the measured value of κ_{ohc} and its dependence on the resting potential [11, 12]. The linearized piezoelectric strain coefficient d_{nl} and the electromechanical coupling coefficient $e_{nl} = d_{nl}/\kappa_{ohc}$ have likewise been characterized with significant variability (with e_{nl} varying from 0.1 nN/mV to 0.005 nN/mV) [11, 15]. Equations 13-16 show the sensitivity of these parameter estimates to the electrical and mechanical boundary conditions as well as the equilibrium values. Experimental measurements of quantities such as the blocked capacitance, while challenging, would inform modeling efforts at low and high frequencies.

OHC loaded by an external impedance

To examine a generic loading condition for the OHC, an external spring-mass-damper is introduced to simulate the external load during *in vitro*, isolated OHC measurements or as a simple surrogate for loading from the organ of Corti. In Fig. 1(c), the external spring-mass-damper system is modeled as an external impedance that is termed as $\tilde{K}_{ext}$ so that $\delta\tilde{T}_p = -\tilde{K}_{ext}\tilde{\delta\varepsilon}$. Eliminating $\delta\tilde{T}_p$ from Eq. 17 we obtain the voltage driven electromotility (eM) as

$$\tilde{\delta\varepsilon} = \frac{\tilde{d}_{nl}\delta\tilde{V}}{\tilde{K}_{ext}\tilde{\kappa}_{ohc} + 1} \quad , \tag{17}$$

where $\tilde{K}_{ext}$ is defined as

$$\tilde{K}_{ext} = -m\omega^2 + K + j\omega\eta \quad . \tag{18}$$

In Eq. 18, m, K, and η are respectively the mass, stiffness, and viscous damping of the external mechanical system. The corresponding nonlinear charge can be determined by substituting Eq. 17 into Eq.16. For the loading condition shown in Fig. 1c, under a voltage clamp condition, the power delivered to the external load from the OHC is

$$P_{input} = \frac{1}{2}\omega \ \mathrm{Im}\big[\frac{\tilde{d}_{nl}}{\tilde{\kappa}_{ohc}}\delta\tilde{V}\tilde{\delta\varepsilon}^*\big] \quad , \tag{19}$$

where $\tilde{\delta\varepsilon}^*$ indicates the complex conjugate of $\tilde{\delta\varepsilon}$, Im[ ] indicates the imaginary part of the complex number.

RESULTS

Using these models, a variety of different experiments may be simulated. For this paper, we limit our focus to voltage driven (so-called voltage-clamped) experiments. In Fig. 2 we show simulations of eM (or displacement) using the parameters listed in Table I in Eq. 17 match the experiments of [8] quite well. Using the same model with slightly different parameters, we can also simulate experimental results from [22] measuring the eM and maximum charge (Qmax, where Qmax is the maximum nonlinear charge associated with the prestin conformation change, which relates to the nonlinear capacitance [22]). As shown in Fig. 3, with the increase of frequency, both eM and Qmax (or nonlinear capacitance) decrease, which replicates the low-pass characteristic of the OHC seen in these results.

In Fig. 4 we examine the effect of varying τ_{nl} on the output power of a single OHC loaded by $\tilde{K}_{ext}$ with varying resonance frequencies. We choose η so that the quality factor (QF) of the dynamic system is constant:

$$\eta = \frac{2}{QF}\sqrt{(K + \frac{1}{\kappa_{ohc}})m} \quad . \tag{20}$$

TABLE I. Parameter Values

	κ_{nl}[m/N]	κ_e[m/N]	κ_{ohc}[m/N]	d_{nl}[nm/mV]	$\tau_{nl}[\mu s]$	QF	m[pg]
OHC 65	1.7	131.3	133	3	0.7	0.82	131
OHC 84	0.3	57.3	57.6	0.6	0.1	1.18	101
OHC (Fig.3)	7	93	100	12.5	50	1.45	338

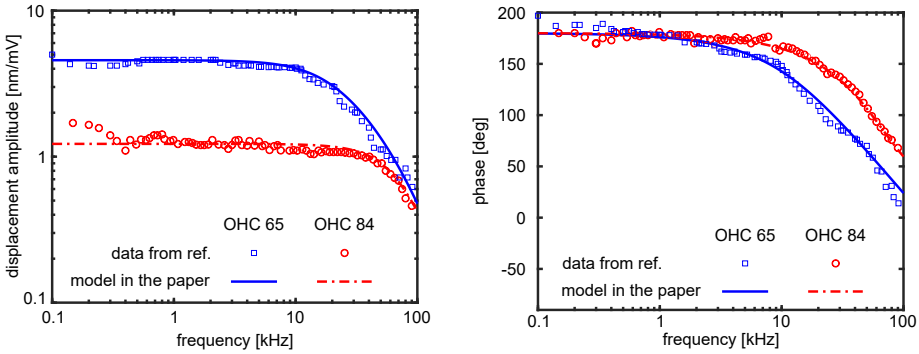


FIGURE 2. OHCs displacement and phase simulated under voltage clamp condition using parameters in Table I compared to data from [8].

As expected, the peak of deposited power occurs at the resonance of the OHC-external loading system. By changing the time constant of the state-dependent transformations, τ_{nl} of 50 μs or a corner frequency, $f_c = 1/2\pi\tau_{nl}$ of 3.2 kHz, and 10 μs ($f_c = 16$ kHz), dramatic changes in the deposited power can occur even when the corner frequency less than the resonance.

CONCLUSION

We have developed a constitutive theory for OHC somatic electromotility that is similar to those presented previously [19, 22], with minor differences in how the charge polarity changes are related to membrane deformation and how the measured nonlinear capacitance is computed. Although various embellishments of the model are possible [22], we wanted to see if a relatively simple model could fit the existing data, and we found that it can; however, the parameter space of the intrinsic and extrinsic losses are difficult to uniquely uncouple with present experimental data. We also found that driving an underdamped system, $QF = 6$, that variations in the state-dependent rates affected the output power. It should be emphasized that the corner frequency of a filter is not a cut-off frequency, and effective forcing can occur above f_c. In addition,

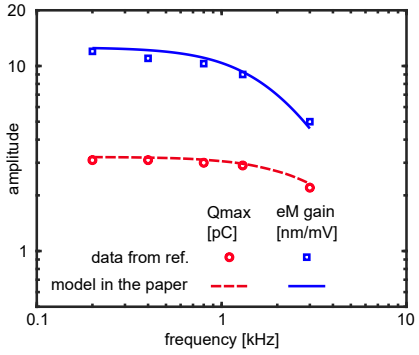


FIGURE 3. OHC eM gain and nonlinear charge simulated with parameters in Table I compared against data from [22].

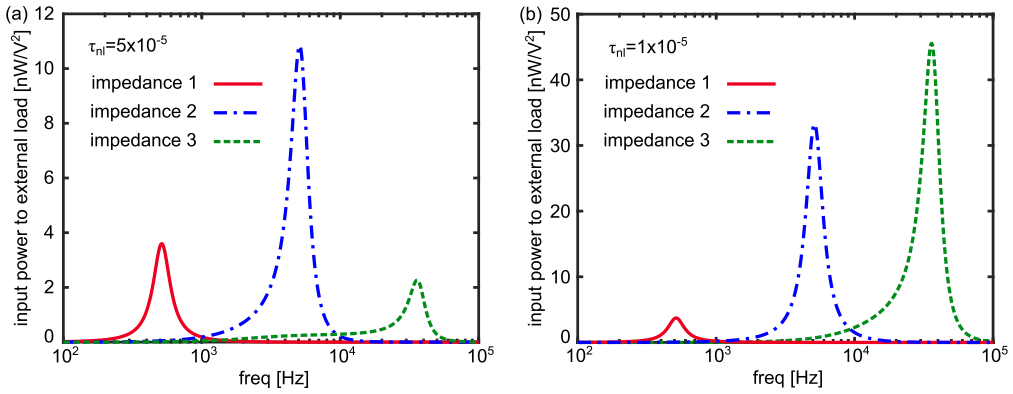


FIGURE 4. Mechanical power input to three different external impedances with (a) $\tau_{nl} = 50\mu s$ (or $f_c = 3.2\ kHz$) and (b) $\tau_{nl} = 10\mu s$ (or $f_c = 16\ kHz$). Other parameters for the OHC are the same as in Fig.3. The same stiffness $K = 8.28$ mN/m and the quality factor $QF = 6$ is used for the three simulation results in this figure. By adjusting the mass m we set the resonance of the three impedance systems respectively at 0.51 kHz (impedance 1), 5 kHz (impedance 2), and 36 kHz (impedance 3).

while a QF of six is a reasonable value for a healthy, active cochlear best frequency (e.g., [10]), the passive cochlear milleu is likely nearly critically damped.

The present study is incomplete. We have only studied voltage clamped behavior, and controlled current excitation will also provide important insights to *in vivo* active processes. Further, the full model parameter space of intrinsic damping of the OHC and the external load has yet to be explored and analyzed for implications to *in vivo* cochlear function and interpretation of experiments. Inclusion of intrinsic damping due to rate transformations into a viscous model of the entire organ of Corti or a global cochlear model [18] will also provide more insights as to the interplay between and importance of τ_{nl} and global cochlear micro- and macro-mechanics.

ACKNOWLEDGEMENT

This work was supported by National Institutes of Health grant NIDCD R01–04084.

REFERENCES

1. M. Adachi, M. Sugawara, and K. H. Iwasa. Effect of turgor pressure on outer hair cell motility. *J. Acoust. Soc. Am.*, 108(5.1):2299–2306, Nov 2000.
2. J. Ashmore. Cochlear outer hair cell motility. *Physiol. Rev.l.*, 88:173–210, 2008.
3. J. F. Ashmore. A fast motile response in guinea-pig outer hair cells: the cellular basis of the cochlear amplifier. *J. Physiol. (Lond.).*, 388:323–347, 1987.
4. J. F. Ashmore. Forward and reverse transduction in the mammalian cochlea. *Neurosci. Res. Suppl.*, 12:S39–S50, 1990.
5. W. E. Brownell, C. R. Bader, D. Bertrand, and Y. De Ribaupierre. Evoked mechanical responses of isolated cochlear hair cells. *Science.*, 227:194–196, 1985.
6. N. Deo and K. Grosh. Two-state model for outer hair cell stiffness and motility. *Biophys. J.*, 86(6):3519–3528, 2004.
7. N. Deo and K. Grosh. Simplified nonlinear outer hair cell models. *J. Acoust. Soc. Am*, 117(4):2141–2146, 2005.
8. G. Frank, W. Hemmert, and A. Gummer. Limiting dynamics of high-frequency electromechanical transduction of outer hair cells. *Proc. Natl. Acad. Sci. USA.*, 96:4420–4425, April 1999.
9. J. E. Gale and J. F. Ashmore. The outer hair cell motor in membrane patches. *Pflügers. Arch. Eur. J. Physiol.*, 434:267–271, 1997.
10. K. Grosh, J. F. Zheng, Y. Zou, E. de Boer, and A. L. Nuttall. High-frequency electromotile responses in the cochlea. *J. Acoust. Soc. Am.*, 115(5):2178–2184, 2004.
11. R. Hallworth. Passive compliance and active force generation in the guinea pig outer hair cell. *J. Neurophysiol.*, 74(6):2319–2328, 1995.
12. D. Z. Z. He and P. Dallos. Properties of voltage-dependant somatic stiffness of cochlear outer hair cells. *J. Assoc. Res. Otolaryngol.*, 01:64–81, 2000.
13. K. H. Iwasa. Effect of stress on the membrane capacitance of the auditory outer hair cell. *Biophys. J.*, 65:492–498, July 1993.
14. K. H. Iwasa. A two-state piezoelectric model for outer hair cell motility. *Biophys. J.*, 81:2495–2506, November 2001.
15. K. H. Iwasa and M. Adachi. Force generation in the outer hair cell of the cochlea. *Biophys. J.*, 73:546–555, July 1997.
16. J. Sneyd J. Keener. *Mathematical physiology: II: Systems physiology.* NY: Springer New York, 2009.

17. S. Kakehata and J. Santos-Sacchi. Membrane tension directly shifts voltage dependance of outer hair cell motility and associated gating charge. *Biophys. J.*, 68:2190–2197, May 1995.
18. J. Meaud and K. Grosh. The effect of tectorial membrane and basilar membrane longitudinal coupling in cochlear mechanics. *J. Acoust. Soc. Am.*, 127(3):1411–1421, 2010.
19. R. Rabbitt. The cochlear outer hair cell speed paradox. *Proc. Natl. Acad. Sci. USA.*, 117(36):21880–21888, 2020.
20. J. Santos-Sacchi. Reversible inhibition of voltage-dependant outer hair cell motility and capacitance. *J. Neurosci.*, 11(10):3096–3110, October 1991.
21. J. Santos-Sacchi and J. P. Dilger. Whole cell currents and mechanical responses of isolated outer hair cells. *Hear. Res.*, 35:143–150, 1988.
22. J. Santos-Sacchi, K.H. Iwasa, and W. Tan. Outer hair cell electromotility is low-pass filtered relative to the molecular conformational changes that produce nonlinear capacitance. *J. Gen. Physiol.*, 151(12):1369–1385, 2019.
23. A.J. Hudspeth Y. Choe, M.O. Magnasco. A model for amplification of hair-bundle motion by cyclical binding of ca2+ to mechanoelectrical-transduction channels. *Proc. Natl. Acad. Sci. USA.*, 95(26):15321–15326, 1998.

Micro-mechanical motion of the guinea-pig organ of Corti

Nigel P. Cooper1[1, a], Anna Vavakou[1, b] and Marcel van der Heijden[1, c]

[1]*Erasmus MC, Rotterdam, The Netherlands.*

[a)]Corresponding author: nigel@mechanicsofhearing.com
[b)]a.vavakou@erasmusmc.nl
[c)]marcel.vdh.work@gmail.com

Abstract. One-dimensional measurements of sound-evoked motion were made by applying optical coherence tomography to the high frequency, basal turn of the cochlea in deeply anesthetized guinea-pigs. Wideband acoustic stimuli were used to evoke nanometer-scale vibrations of the basilar membrane (BM) and organ of Corti (OoC). Both macro- and micro-mechanical aspects of the responses were analysed, where macro-mechanical refers to 'motions observed on the BM', and micro-mechanical covers the motions of other structures 'relative to those observed on the BM'. Both macro- and micro-mechanical motions exhibited compressive nonlinearity across wide dynamic ranges (typically from ~30 to >80 dB SPL near a site's characteristic frequency or CF). Only the macro-mechanical motion was sharply tuned, however – the micro-mechanical motion, which included a transverse 'squeezing' type of OoC deformation, was low-pass filtered (well below CF), and *not* sharply tuned in its own right. The micro-mechanical motion was also delayed, accumulating ~0.5 cycles of phase-shift with respect to the macro-mechanics between 0 Hz and CF. Outer hair cell (OHC) and Deiters cell (DC) length changes were also analysed: OHC length changes exceeded macro-mechanical movements by factors of ~2-4 at low frequencies, but decreased in (relative) magnitude with increasing frequency as the CF was approached. At low frequencies, OHCs became longer and shorter when the BM moved towards the scala tympani and scala vestibuli, respectively. At frequencies closer to CF, however, this phase relationship reversed, in accordance with the ~0.5 cycles of phase-accumulation in the micro-mechanics. These observations are discussed with respect to previous findings, including a recent theoretical study that considered frequency-dependent changes in the mode-shape of a wave propagating along a simple physical waveguide [1].

INTRODUCTION

Recent experimental studies (e.g., [2–8]) have provided new insights into the way that healthy mammalian cochleae respond to sound stimuli, advancing the field of cochlear mechanics far beyond the level of the BM's classically studied macro-mechanics [9–11] (see Table 1 for abbreviations). These studies use novel optical techniques to probe the sound-evoked vibrations of light-scattering structures deep inside the cochlear partition, without the need to physically open the cochlea itself. Here we apply one such technique to study the spatial distributions of sound-evoked vibrations across the width and depth of the cochlear partition in the basal turn of the guinea-pig cochlea. We focus on the differential motions that occur between anatomical structures within single 'radial/transverse' cross-sections, i.e. motions that relate directly to the micro-mechanics of the OoC.

TABLE 1. Abbreviations.

BM	Basilar membrane	OoC	Organ of Corti
DC	Deiters' cell	RL	Reticular lamina
IHC	Inner hair cell	SM	Scala media
ME	Middle-ear	SPL	Sound pressure level
OCT	Optical coherence tomography	ST	Scala tympani
OHC	Outer hair cell	TM	Tectorial membrane

METHODS

Sound-evoked vibrations were recorded from the cochlear partition of healthy but deeply anaesthetized guinea-pigs, in accordance with the guidelines of the Animal Care and Use Committee at Erasmus MC. Vibrations were measured under open-bulla conditions from the first turn and hook region of the cochlea, as viewed through the intact round window membrane. A glass coverslip, back-filled with a drop of saline, provided a stable air-fluid interface above the round window region, and facilitated near-perpendicular, i.e. almost transverse, imaging of the underlying BM/OoC (see Fig. 1a). Control vibrations were measured from the incudo-stapedial joint in the middle ear (ME).

Multi-tone 'zwuis' complexes [12] were used as acoustic stimuli, with 48 irregularly-spaced components spanning most of the animal's audiogram (typically from ~0.4 to ~38 kHz). Each stimulus was presented for ~12 seconds, with inter-stimulus intervals of ~1 minute.

A spectral-domain OCT system (ThorLabs Telesto III, λ = 1310 nm) was used for imaging and 1-dimensional vibration measurements [6]. The system had a lateral resolution of ~10 µm, and an axial point spread function with a full-width-at-half-maximum-amplitude of ~6 µm. Cross-sectional (B-scan) and axial OCT images (A- and M-scans) were phase-locked to an acoustic stimulation system (Tucker Davies Technologies system III) with a sampling rate of 111.6 kHz. The M-scans provided both depth-resolved images (i.e. maps of scattering intensity vs. axial depth) and vibration information (i.e. maps of scatterer-displacement in the axial direction vs. axial depth).

Responses were analyzed by Fourier transformation of the vibration waveforms derived from contiguous groups of 3 M-scan 'pixels', where each pixel is ~2.5 µm deep. The statistical significance of each response component was assessed using Rayleigh tests of its phase-stability across time, and a rejection criterion of p=0.001 (see [12] for details). The measurement noise-floor depended on the tissue's optical accessibility and light scattering power, typical values ranging from ~30 pm/√Hz in the cochlea down to ~3 pm/√Hz in the ME.

RESULTS

Frequency-dependent vibration patterns

Figure 1 illustrates how the sound-evoked micro-mechanical vibrations of the OoC superimpose on the macro-mechanical vibrations of the BM. The spatial pattern of the BM's vibration is fairly simple and almost independent of frequency; BM vibrations peak just lateral to the junction between the BM's arcuate and pectinate zones, and vary little in phase across the BM's width (cf. [13]). In contrast, the vibration patterns within the main body of the OoC are more complex, and highly frequency-dependent:-

At low frequencies (e.g. at 1.3 kHz in Figs. 1c,f) the largest vibrations center around a vibration hotspot near the interface between the DCs and OHCs – deep within the body of the OoC. In this frequency region, the OHC/DC interface vibrates almost in-phase with the BM, but in anti-phase to the RL and TM (Fig. 1f). The RL and TM vibrate slightly less than the BM, and in anti-phase to it.

At slightly higher frequencies (e.g. the 10.7 kHz data in Figs. 1d,g) the largest vibrations remain focussed around the OHC/DC hotspot, but lag the smaller vibrations seen on the BM or the RL/TM by ~0.2 cycles of phase. For most of these 'tail frequency' responses, the BM and RL/TM vibrate with similar magnitudes (cf. Fig. 1b,d), but around 0.1-0.25 cycles out of phase (Fig. 1g).

At high frequencies (e.g. the near-CF, 34.1 kHz data in Figs. 1e,h) the largest vibrations occur on the RL and TM, and almost the entire OoC vibrates in-phase with the BM. The only region that vibrates in anti-phase to the BM at these frequencies is the tectal cell region that adjoins the lateral edge of the RL (and possibly the TM). This region vibrates much (i.e. 10 times or 20 dB) less vigorously than the RL and/or TM, however. The OHC/DC interface region (that dominates the vibrations at low frequencies) also vibrates at least 3 dB less than the nearby BM, RL or TM – i.e. what was a vibration hotspot at low frequencies becomes a vibration 'cold spot' at CF.

A fuller picture of the frequency-dependence of the OoC's vibration pattern is provided in Figure 2. The color-coded panels here show vibration magnitude-ratios (left column) and phase-differences (right column), relative to the BM's vibrations, as a function of both frequency and depth within a single 'column' of the OoC – a column that runs vertically downwards from the BM, through the DC and OHC bodies, and ends after passing through the TM (these structures are schematized by the labelled blocks in the central inter-column of Fig. 2; the *x*-position of the OCT beam underpinning the recordings is shown by the red line in Fig. 1a). As the vibration patterns in Fig. 2 are

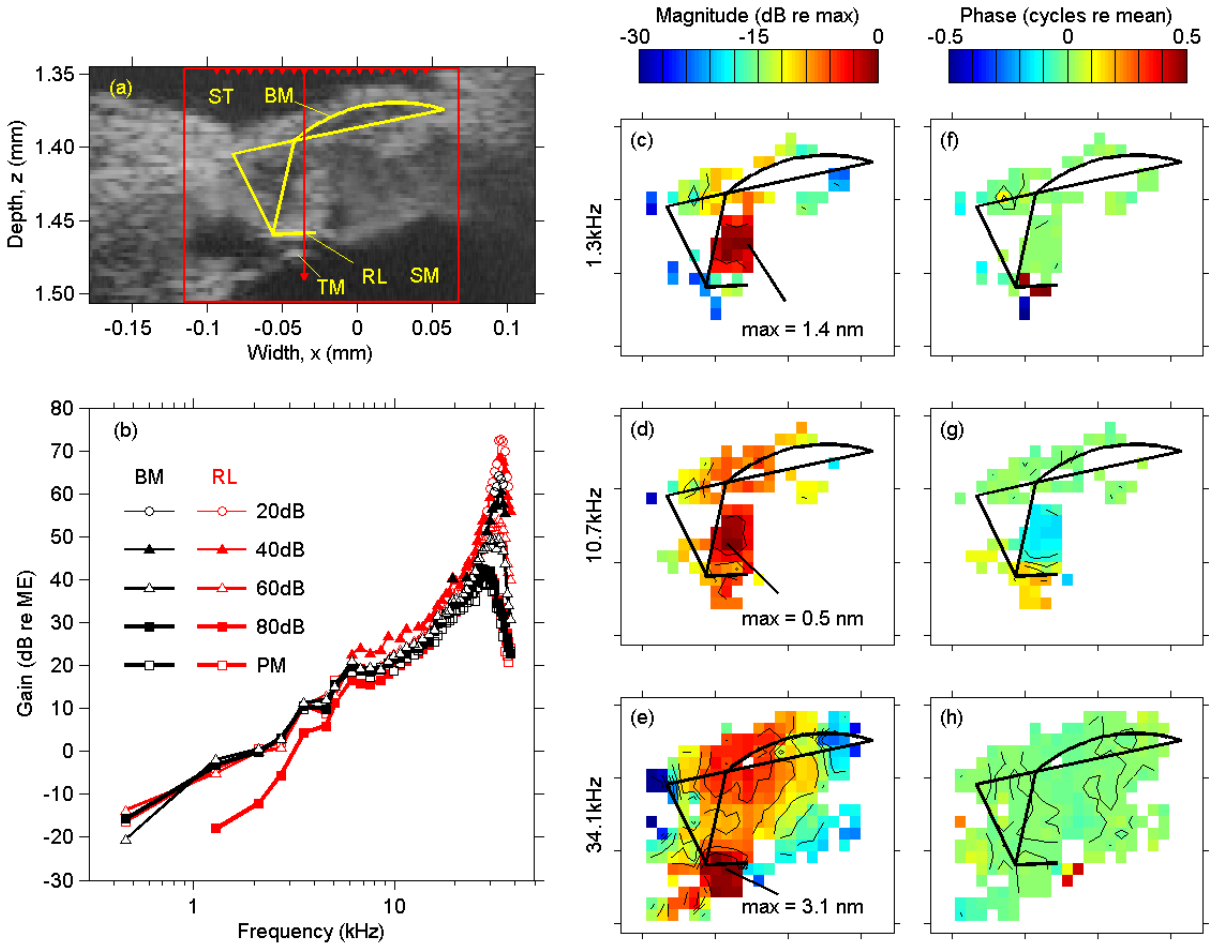


FIGURE 1. Sound-evoked vibrations in the 34 kHz region of the guinea-pig organ of Corti. (a) Cross-sectional image of the OoC (greyscale), with schematic skeletal elements of the BM/OoC superimposed in yellow (see Table 1 for abbreviations). Red line indicates one of the 20 beam positions (arrowheads) used for vibration recordings shown in (c-h), and red square indicates the boundaries of the area mapped out in (c-h). (b) Level-dependent tuning evident at the BM (black) and RL (red) between 20 and 80 dB SPL (see legend – the *post-mortem* (PM) data were collected at 75 dB SPL). (c-h) Spatial maps of vibration magnitude (c-e) and phase (f-h) locked to the 1.3 kHz (c,f), 10.7 kHz (d,g) and 34.1 kHz (e,h) components of a single 48-tone complex. Mean stimulation level in (c-h) was 60 dB SPL per spectral component. Experiment GP193023.

expressed relative to the vibrations of the BM, the maps in Fig. 2 can be thought of as 'micro-mechanical' vibration maps. The maps are almost (but not quite) uniform *post-mortem* (Figs. 2g-h), and become progressively more rich in features with decreasing SPL *in vivo* (compare rows 1-3 in Fig. 2):-

The anti-phasic (to BM) motion of the RL and TM (cf. Fig. 1f) is only seen at the lowest frequencies in our data (cf. bottom left of Figs. 2b,d), and may well be caused by OHC electro-motility ([14], see Discussion); the 0.5 cycle phase-transition between the BM and RL is very tightly focused in space, and centers near the mid-point (depth-wise) of the OHC's bodies. In this very low frequency region, the RL usually moves less than the BM (at least at levels above 60dB SPL - cf. Figs. 1b and blue pixels at lower left of Figs. 2c,e).

The OHC/DC vibration hotspot depicted in Figs. 1c,d underpins the most eye-catching feature of the frequency-versus-depth maps in Fig. 2, and progressively transitions though various states (from hotspot to cold-spot) with increasing frequency – most clearly at the lowest level illustrated (60 dB, cf. Figs. 2a,b). The OHC/DC vibrations occur in-phase with the BM at the lowest frequencies, but quickly accumulate (and subsequently lose) a phase lag of up to 0.25 cycles relative to the BM as frequency increases (cf. Fig. 2b). The OHC/DC hotspot becomes a cold-spot above ~18 kHz (i.e. ~CF/2 in Fig. 2a). The OHC/DC region loses its phase-lag (re: the BM) above ~25 kHz, and moves almost in-phase with the BM at CF (34 kHz; Fig. 2b).

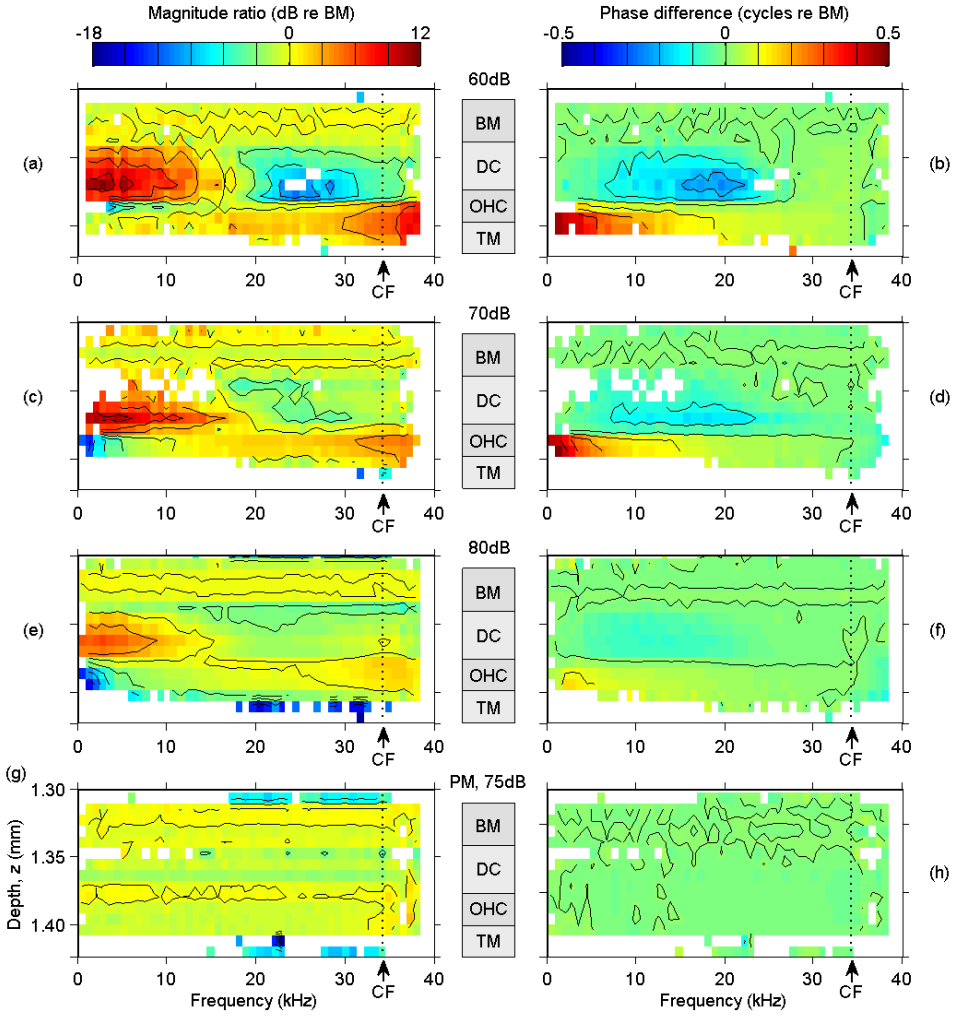


FIGURE 2. Level- and condition-dependent vibration patterns in the 34 kHz region of the guinea-pig organ of Corti. Color-coded vibration magnitude-ratios (left column; a,c,e,g) and phase-differences (right column; b,d,f,h) with respect to the BM are shown as a function of frequency and depth along a vertical (quasi-transverse) pathway (cf. red beam in Fig. 1a) through the BM, DCs, OHCs and TM (depth ranges depicted in central column). All vibrations are shown relative to the vibrations averaged across the BM's depth-range. Mean stimulation levels were 60 (a,b), 70 (c,d), 80 (e,f) and 75 dB SPL (g,h) per spectral component. The 75dB data (g, h) were collected approximately 1 hour *post-mortem*. Experiment GP19023.

At frequencies near and especially above the preparation's macro-mechanical CF (34 kHz), the largest vibrations (in both relative and absolute terms) are seen on the RL and TM (bottom right of Figs. 2a,c), and the entire OoC vibrates almost in-phase with the BM (Figs. 2b,d) The OHC/DC cold-spot also fades out just above the CF, i.e. the central region of the OoC appears to follow the movements of the BM and/or RL quite faithfully above CF; Figs. 2a,b).

Further analysis of micro-mechanical vibrations – mode-shape swapping?

The data of Figs. 1 and 2 illustrate the complex appearance of the vibration patterns observed across the cross-sectional area of the OoC, with Fig. 2 emphasizing the systematic differences between the OoC and BM vibrations. These patterns can potentially be simplified by separating the observed vibrations into multiple components or modes, such as the 'rotating-wedge' and 'squashed sandwich' deformation patterns (mode-shapes) considered in

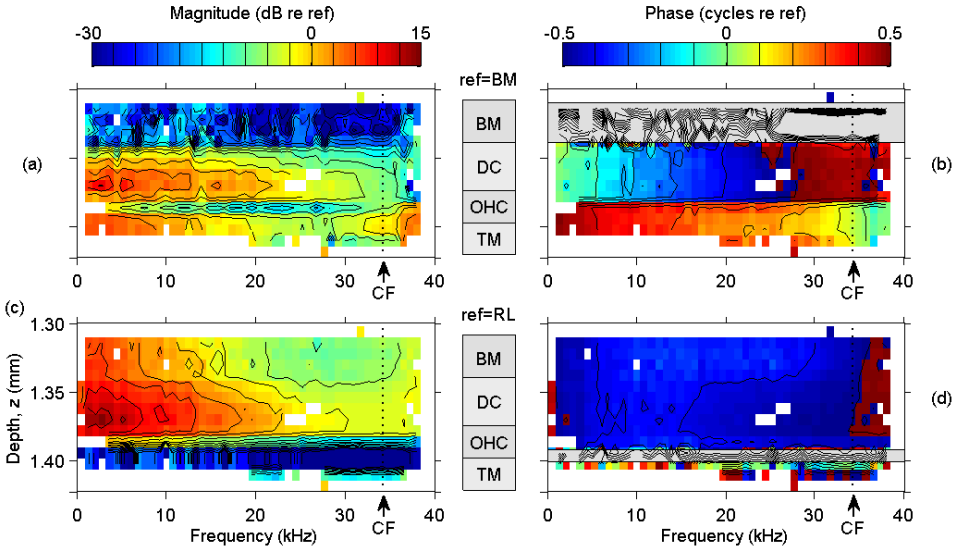


FIGURE 3. Micro-mechanical vibration patterns in the 34 kHz region of the guinea-pig organ of Corti. Color-coded vibration magnitudes (a,c) and phases (b,d) are shown as a function of frequency and depth along a vertical (quasi-transverse) pathway through the BM, DCs, OHCs and TM (depth ranges depicted in central column). Magnitudes and phases are shown after subtraction of, and normalization by, the vibrations observed on either the BM (a,b) or RL (c,d) (see text). Response phases in the BM and RL region have been greyed out in (b) and (d), respectively, to mask random variations attributable to measurement/analysis 'noise'. Mean stimulation level was 60 dB SPL per spectral component. Experiment GP19023.

numerous previous reports [1,15–18]. We attempted to decompose the observed responses into two modes, one reflecting the assumed (basic, kinematic, macro-mechanical?) drive to the system's micro-mechanics, and one reflecting the micro-mechanical deformations themselves. After representing the observed vibrations in terms of complex response vectors, we simply *subtracted* various estimates of this 'drive to the micro-mechanics' to reveal the micro-mechanical motion itself.

Figure 3 shows two examples of such analyses, both based on the same (60 dB) dataset that was illustrated in Figs. 2a,b. The maps are presented in a similar manner to those in Figs. 2, but in Fig. 3 the magnitudes and phases are those of the complex response vectors at each frequency and recording depth, after *subtracting* (as well as normalizing by) the response vectors at either one of two 'reference' structures - the BM in Figs. 3a,b, or the RL in Figs. 3c,d. The resultant (truly micro-mechanical?) pictures are much simpler than those shown in Fig. 2a,b: At each frequency (across the entire range of the response area) every point along the vertical column of OoC cells and structures is seen to vibrate in one of just two spatially-separated sections. The first section, which incorporates the basal poles of the OHCs and the DCs, moves in almost perfect synchrony with (although in anti-phase to) the RL (Fig. 3d), but exhibits a slight phase-lag relative to the BM (Fig. 3b) – a phase lag that accumulates ~0.5 cycles between 0 Hz and CF (34 kHz). The second section, which incorporates the apical poles of the OHCs, the RL and the TM, moves almost in anti-phase to the first section at every frequency (cf. Figs. 3b,d). The micro-mechanical vibration vectors are still highly frequency-dependent, however; their magnitudes decrease progressively from a maximum below ~2 kHz (where the micro-mechanical vectors are 2-4 times larger than the BM and RL reference vectors, respectively), and become similar to the reference vectors at ~20-23 kHz. The micro-mechanical vector magnitudes decrease further to become smaller than both reference vectors by CF. The micro-mechanical vibration vectors hence appear to be low-pass filtered, with a corner frequency well below the preparation's CF, and to involve anti-phasic movements on either side of a dividing line half-way between the apical and basal poles of the OHCs. Since the micro-mechanical vectors exceed the macro-mechanical (/reference) motion at low, but not at high frequencies, the OoC's overall 'mode-shape' must change with frequency: the anti-phasic overall motion of the BM and RL below 2 kHz (cf. Fig. 1c) indicating the dominance of 'anti-parallel' or 'squashing' deformations at low frequencies, and the in-phase BM and RL motion near CF (cf. Fig. 1e) indicating the dominance of 'parallel' deformations or 'rotational' motion at higher frequencies (cf. [1]).

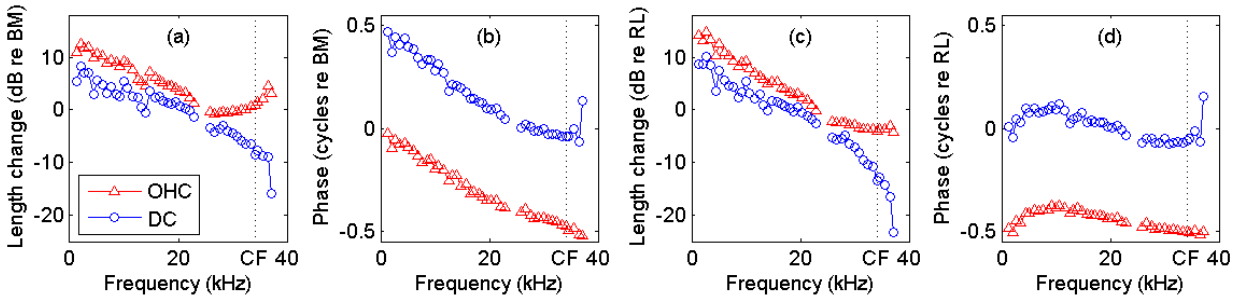


FIGURE 4. Frequency-dependent cell-length changes of OHCs (triangles) and DCs (circles) in the 34 kHz region of the guinea-pig OoC. Magnitudes (a,c) and phases (b,d) are shown relative to the vibrations observed on either the BM (a,b) or RL (c,d) - see text. Mean stimulation level was 60 dB SPL per spectral component. Experiment GP19023.

Cell length changes

With certain assumptions (see [6]), the data from this study can be interpreted in terms of underlying cell- or structural length-changes, as illustrated in Fig. 4. At low frequencies, these data suggest that the OHCs become longer and shorter at the same time(s) that the BM moves 'down' and 'up' (Fig. 4b), and the RL moves 'up' and 'down' (Fig. 4d), respectively (where, by convention, 'down' means towards the ST, and 'up' means towards the SM). At higher frequencies, however, the OHC length-changes lag the BM's motion by up to 0.5 cycles in phase (Fig. 4b). Near CF, then, the OHCs become shorter and longer as the BM moves 'down' and 'up' (the opposite of the relationship at low frequencies). The DC length-changes appear smaller than, and anti-phasic to the OHC length-changes at all frequencies. Both DC and OHC length changes appear almost phase-locked (i.e. synchronous) with the RL's motion at all frequencies (Fig. 4d).

DISCUSSION

The raw data presented here are consistent with many of those presented in previous studies, both recent and historical (e.g. [2–4,6,8–11]). There are a few apparent discrepancies, including the relative 'phase-inversion' (i.e. lag vs lead) of the OHC/DC hotspot compared to that reported in our earlier (gerbil) work [6]. Such discrepancies may not be 'real', though; the OHC/DC phase-lags seen in the present report (and elsewhere) can easily be turned into phase-leads (as reported in [6]) simply by altering the viewing angle of the OCT's recording beam in the longitudinal dimension (as shown in [6], and verified by unpublished guinea-pig data from the current study). The same explanation may also account for our earlier inability [6] to directly 'see' anti-phasic motion at the two ends of the OHCs – this motion may only have been uncovered in the current study by observing both the macro- and micro-mechanical motion in an almost truly 'transverse' plane (cf. [6]).

Perhaps the most interesting finding of the present report is that every point within a single depth-wise column of the OoC vibrates in one of just two ~anti-phasic micro-mechanical sections, regardless of the stimulating frequency. In contrast to the macro-mechanical tuning of the BM (and/or RL), this micro-mechanical motion is not sharply tuned at all – it is low-pass filtered with a corner frequency more than two octaves below the macro-mechanical CF. Moreover, the micro-mechanical motion exhibits a significant delay with respect to the macro-mechanical motion of the BM, accumulating ~0.5 cycles of phase-lag by the point that the CF is reached. Both this phase-lag and the low-pass filtering prevent the micro-mechanics from evoking anti-phasic motion of the BM and RL above ~2 kHz in our data (in the 34 kHz CF region). While anti-phasic BM and RL motion can be evoked at higher frequencies using electrical stimuli both *in vivo* [5] and *in vitro* [18], it is only manifest at very low frequencies (including *dc*, i.e. 0 Hz) in responses to acoustic stimuli [4,7,8]. A further consequence of the ~0.5 cycles of delay in the micro-mechanics is that it makes both the macro- and micro-mechanical aspects of the responses synergistic at CF (when viewed from the RL; cf. [7]).

Our finding that the micro-mechanical motion is low-pass filtered contrasts with the predictions, implications or suggestions of many previous studies, although the observation that the micro-mechanics is not sharply tuned in its own right is not new (see Fig. 13 of [19] for an early indication of this). Our observation that the RL moves less than, and in anti-phase to the BM at very low frequencies leads us to suggest that one function of OHC electro-

motility [14] might be to *decrease* (i.e. to suppress or inhibit) the mechanical stimulation of the IHCs (and OHCs) at low-frequencies. Exactly how the OHCs operate in this (or any other) frequency region remains a mystery, however, and various aspects of their motion seem inconsistent with their being any part of a simple (positive or negative) feedback loop (cf. [20,21]).

One point not emphasized (so far) in this paper is that both the macro- and micro-mechanical aspects of the OoC's vibrations depend nonlinearly on the level of stimulation. This is most clearly demonstrated in Fig. 1b (where the BM compresses near-CF tones between ~30 and 80 dB SPL, and the RL compresses them even more), but it is apparent in Fig. 2 as well (where the *in vivo* micro-mechanical 'features' such as the hot- and/or cold-spots become progressively more apparent at lower SPLs). In agreement with numerous previous reports, we believe that most, if not all of this nonlinearity is introduced by the mechano-electrical transduction process [22] in the OHCs [23].

ACKNOWLEDGMENTS

Supported the Netherlands Organization for Scientific Research, ALW 823.02.018.

REFERENCES

1. van der Heijden, M. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 14548–14552.
2. Chen, F.; Zha, D.; Fridberger, A.; Zheng, J.; Choudhury, N.; Jacques, S.L.; Wang, R.K.; Shi, X.; Nuttall, A.L. *Nat. Neurosci.* **2011**, *14*, 770–774.
3. Lee, H.Y.; Raphael, P.D.; Park, J.; Ellerbee, A.K.; Applegate, B.E.; Oghalai, J.S. *Proc. Natl. Acad. Sci.* **2015**, *112*, 3128–3133.
4. Ren, T.; He, W.; Kemp, D. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 9910–9915.
5. Ren, T.; He, W.; Barr-Gillespie, P.G. *Nat. Commun.* **2016**, *7*, (10282) 1-9.
6. Cooper, N.P.; Vavakou, A.; van der Heijden, M. *Nat. Commun.* **2018**, *9*, 1–12.
7. He, W.; Kemp, D.; Ren, T. *eLife* **2018**, *7*, e37625.
8. Dewey, J.B.; Altoè, A.; Shera, C.A.; Applegate, B.E.; Oghalai, J.S. *Proc. Natl. Acad. Sci.* **2021**, *118*, e2025206118.
9. Békésy, G. von. *Experiments in Hearing*; McGraw-Hill: New York, **1960**.
10. Rhode, W.S. *J. Acoust. Soc. Am.* **1971**, *49*, 1218–1231.
11. Sellick, P.M.; Patuzzi, R.; Johnstone, B.M. *J. Acoust. Soc. Am.* **1982**, *72*, 131–141.
12. Versteegh, C.P.C.; van der Heijden, M. *J. Assoc. Res. Otolaryngol.* **2012**, *13*, 785–798.
13. Cooper, N.P.; van der Heijden, M. In: C Bergevin and S. Puria (eds) To Ear and Back Again: Advances in Auditory Biophysics; AIP Publishing, Melville, NY: Brock University, Canada, **2018**; Vol. 1965, pp. 080001.
14. Brownell, W.E.; Bader, C.R.; Bertrand, D.; de Ribaupierre, Y. *Science* **1985**, *227*, 194–196.
15. Richter, C.-P.; Dallos, P. In: AW Gummer (ed) The Biophysics of the Cochlea: Molecules to Models; World Scientific: Titisee, Germany, **2003**; pp. 278–284.
16. ter Kuile, E. *Pflüg. Arch. Ges Physiol.* **1900**, 146–157.
17. Mammano, F.; Ashmore, J.F. *Nature* **1993**, *365*, 838–841.
18. Jabeen, T.; Holt, J.C.; Becker, J.R.; Nam, J.-H. *Biophys. J.* **2020**, *119*, 314–325.
19. Neely, S.T. *J. Acoust. Soc. Am.* **1993**, *94*, 137–146.
20. Mountain, D.C.; Hubbard, A.E.; McMullen, T.A. In: E. de Boer and M.A. Viergever (eds) Mechanics of Hearing; Martinus Nijhoff: The Hague, **1983**; pp. 119–126.
21. Zwicker, E. *Biol. Cybern.* **1979**, *35*, 243–250.
22. Hudspeth, A.J.; Corey, D.P. *Proc. Natl. Acad. Sci.* **1977**, *74*, 2407–2411.
23. Kim, D.O. *Hear. Res.* **1986**, *22*, 105–114.

Modeling the Fine Structure of Ear Canal Pressure and Cochlear Microphonics in Response to a Pure Tone

Julien Meaud,[1, 2, a)] Haiqi Wen,[2, b)] and George Samaras[2, c)]

[1)] *Pettit Institute of Biosciences and Bioengineering, Georgia Institute of Technology, Atlanta GA 30332.*
[2)] *Woodruff School of Mechanical Engineering , Georgia Institute of Technology, Atlanta GA 30332.*

[a)] *Corresponding author: julien.meaud@me.gatech.edu*
[b)] *Electronic mail: hwen@gatech.edu*
[c)] *Electronic mail: gsamaras3@gatech.edu*

Abstract. Experimental measurements of the ear canal pressure (ECP) and cochlear microphonics at the round window (RWCM) in response to a pure tone exhibit quasi-periodic ripples in the amplitude. This fine structure, which is hypothesized to be due to a linear reflection mechanism caused by inhomogeneities in the properties of the cochlear partition, can be used to provide noninvasive information about cochlear function. In this work, a previously developed model of the gerbil cochlea is used to study the generation of the fine structure on the ECP and RWCM. The model includes cochlear roughness due to inhomogeneities in the properties of outer hair cells (OHCs) and a longitudinal cable model to represent the propagation of electrical currents in the cochlear ducts. The fine structure predicted in both ECP and RWCM is in line with experimental measurements in other species. Model simulations are used to derive the reflection components of the ECP and RWCM by taking the difference between predictions with a model with roughness and a smooth model. Results from simulations are analyzed to determine the cochlear regions that contribute to the RWCM fine structure.

INTRODUCTION

The cochlear microphonics (CM) is an extracellular potential generated due to the electrical response of OHCs [1]. It can be measured either at the round window (which will be called RWCM) or near the BM in the scala tympani (which will be called local CM or LCM). Charaziak et al. [2] recently demonstrated that, in chinchillas, the RWCM exhibits spectral ripples similar to what is observed in the ear canal pressure. Three different components of different latencies were identified. Both the short and the long latency components are generated by basal OHCs, either from direct stimulation by the forward traveling wave, or from the stimulation by the reverse wave that arises from coherent reflection near the peak of the forward wave [3]. However, Charaziak et al. demonstrated what the RWCM may also include contributions from a mid-latency component directly generated by OHCs located near the peak of the traveling wave. The presence of this mid-latency component suggests that the function of non-basal OHCs may be monitored by measuring the CM at the round window. In this work, a previously developed physiologically-motivated model of the cochlea [4, 5] is used to simulate the spectral ripples in the ear canal pressure (ECP) and RWCM and to identify the spatial regions that contribute to these spectral ripples.

PHYSIOLOGICALLY MOTIVATED MODEL OF THE COCHLEA WITH COCHLEAR ROUGHNESS AND LONGITUDINAL CABLES

We have recently developed and calibrated a physiologically-motivated model of the gerbil cochlea [4]. The model was calibrated based on both mechanical measurements (scala tympani (ST) fluid pressure) and electrical measurements (LCM) at a basal location (23.5 kHz best place). In this model, electrical longitudinal cables are included in the scalae of the cochlea to represent the spread of electrical currents within the ducts [6]; as shown in [4], these cables are essential in order to predict similar characteristics for the LCM as in measurements.

As in our recent work which focused on the generation of stimulus frequency otoacoustic emissions (SFOAEs) [5], the model includes roughness by adding small random perturbations in the value of the OHC electromechanical coupling coefficient, ε_3, which relates the electromotile force to the transmembrane potential. The size (standard deviation) of these random perturbations, ΔR, was adjusted to 0.3% to predict similar RWCM spectral ripples as in the experiments from Ref. [2] (peak to notch level differences of about 25 dB, with deep notches). Application of roughness in BM stiffness or in conductance of MET channel was also considered, but it did not have significantly change model predictions.

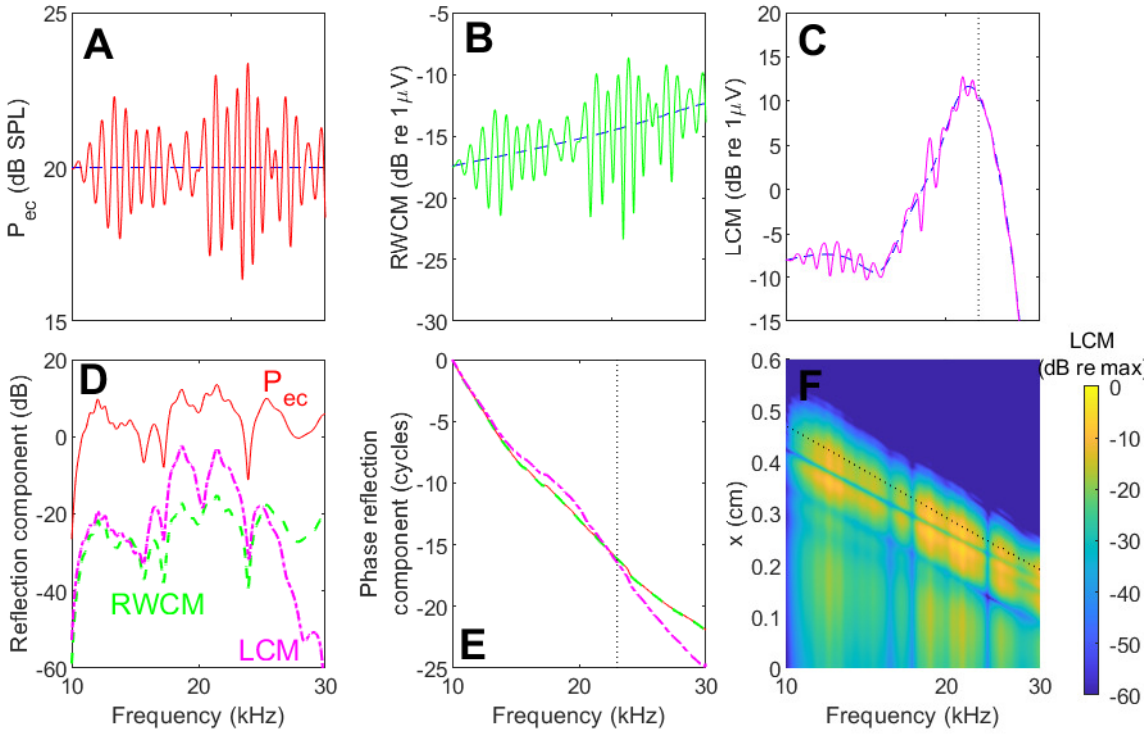


FIGURE 1. Model simulations of the ECP, RWCM, and LCM in response to 20 dB SPL tones. In C, the LCM is plotted at the 23.5 kHz BP. In A-C, the dashed lines correspond to the predictions from the smooth model and the solid lines to the predictions of the model with cochlear roughness. Panels D and E show the amplitude and phase of the reflection component in the ECP, RWCM and LCM. F. Amplitude of the LCM as a function of stimulus frequency and longitudinal position. The dotted line corresponds to the place frequency map of the smooth model.

RESULTS

Response of ECP and LCM to a pure tone

Figure 1 shows the response to a pure tone predicted by the model with cochlear roughness. Spectral ripples are observed in the predictions for the ECP, RWCM and LCM in panels A-C. Throughout this article, the model predictions for the ST potential as a function of the longitudinal position $x > 0$ will be called the LCM, while the predictions at the base, $x = 0$, will be called the RWCM.

The response in ECP (Fig. 1A) and RWCM (Fig. 1B) includes peaks and deep notches, as observed experimentally [2]. Taking the complex difference between simulations with cochlear roughness and without roughness yields the reflection component, as discussed in [5]. The reflection component of the ECP corresponds to the SFOAE. The reflection component of the RWCM may include contributions from a mid-latency component, generated near the peak region, and from a long-latency component, generated by basal OHCs due to their excitation by the reverse traveling wave. As seen in Fig. 1D, the reflection components of the ECP and the RWCM exhibits identical quasi-periodic variations in the amplitude (fine structure), while the reflection component of the LCM rises above the RWCM around the best frequency of the intracochlear location. Similarly, the phase of the SFOAE and reflection component of RWCM are identical, with a steep slope indicative of a long group delay (Fig. 1E). Analysis in [5] shows that the model SFOAE group delay is approximately $2\times$ the group delay of the BM response at its peak, while the LCM phase deviates from the SFOAE and RWCM phase. These group delays are consistent with the generation of a reverse wave near the peak location of the traveling wave. The response of the reflection component of the LCM is shown as a function of stimulus frequency and longitudinal position in panel F. This reflection component tends to maximum near the best place, but extends significantly towards the base of the cochlea.

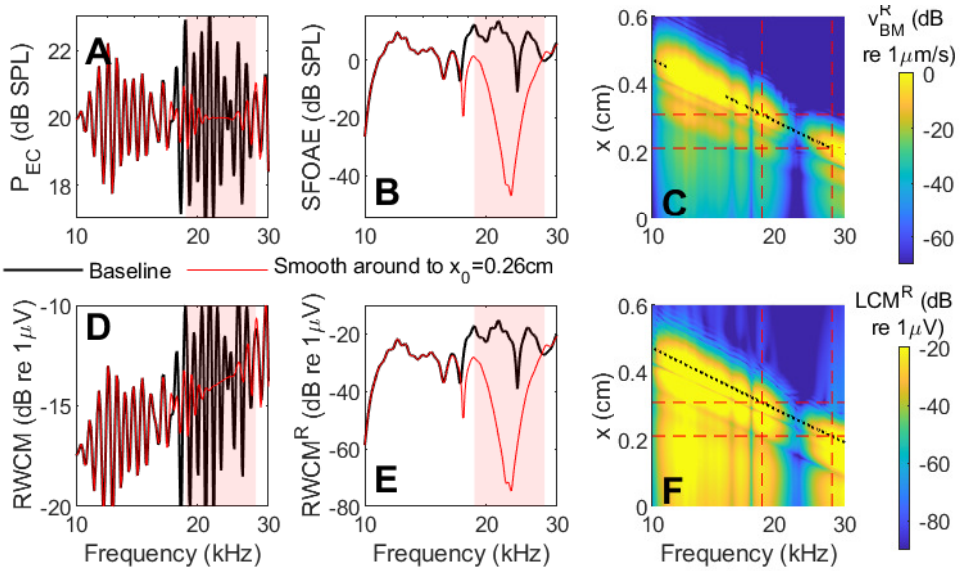


FIGURE 2. Effect of local smoothing of the model on predictions for the ECP and CM. A. ECP. B. SFOAE. C. Reflection component of BM velocity as a function of frequency and position. D. RWCM. E. Reflection component of RWCM. F. Reflection component of LCM as a function of frequency and position. In A-B and D-E, the black line corresponds to the baseline model with roughness applied at all longitudinal locations; the red line corresponds to the response in a model were roughness was eliminated around $x_0 = 0.26$ cm (23.5 kHz best place). In C and F, the black dotted line is the best place; the horizontal dashed lines are the boundaries of the zone without roughness. The vertical dashed lines in C and F, and the shaded area in A-B and D-E, are the boundary of the range of best frequency corresponding to the region without roughness.

Smoothing the cochlea around the peak region eliminates spectral ripples in both ECP and RWCM

To assess where the spectral ripples in the ECP and RWCM are generated, we first examined the effect of local smoothing of the cochlea by eliminating cochlear roughness around the 23.5 kHz best place (Fig. 2). As shown in [5], smoothing the cochlea eliminates the spectral ripples in the ECP around 23.5 kHz (Fig. 2A) because it eliminates the reflection component (*i.e.*, the SFOAE) in this frequency range (Fig. 2B). The effect of the local smoothing on the spectral ripples of the RWCM is nearly identical (Fig. 2D). These results show that the reflection component (both in ECP and RWCM, Fig. 2E) requires reflection from the peak region of the traveling wave. Panels C and F show how the reflection components of the BM velocity and of the CM are eliminated at all locations for frequencies around the BF of the center of the smooth region, while other frequencies are not affected.

Eliminating mechano-electrical transduction at the base eliminates the reflection component of RWCM

We then determined if non-basal OHCs contribute to the RWCM by setting the conductance of the OHC mechano-electrical transduction channels to zero at basal locations (up to $x_0 = 0.05$ cm). If the RWCM spectral ripples include contributions from a mid-latency component generated by OHCs located near the peak of the traveling wave, the reflection component of the RWCM is expected to remain significant when the basal OHCs are damaged.

Eliminating MET at basal locations has a very limited effect on both the overall amplitude and the spectral ripples in ECP (Fig. 3A). However, it reduces by about 30 dB the reflection component of the RWCM (Fig. 3B). The response of the reflection component of the CM essentially goes to zero at the basal locations where MET is eliminated (Fig. 3C). It implies that spectral ripples in the ECP originates from wavelets generated near the peak region while the RWCM (including the spectral ripples) is only generated by basal OHCs in this model.

In the current model, the space constant of the ST longitudinal cable, λ_{ST}, was tuned to obtain a tuned LCM

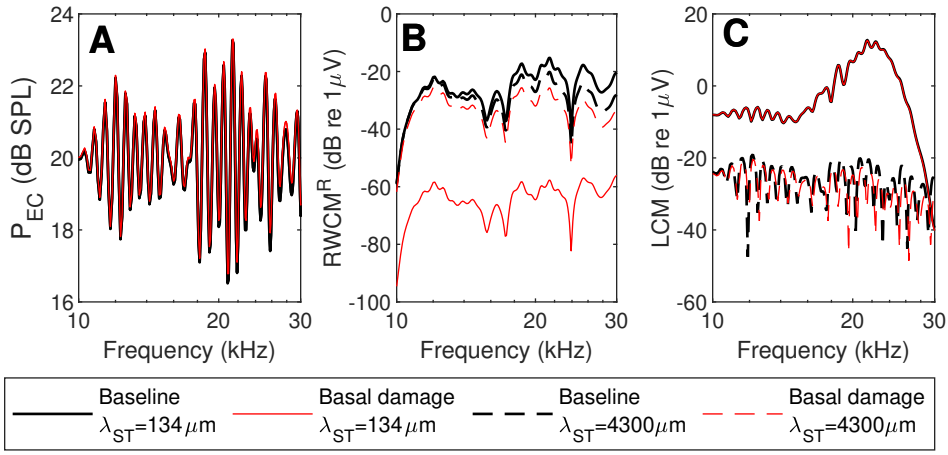


FIGURE 3. Effect of eliminating mechano-electrical transduction (MET) at basal locations (up to $x_0 = 0.05$cm) on ECP and RWCM in response to 20 dB SPL tones. A. ECP. B. Reflection component of RWCM. C. LCM. The black lines and red lines are the response of the baseline model and of the model without MET at basal locations, respectively. The solid lines correspond to the model used in Figs. 1 and 2; the dashed line correspond to a model with large ST cable space constant of 4300 μm.

response which is similar to experimental measurements in gerbil cochlea (see Fig. 2 in [4]). This space constant of 134 μm is considerably smaller than the large space constant of 4300 μm used by Charaziak et al. in the model they used to interpret their RWCM measurements. We also examined the response of the model when λ_{ST} is set to 4300 μm (dashed lines in Fig. 3). This increase in λ_{ST} causes very limited effects on the ECP spectral ripples and a a slight reduction in the magnitude of the reflection component of the RWCM (compare the black solid and dashed lines in Fig. 3B). The main difference is that in this model with a large λ_{ST} value, eliminating mechano-electrical transduction at the base only causes a small reduction in the reflection component of RWCM, implying that the spectral ripples include significant contributions from more apical locations. These contributions are expected to correspond to a mid-latency component since their delay would corresponds roughly to the delay of the forward traveling wave. However, as seen in Fig. 3C, this model with a large λ_{ST} exhibits a very flat LCM response which is inconsistent with the tuned LCM response observed at low stimulus level in experiments [4]. This flat response is observed because the ST longitudinal cables couple contributions from OHCs that are out of phase near to the peak of the traveling wave, causing a cancellation effect.

DISCUSSION

In this article, a physiologically-motivated model of the gerbil cochlea with cochlear roughness is used to simulate spectral ripples in the ECP and RWCM. The spectral ripples are qualitatively similar to measurements in chinchilla. However, while the measurements imply the presence of three components, model simulations are consistent with the presence of only two components: a component generated by basal OHCs by the forward traveling wave (which corresponds to the short latency component) and a component generated by basal OHCs by the reverse traveling wave that originates from the peak region (which corresponds to the long latency RWCM component).

Generation of a RWCM by OHCs located near the peak of the traveling wave is only observed when the space constant of the longitudinal cable of the scala tympani model is increased to a large value similar to what Charaziak et al. used. However, the use of such a large space constant makes the LCM predictions unrealistic compared to in vivo measurements. Perhaps coupling of the current cochlear model to of a three-dimensional (3D) model of the electrostatics in the scala tympani, such as [7, 8], is needed in order to simultaneously match the LCM and RWCM measurements. In these 3D models of the scala tympani potential, the LCM predicted near the BM includes primary local contributions while the LCM away from the BM includes non-local contributions

ACKNOWLEDGMENTS

This research was supported by National Institutes of Health (NIH) Grant Nos. R01 DC016114.

REFERENCES

1. R. B. Patuzzi, G. K. Yates, and B. M. Johnstone, "The origin of the low-frequency microphonic in the first cochlear turn of guinea-pig," Hearing research **39**, 177–188 (1989).
2. K. K. Charaziak, J. H. Siegel, and C. A. Shera, "Spectral ripples in round-window cochlear microphonics: evidence for multiple generation mechanisms," Journal of the Association for Research in Otolaryngology **19**, 401–419 (2018).
3. C. A. Shera and J. J. Guinan Jr, "Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian oaes," The Journal of the Acoustical Society of America **105**, 782–798 (1999).
4. T. Bowling, H. Wen, S. W. Meenderink, W. Dong, and J. Meaud, "Intracochlear distortion products are broadly generated by outer hair cells but their contributions to otoacoustic emissions are spatially restricted," Scientific Reports **11**, 1–14 (2021).
5. H. Wen and J. Meaud, "Link between stimulus frequency otoacoustic emission peaks and cochlear resonances," Journal of Acoustical Society of America.
6. S. Ramamoorthy, N. V. Deo, and K. Grosh, "A mechano-electro-acoustical model for the cochlea: response to acoustic stimuli," The Journal of the Acoustical Society of America **121**, 2758–2773 (2007).
7. P. D. Teal and G. Ni, "Finite element modelling of cochlear electrical coupling," The Journal of the Acoustical Society of America **140**, 2769–2779 (2016).
8. B. Frost and E. S. Olson, "Model of cochlear microphonic explores the tuning and magnitude of hair cell transduction current," Biophysical Journal **120**, 3550–3565 (2021).

Fluid focusing contributes to the BM vibration amplification by boosting the pressure

Renata Sisto[1, a)], Daniele Belardinelli[1, b)], Alessandro Altoè[2, c)], Christopher A. Shera[2, 3, d)], and Arturo Moleti[4, e)]

[1]*Department of Occupational and Environmental Medicine, Epidemiology and Hygiene, INAIL-National Research Centre for Safety and Prevention at Workplace, Monteporzio Catone (Rome), ITALY*
[2]*Caruso Department of Otolaryngology,, University of Southern California, Los Angeles, CA, USA.*
[3]*Department of Physics and Astronomy, University of Southern California, Los Angeles, CA, USA.*
[4]*Department of Physics –University of Rome 'Tor Vergata', and NAST Center, Rome, ITALY*

[a)]Corresponding author: r.sisto@inail.it
[b)]daniele.belardinelli@roma2.infn.it
[c)]altoe@usc.edu
[d)]christopher.shera@usc.edu
[e)]moleti@roma2.infn.it

Abstract. Two hydrodynamic effects are introduced in the standard transmission-line formalism, the focusing of the pressure and fluid velocity fields near the basilar membrane and the viscous damping at the fluid-basilar membrane interface, which significantly affect the cochlear response in the short-wave region. In this region, in which the wavelength is shorter than the cochlear duct height, only a layer of fluid of order of the wavelength is effectively involved in the traveling wave. This has been interpreted [8] as a reduced fluid contribution to the system inertia in the peak region, which is a viewpoint common to the 3-D FEM solutions. In this paper we propose an alternative approach, from a slightly different physical viewpoint. Invoking the fluid flux conservation along the traveling wave propagation direction, we can derive a rigorous propagation equation for the pressure integrated along the vertical axis. Consequently, the relation between the average pressure and the local pressure [4] at the fluid-BM interface can be written. The local pressure is amplified by a factor dependent on the local wavenumber with respect to the average pressure, a phenomenon we refer to as "fluid focusing", which plays a relevant role in the BM total amplification gain. This interpretation of the hydrodynamic boost to the pressure provides a physical justification to the strategy [10] of fitting the BM admittance with a polynomial containing both a conjugated pole and a zero. In the short-wave region, the sharp gradients of the velocity field yield a second important effect, a damping force on the BM motion, proportional to the local wavenumber, which stabilizes active models and shifts the peak of the response towards the base, with respect to the resonant place. This way, the peaked BM response is not that of a proper resonance, corresponding to a sharp maximum of the admittance, but rather a focusing-driven growth toward the resonant place, which is "aborted" before reaching it by the sharply increasing viscous losses. The large values of the wavenumber that ensure strong focusing are ultimately fueled, against viscosity, by the nonlinear OHC mechanism, hence the otherwise puzzling observation of a wide nonlinear gain dynamics with almost level-independent admittance.

INTRODUCTION

It is well-known that the WKB 1-d transmission line cochlear models require non-trivial parameter tuning to give accurate predictions in the short-wave region, where the wavelength of the traveling wave (TW) is shorter than the cochlear duct height *H*, and 2-d, 3-d fluid effect must be included in the model [4]. The main effects are the focusing of the TW in a thin fluid layer close to the basilar membrane (BM) and the strong viscous damping of the BM motion associated with the sharp gradients of the velocity field in the BM-fluid interface region [6]. In the short-wave region,

only a layer of fluid of thickness of the order of the wavelength λ is involved in the traveling wave phenomenon. This occurrence has been interpreted [8] as a reduced fluid inertia. FEM models commonly use this strategy to include 3-d hydrodynamic effects without solving the equations over the whole volume.

In this paper, we propose a different approach, based on a variant of the 2-d argument introduced by Shera et al. [4], who showed that the local pressure forcing the transverse BM motion is amplified by a factor α with respect to its average along *z*. In the long-wave limit the pressure at the BM–fluid interface and the average pressure are approximately coincident. The more the wave vector increases, the more the pressure at the fluid-BM interface increases. This pressure amplification is the phenomenon we refer to as "fluid focusing". The boost to the pressure, given by the fluid hydrodynamics, plays a relevant role as a strong contribution to the BM total amplification gain.

The interpretation of the hydrodynamic effect at the peak as a boost to the pressure supports the strategy proposed by Zweig [10] of fitting the BM admittance of 1-d transmission line models with a polynomial ratio containing a conjugated pole and a zero. Indeed, the BM velocity amplification in the peak region is not only related to the poles of the admittance but also to its zeros, which are sensitive to the local amplification of the forcing pressure. For TW power to spatially increase while maintaining stability, the amplifier ultimately manifests as a "zero" in the BM admittance, or something mathematically equivalent. Taking into account both focusing and fluid viscosity, the experimentally observed [2] approximate level-invariance of the BM admittance can be compatible with the notion of a strong active power input from the outer hair cells (OHC) anti-damping force [6].

MODEL

1. Average pressure method

Invoking the volume conservation for the incompressible fluid, one can derive a rigorous propagation equation along the longitudinal direction *x* for the differential pressure *p* averaged along the vertical axis *z* [4, 6]. We will start from the 2-d formulation, which is formally simpler than, and almost equivalent to, the 3-d formulation. The basic equations and boundary conditions for an incompressible viscous fluid in a simplified 2-d box model of length *L* and height *H*, are:

$$\vec{\nabla}\cdot\vec{u}=0 \qquad \text{Incompressibility equation,} \tag{1}$$

$$\rho\frac{\partial\vec{u}}{\partial t}=-\vec{\nabla}p+\mu_b\nabla^2\vec{u} \qquad \text{Linear Navier-Stokes equation,} \tag{2}$$

$$\vec{u}(z=H)=(0,0) \qquad \text{Boundary condition at } z=H, \tag{3}$$

where ρ and μ_b are the fluid density and viscosity coefficient, *x*, *z* are the cochlear longitudinal and vertical axes, orthogonal to the BM. Let *u* and *w* be the *x* and *z* components of the fluid velocity: $\vec{u}\equiv(u,w)$. The adhesion boundary condition at the BM-fluid interface requires:

$$\vec{u}(z=\xi\cong 0)=\left(0,\ \frac{\partial\xi}{\partial t}\right), \text{ i.e., } \begin{bmatrix} u=0 \\ w=\dot{\xi} \end{bmatrix}, \tag{4}$$

where $\dot{\xi}$ is the BM velocity. $u = 0$ is the classical no-slip condition at the interface. Fluid incompressibility implies:

$$\frac{\partial u}{\partial x}+\frac{\partial w}{\partial z}=0\ . \tag{5}$$

Integrating along *z* one gets:

$$\int_\xi^H\frac{\partial u}{\partial x}dz=-\int_\xi^H\frac{\partial w}{\partial z}dz=-(w(z=H)-w(z=\xi))=\frac{\partial\xi}{\partial t}\ . \tag{6}$$

On the other hand:

$$\int_\xi^H\frac{\partial u}{\partial x}dz=\frac{\partial}{\partial x}\int_\xi^H u dz\ , \tag{7}$$

and, consequently:

$$\frac{\partial}{\partial x}\int_{\xi}^{H} u dz = \frac{\partial \xi}{\partial t}\,. \tag{8}$$

Equation (8) is an exact flux conservation equation, and the starting point of our reasoning. From Eq. (8) the propagation equation for the pressure averaged over the *z* axis is straightforwardly obtained. Although the integration starts from $|\xi| \ll H$, in the following we will assume $\xi = 0$, for simplicity. Taking the time derivative:

$$\int_{o}^{H} dz \frac{\partial}{\partial x}\frac{\partial u}{\partial t} = \ddot{\xi}\,, \tag{9}$$

and, neglecting the viscous terms in Eq. (2):

$$\int_{o}^{H} dz \frac{\partial^2 p}{\partial x^2} = -2\rho\,\ddot{\xi}\,. \tag{10}$$

where the factor 2 is due to considering the differential pressure, which is the only part of the pressure involved in the slow TW. Using the separation of variables technique, the pressure can be written [3] as a function of *x* multiplied by a vertical profile that is a hyperbolic cosine function of *z*:

$$p(x,z,\omega) = f(x,\omega)\frac{\cosh\big(k(H-z)\big)}{\cosh(kH)} e^{-i\int_0^x dx' k(x',\omega)}\,. \tag{11}$$

By integrating over the vertical direction, the following relation can be obtained between the pressure, averaged over the vertical direction, and the local pressure close to the BM [4]:

$$\bar{p} = \frac{p(x,0,\omega)}{\alpha}, \tag{12}$$

where

$$\alpha = \frac{kH}{\tanh(kH)}\,. \tag{13}$$

This function, introduced by Duifhuis [3], increases sharply, for each frequency component of the TW, approaching its resonant place, as the wavelength becomes shorter than the duct height. In the limit $|k|H \gg 1$ the factor α is well approximated by *kH*. The propagation equation for the focused differential pressure *p*(*x*,0) is therefore:

$$\frac{\partial^2 p(x,0)}{\partial x^2} = -2\rho\frac{\alpha}{H}\ddot{\xi} = -2i\omega\rho\frac{\alpha}{H}\dot{\xi}\,. \tag{14}$$

Let us recall here the second basic cochlear equation, i.e., the equation describing the dynamics of the single nonlinear oscillator stimulated by the fluid differential pressure.

$$\sigma_{bm}\big(\ddot{\xi} + \Gamma(x,\xi,\dot{\xi})\dot{\xi} + \omega_{bm}^2(x,\xi,\dot{\xi})\xi\big) = p(x,0)\,, \tag{15}$$

where σ_{bm} is the BM surface density, and ω_{bm} is the BM local resonant frequency. Equations (14) and (15) relate the BM kinematic variables to the pressure and to its second spatial derivative. In Eq. (15), the damping coefficient and resonance frequency may be nonlinear functions of the BM displacement and velocity. For example, the function $\Gamma(x,\xi,\dot{\xi})$ may consist of a linear passive damping term, $\Gamma_p(x)$, assumed to be a scale-invariant function of *x* only, and of an active nonlinear term $\Gamma_a(x,\xi,\dot{\xi})$, which schematizes the anti-damping effect of the OHC forces. This nonlinear dependence can be neglected, for simplicity, considering a set of linear models with different effectiveness of the active mechanism, roughly representing the response of the system at different stimulus levels.

If the fluid viscosity is taken into account, a friction force acting along the vertical direction on the BM–fluid interface has to be added to the right-hand side of Eq. (15), as in [6]:

$$F_{fv}(x,0) = -4\mu_b\frac{\partial w}{\partial z} = -4\mu_b k\dot{\xi}\,, \tag{16}$$

where the boundary condition Eq. (5) has been used, and a multiplicative factor 2 is due to the presence of two identical forces coherently acting on the BM in the two scalae. Eq. (15) becomes:

$$\sigma_{bm}\left(\ddot{\xi}+\Gamma(x,\xi,\dot{\xi})\dot{\xi}+4\frac{\mu_b}{\sigma_{bm}}k\dot{\xi}+\omega_{bm}^2(x,\xi,\dot{\xi})\xi\right)=p(x,0)\,. \tag{17}$$

From Eq. (17), written in the frequency domain, one can define the BM admittance as the ratio between the BM velocity and the local focused differential pressure:

$$\frac{\dot{\xi}}{p}=Y_{bm}=\frac{1}{\sigma_{bm}}\frac{i\omega}{\left(\omega_{bm}(x)^2-\omega^2+i\omega\left(\Gamma+\frac{4\mu_b}{\sigma_{bm}}k\right)\right)}\,. \tag{18}$$

Equation (14) synthesizes the longitudinal transmission and vertical focusing properties of the TW in a local relation among pressure, BM velocity, frequency and wavenumber, dependent on the dynamics of an incompressible ideal fluid and on the adhesion boundary conditions. Equation (18) is a second relation among the same physical quantities, dependent on the physics of the local oscillator and on the viscous force at the BM-fluid interface. Putting them together, one gets the TW equation for the pressure, and a self-consistent local relation between the wave vector and the frequency, using an iterative procedure [6]. Indeed, considering that:

$$\frac{\partial^2 p(x,0)}{\partial x^2}\cong -k^2 p(x,0)\,, \tag{19}$$

one gets an implicit relation (a and Y_{bm} are both functions of k) that yields the local value of the wave vector as a function of the parameters of the local oscillator (and of viscosity):

$$k^2=-2i\omega\rho\frac{\alpha}{H}Y_{bm}\,. \tag{20}$$

In the limit $|k|H\gg 1$, one finds the short-wave relation between k and the admittance [5]:

$$k=-2i\omega\rho Y_{bm}=\frac{1}{\sigma_{bm}}\frac{2\rho\omega^2}{\left(\omega_{bm}(x)^2-\omega^2+i\omega\left(\Gamma+\frac{4\mu_b}{\sigma_{bm}}k\right)\right)}\,. \tag{21}$$

2. Average Lagrangian method

The dispersion relation or, in other words, the relation between the wave vector and the frequency can be alternatively determined by following another strategy, proposed by Steele and Taber [8]. In Whitham [9], the validity of the variational principle is demonstrated for the average Lagrangian, which, for systems in which the wave vector is a smoothly varying function, can be factorized as:

$$\mathcal{L}(\omega,k,W)=\mathcal{F}(\omega,k)W^2\,, \tag{22}$$

where W is the vibration amplitude, and the dispersion relation is given by imposing:

$$\mathcal{F}(\omega,k)=0\,. \tag{23}$$

Let us apply this strategy to the previously described 2-d model, to get the dispersion relation, or, in other words, the relation between the wave vector and the BM admittance, in the short-wave limit. In this sense, the average pressure method and the average Lagrangian method are almost equivalent, the difference being mainly related to practical computational aspects and "philosophical" interpretational preferences.

The Lagrangian of the conservative forces is made up by the kinetic energy of the fluid on both sides of the OoC, the kinetic energy of the BM, and the potential energy of the BM. The kinetic energy of the fluid (per unit surface), due to focusing, only involves an effective layer of fluid of effective height:

$$H_{eff} = \frac{H}{\alpha}, \tag{24}$$

Therefore, the average Lagrangian for the conservative force can be written as:

$$\mathcal{L}(\omega, k, W) = \omega^2 \rho H_{eff} W^2 + \frac{1}{2}\omega^2 \rho_{bm} d W^2 - \frac{1}{2} K W^2 \ , \tag{25}$$

where *d* and *K* are the thickness and stiffness per unit surface of the BM:

$$K(x) = \rho_{bm} d \omega_{bm}^2(x) \equiv \sigma_{bm} \omega_{bm}^2(x) \ . \tag{26}$$

Following the strategy by Steele and Taber [8], a Lagrangian term may be written also for dissipative forces. The BM damping gives a contribution, which may also include the anti-damping effect of the OHCs:

$$D_{diss} = -\frac{1}{2} i\omega \Gamma \rho_{bm} d W^2 \ . \tag{27}$$

Here, we also include the friction force at the fluid–BM interface, Eq. (16), giving to the Lagrangian the contribution:

$$D_{fv} = -\frac{1}{2} 4 i\omega \frac{\mu_b}{\sigma_{bm}} k \rho_{bm} d W^2 \ . \tag{28}$$

A term coming from the viscosity in the fluid could also be kept into account. In Steele and Taber [8], the wave vector β is introduced for a viscous fluid:

$$\beta^2 = k^2 + \frac{i\omega\rho}{\mu_b} \ . \tag{29}$$

and the effective height is calculated (see their Eq. 25.b). In the limit $\frac{k}{\beta} \approx 0$, which typically holds in the mid-high frequency range:

$$H_{eff\_v} = H_{eff} \left[1 + \frac{1}{\beta H}\right] + \frac{1}{\beta} \ , \tag{30}$$

Summing up all the terms contributing to the Lagrangian and considering the effective height of the viscous fluid:

$$\mathcal{L}(\omega, k, W) = \omega^2 \rho H_{eff\_v} W^2 + \frac{1}{2}\omega^2 \rho_{bm} d W^2 - \frac{1}{2} i\omega \left(\Gamma + 4\frac{\mu_b}{\sigma_{bm}} k\right) \rho_{bm} d W^2 - \frac{1}{2} K W^2 \ . \tag{31}$$

Using Eq. (31) into Eq. (23):

$$\mathcal{F}(\omega, k, \mu) = \omega^2 \rho H_{eff\_v} + \frac{1}{2}\omega^2 \rho_{bm} d - \frac{1}{2} i\omega \left(\Gamma + 4\frac{\mu_b}{\sigma_{bm}} k\right) \rho_{bm} d - \frac{1}{2} \rho_{bm} d \omega_{bm}^2(x) = 0 \ , \tag{32}$$

assuming $\rho = \rho_{bm} \equiv \sigma_{bm}/d$, and substituting the viscous effective height, one gets, in the short-wave limit:

$$\frac{d\left(\omega_{bm}^2(x) - \omega^2 + i\omega\left(\Gamma + 4\frac{\mu_b}{\sigma_{bm}} k\right)\right)}{2\omega^2} = H_{eff} \left[1 + \frac{1}{\beta H}\right] + \frac{1}{\beta} = \frac{H}{\alpha} + \frac{1}{\beta} + \frac{1}{\alpha\beta} \cong \frac{H}{\alpha} + \frac{1}{\beta}, \tag{33}$$

$$k \cong \frac{1}{\sigma_{bm}} \frac{2\rho\omega^2}{\left(\omega_{bm}^2(x) - \omega^2 + i\omega\left(\Gamma + 4\frac{\mu_b}{\sigma_{bm}} k\right)\right)} \left(1 + \frac{k}{\beta}\right). \tag{34}$$

In the mid-high frequency range, Eq. (34) represents therefore the same dispersion relation as that of Eqs.(18) and (21), confirming the approximate equivalence of the two formulations.

DISCUSSION

Although the two points of view, that attributing the BM vibration amplification to the focused pressure amplitude at the BM-fluid interface and that attributing it to the reduced system inertia, lead to the same quantitative results, their physical meanings are different. The quantity α, which is a key element of both formulations, has a double meaning. It approximately represents the fraction of the fluid height that is involved in the traveling wave phenomenon and, consequently, it may be interpreted as the factor locally expressing the reduction of the fluid inertia in the short-wave region. On the other hand, α represents the amplification of the pressure, due to the increase of the real part of the wave vector k in the peak region. In a classical locally-resonant transmission line, the peaked BM response is entirely due to the local properties of the admittance. Here, it is also a consequence of the local focusing of the force applied to the BM, i.e., the differential fluid pressure. Indirectly, focusing is still due to the resonant nature of the BM, which implies a sharp increase of the wavenumber approaching the resonant place.

The Lagrangian method considers the fluid and the BM as a closed system, and immediately suggests the idea that hydrodynamic focusing in the peak region reduces the effective height of fluid involved in the TW and, consequently, the idea that the variable effective inertia of the dynamical system is the reason for the boost of the velocity.

The average pressure method, related to the flux conservation equation, focuses on the BM as a dynamical system forced by an external pressure force, and, therefore, fosters the idea that the fluid focusing phenomenon boosts the forcing term in the resonant oscillator equation of motion, and, consequently, the BM motion, without changing the BM admittance.

We see that the two representations are almost equivalent, but one of them may be more suitable than the other for describing in an intuitive way important features of the cochlear mechanics, as the role of the active forces within the OoC, the different motion of different structures within the OoC, the role of the delays (phase-shifts) in the activation of the OoC active elements, their nonlinearity, the power balance along the TW propagation, etc.

For example, in [6], a 2-d WKB model including fluid focusing and viscous damping at the fluid-BM interface was used to show that, similarly to what happens in a 3-d FE viscous model, several features of the BM and pressure experimental response are easily predicted and interpreted. In the WKB and FE models, a peak-gain variation range of about 30-40 dB was obtained by varying the anti-damping term strength of a set of linear systems. Both the BM velocity and pressure response showed a peaked behavior, associated with fluid focusing, with almost invariant admittance. The maximum of the peak was increasingly shifted towards the apical region with increasing anti-damping. A similar phenomenology was observed in experimental studies measuring the pressure near the BM and the BM velocity [2] at different stimulus levels. The role of the anti-damping active forces within the OoC is intuitively explained in the proposed framework as follows: approaching the resonant place, the active OHC mechanism increases the local wavenumber boosting the pressure on the BM, but the friction force at the BM–fluid interface, proportional to the wavenumber, also increases. The overall effect is that the peak admittance is almost level-invariant. For the same reason, the cochlear effective tuning Q, although significantly dependent on the stimulus level, also exhibits a moderate variability range, compared to that of the BM gain, contrary to what would be expected for a standard resonant system [7]. Another important consequence of the dependence of the admittance on k consists in the fact that the traveling wave is extinguished due to viscous damping in a region where the imaginary part of the wave vector is not yet comparable to the real part. In other words, the nominal resonant place is not reached by the TW, but only approached, more closely for the more active models, which more effectively counteract the viscous losses allowing to reach larger values of k, and more effectively boosting the pressure.

Strictly speaking, the amplification phenomenology observed in the peak region is not that of a classical resonance, but the large values of the wavenumber that sustain a large pressure boost are possible only if the OHC active anti-damping force is effective enough, i.e., at low stimulus levels. This way, one may understand how nonlinear anti-damping forces may yield nonlinear BM-velocity and pressure response with almost level-invariant admittance. The small experimental nonlinear variation of the admittance [2] would represent a puzzle if the dependence on wavenumber viscosity were not taken into account.

The pressure amplification due to the focusing in the peak region suggests that the BM transfer function can be naturally approximated by the “conjugated pole plus zero” function, as proposed by Zweig [10, 11] and developed by Altoè and Shera [1]. The zero at the numerator of the BM transfer function is related to the forcing term (i.e. the pressure). In particular, a real zero model seems to perform particularly well for fitting the BM response. In the peak region, the BM gain can be interpreted as the sum of two contributions, one coming from the denominator and related to the increase of the effective tuning Q and the other coming from the numerator or, in other words, from the increase of the term proportional to the complex frequency in the numerator.

Going into finer details, 3-d FEM and WKB solutions show that a thin layer of fluid is actually moving with the BM, and that the viscous force may be considered as applied to the external surface of this layer, and through it, to the BM. The thickness of this layer is determined by viscosity and it can be significantly smaller than the wavelength of the longitudinal TW propagation. Therefore, when considering the effective inertia of the oscillator, that of the thin layer, and the local dependence of its thickness, should be considered also in our approach, in addition to, and affecting the size of, the pressure boost effect.

CONCLUSION

Pressure focusing and viscous damping in the peak region are discussed using two different approaches to cochlear modeling, yielding almost equivalent results. The role of viscosity is that of a stabilizing force able to cause the total dissipation of the TW power well before the resonant tonotopic place is reached. The pressure focusing amplification, due to a geometrical form factor, has also been interpreted in terms of a decrease of the effective system inertia. This interpretation, although almost equivalent in terms of numerical results, is a conceptually different one. The pressure focusing phenomenon explains why the best fit to the BM transfer function suggests to introduce a "conjugated poles plus zero" function. In other words, the amplification is not only dependent on the *Q* of the local resonant oscillator but also on the increase with increasing wavenumber of the external force (i.e., the pressure) associated with focusing [4]. In this framework, the important stabilizing role of the viscous force on the BM must be acknowledged, to reconcile the concept of a nonlinear anti-damping OHC force as the basic source of the high-gain nonlinear BM response with the experimental evidence for level-independent admittance [6].

ACKNOWLEDGMENTS

This work was partly supported by grants NIH/NIDCD R21 DC019712 and R01 DC003687.

REFERENCES

1. Altoè A, Shera C.A., (2020). Phys Rev Res 2, 013218
2. Dong, W., and Olson, E.S. (2013). Biophys. J. 105, 1067
3. Duifhuis H.E., (1988). Cochlear macromechanics. Auditory Function: Neurological Bases for Hearing, 189
4. Shera, C.A., Tubis, A., and Talmadge, C. L. (2005). J. Acoust. Soc. Am. 118, 287
5. Siebert, W.M. (1974). J. Acoust. Soc. Am. 56, 594
6. Sisto R., Belardinelli D., Moleti A. (2021) J. Acoust. Soc. Am. 150, 4283
7. Sisto, R., Moleti, A., and Altoe`, A. (2015). J. Acoust. Soc. Am. 138, EL155
8. Steele C.R. and Taber L. (1979), J. Acoust. Soc. Am. 65, 1007
9. Whitham G.B. (1970), J. Fluid. Mech. 44, part 2, 373
10. Zweig G. (2015), J Acoust Soc Am. 138, 1102
11. Zweig G. (2016), J Acoust Soc Am. 139, 2561

My 50 years of cochlear modeling

Jont B. Allen[a)]

University of Illinois.edu,
ECE, Rm 3062,
Urbana Illinois, USA
Version 1.4a–6/25/22;

[a)]*Electronic mail: jontallen@ieee.org*

Abstract. The goal of this presentation is multi-fold: The primary goal is to discuss my present understanding of cochlear function. A secondary goal to review the history of my work in modeling the cochlea and the role of four close friends: Egbert de Boer, Steve Neely, Paul Fahey and George Zweig.
Regarding my present understanding of how the cochlea works, including experimental data on cochlear function, the tympanic membrane, middle ear, basilar and tectorial membranes, inner and outer hair cells, auditory nerve, and the cochlear amplifier. These views have been greatly sharpened by looking back and piecing this complex puzzle together. A great deal of progress has been made in the last 50 years.
Conclusions: My recent review of neural tuning curve data from 1985, using nonlinear (NL) distortion product generation, has revealed a deeper understanding of cochlear function. The most important, and surprising result, is that the cochlea is much more linear in its filtering properties than I previously assumed. When the suppressor frequency f_s is at least 1/2 octave lower than the characteristic frequency f_{cf}), we call this "low-side" suppression. This BM linearity is seen for "low-side" suppressors below 65 [dB-SPL]Fahey and Allen [23]. Above 65 [dB] the suppression dominates with a slope greater than 1 [dB/dB]. The obvious explaination is that the threshold of excitation of the inner hair cells and of suppression threshold of the outer hair cells, which control the NL suppression, are the same. This is unexpected result, because low-side suppression on the basilar memrane has a 20-30 [dB] higher threshold[20,24].
This implies that low-side suppression and loudness recruitment (the reduced loudness of low-intensity sounds in the hearing-impaired ear) are closely related (i.e., are likely the same phenomena). The ramifications of this observation seem important as they will impact the diagnosis of cochlear hearing loss, and the fitting of hearing aids. In other words, two-tone suppression acts like an automatic gain control, elevating the loudness threshold, with very little distortion. This unexpected conclusion, about the linear nature of the cochlear response for low-side suppressors, is surprising.

INTRODUCTION

The path to new friends: After obtaining my MS in Electrical Engineering from the Univ. Penn on modeling the electrical thawing of frozen dog kidney (1968), and PhD on modeling a 1 [m] $\times$ 1 [cm] 18 kC° plasma arc-jet (1970). By 1971 I joined AT&T Bell Labs and in 1975 I was transferred to the *Acoustics Research Department* in Murray Hill NJ.

I was soon introduced to visiting scholar Egbert de Boer, who was widely known for his significant wide-range of hearing-science contributions. Egbert soon informed me that I would start modeling the cochlea (inner ear). I took the "hint" from both him and my highly knowledgeable supervisor David Berkley, who had published on this topic. I was immediately in the able hands of Mohan Sondhi who greatly upgraded my otherwise meager understanding of the necessary mathematics. Several papers followed[6,12,19,31–33].

Researching the inner ear: Bell Lab's Acoustics Research Department was mostly involved in automatic speech recognition (ASR) and various acoustic applications (speaker phone and electret microphones), but historically there was always one person doing hearing research, and I quickly morphed into that lucky soul. At Bell Labs, limits were only determined by one's imagination.

The next major event in my life was the arrival of Steven Neely from Cal Tech, who was assigned to work with me, and work we did. This turned into a life-long seminal friendship, with many important joint publications. Within a few years (circa 1982) I was setting up a lab at Columbia University in the Black Building (168th St, NYC), to record from the auditory nerve of the cat. For this work University of Scranton physicist Paul Fahey joined in, resulting in another life-long friendship/collaboration, along with, in our view, several key publications on how the nonlinear cochlea processes sound.

Cochlear modeling was soon followed by modeling the middle ear. This interest led to a PhD relationship with Sunil Puria, also from Columbia University. Today Sunil is one of the leading, if not the number-one expert, on middle ear science. Throughout this time up to the 2020 MoH meeting (Canada), I was in a continuous fun conversations with de Boer, although we almost never agreed. These years have been carefully documented due to the extensive work of Chris Shera, who designed the MoH website, with all the MoH publications freely available.

Around 1982-83, AT&T began the highly successful development of multi-band wide dynamic range hearing aids, which within five years was sold to the ReSound Corporation[8,10]. ReSound is one of the few remaining large hearing aid companies in a highly competitive industry. My role at ReSound molded my future in many ways, largely due to to its founder, and another very close friend, Stanford's ENT Dr. Rodney Perkins, and also via another equally special friend, Mead Killion the founder of Etymotic Research.

Following the amazing experience with ReSound, I returned to Bell Labs research to continued my research on modeling the middle ear and cochlea, and quantifying auditory speech decoding features[16,26,27]. The publications that best document these novel views are AF-92[14] and FA-85[23].

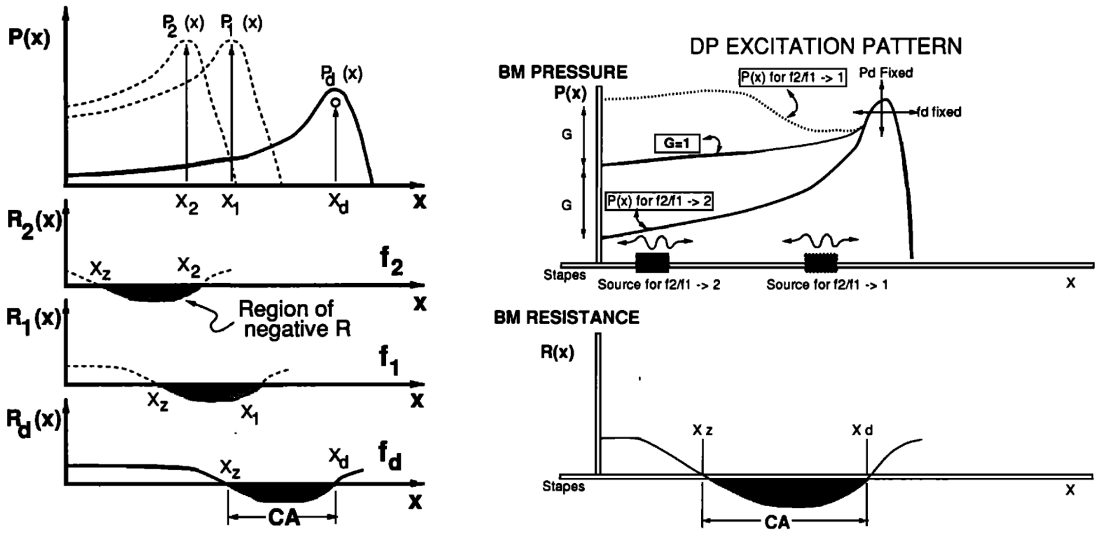


FIGURE 1. *These figures from*[14] *(1992) describe the protocol for measuring the cochlear amplifier gain on the basilar membrane. In the LEFT panel two tones at frequencies $f1 < f2$ excite the cochlea at locations X_1, X_2. Such a pair of tones generate a nonlinear (NL) distortion product at frequency $f_d = 2f_1 - f_2$ with $f_2 > f_1$, propagated to $X_d > X_1 > X_2$. The primary levels of the two tones $P_1(f_1), P_2(f_2)$ are adjusted so that the distortion product pressure P_d maximally excites a neuron having its best frequency at $f_d < f_1 < f_2$. In this way the source at $X_d(f_d)$ is generated along the basilar membrane, basial to the neuron's best frequency, as show in the RIGHT panel. If we assume that there is a Rayleigh-reciprocal*[13] *active region of negative resistance, corresponding to a cochlear amplifier generation site shown in black, then as the signal from the distortion product propagates back to the ear canal, it will be greatly amplified by the active region. Thus the change in $P_{ec}(f_d)$ as a function of f_2/f_1 would be amplified by twice the gain of the CA, because it would pass through the shaded region twice, once on the way in, and again on the way out. RIGHT: Since we know the DP level at the auditory nerve, and its ear canal pressure, we can determine the gain of the CA as the pressure ratio.*

Experiment I (AF-92)

The 1992 experiment (AF-92) is significant for two reasons: First it rigorously defines the term *cochlear amplifier* (CA). It does this by introducing an experimental protocol to measure the magnitude of the CA's cycle by cycle power gain, thus quantifying the gain of the CA.

The experimental paradigm: The procedure begins by searched for a neuron, and once found, measuring its tuning curve. A fixed tone of frequency f_d having an ear canal pressure of $P_{ec}(f_d)$ is then presented 6 [dB] above the neuron's threshold, at the best (characteristic) frequency (BF) of the attached neuron.

Next we moved the source at frequency f_d from the ear canal, to a variable location X_2 on the BM. The ear canal pressure remained approximately the same, as previously demonstrated (Fig. 7) Fahey and Allen [23] (1985). This is what one would be predicted from a model of the traveling wave, as long as there are no large ear canal standing waves.

Thus in this scheme we have swapped the ear canal pressure source with a source on the basilar membrane, at a variable location X_2, generated by a two-tone distortion product having an intensity 6 dB above the neurons threshold at its best frequency f_d.

We then verified that the ear canal pressure is approximately the same value as when it was presented from the ear canal, as we had previously demonstrated[23], but now verified as part of our protocol. As shown on the LEFT panel of Fig. 1, an internally generated distortion product tone at f_d the neurons best frequency f_{cf} generated at $X_2(f_2)$ by the NL action of the outer hair cells. This distortion product signal at frequency f_d is generated on the basilar membrane (BM) at $X_2(f_2) < X_1(f_1) < X_d(f_d)$. This source may be moved along the BM by changing the frequencies f_2 and f_1, where $f_1 = \frac{(f_d+f_2)}{2}$. For example if $f_d = 1, f_2 = 3$ then $f_1 = 2$.

Thus as the two primary frequencies change, their levels were adjusted to keep the ear canal pressure $P_{ec}(f_d)$ at f_d fixed at the neural tuning curve's threshold. The magnitude of this adjustment depends on the CA gain. We shall show that it must be twice the gain of the CA, because this is the round-trip (in and out) gain.

In summary: By varying the DP source location $X_2(f_d)$ along the BM (left-most dashed line), while simultaneously keeping the DP frequency f_d and pressure $P_d(f_d)$ fixed, (RIGHT panel of Fig. 1), we may determine the acoustic CA gain coming from $X_2(f_2)$. The DP pressure P_d propagates from X_2, where it is generated, to $X_d(f_d)$, where it is detected by the target neuron, having its $f_{cf} = f_d$. The canal pressure $P_{ec}(f_d)$ was also held constant as $X_2(f_2)$ is varied.

As the source at X_2 moves through the CA region (dark shaded region), it would be amplified, causing the ear canal pressure P_{ec} to vary by more than the CA gain, as a function of $X_2(f_2)$.

Results: As reported by AF-92, the CA gain was less than 6 dB. As we shall discuss in the next section on FA-85, there was a large amount of nonlinear compression around the best frequency of the tuning curve. We argue that this NL compression is not dependent on the CA gain, since it is only a few [dB]. It is exactly this difference we wish to emphasize.

Discussion

The source transducer design for these experiments was exceptionally linear, with an acoustic source impedance close to that of the ear canal[5]. Thus reflections (standing waves) in the cochlear were measurably negligible. No nonlinear artifacts were observable in an acoustic cavity using this transducer.

Given the AF-92 measurement paradigm, we compared the ear canal pressure level $P_{ec}(f_d)$ to the pressure propagated from the BM source $P_{ec}(f_d)$ on the BM at $X_d(f_d)$, where both are set to the neuron's threshold. If there were a CA, this difference would have been large. However it was less than 6 dB. Thus conclude there is no CA effect.

This experiment assumes that the BM is reciprocal, namely that the cycle by cycle gain is independent of the direction of travel. This assumption, know as *Rayleigh reciprocity*, holds for all existing published cochlear models. For reciprocity to be violated (non-reciprocal), it must have internal magnetic fields. The cochlea is Rayleigh-reciprocity because there being no magnetic forces[13] (Postulate P6, pages 163 and 343)).

Since the CA power gain is independent of direction, the gain adds in both the forward and backward propagated waves. As a result the round trip gain is greater than the one way-gain.

The two primary frequencies are chosen as determined by the *second cochlear map* function (CM_2) $X_d(f_d)$, defined as the frequency $f_d < f_{cf}$ where the DP pressure P_d^* is maximum[15] (Fig 3). As noted in[4], f_d is also the frequency where the *slope* of the FTC jumps from ≈-60 [dB/oct], to ≈ 0 [dB/oct] (Allen[4], Figs. 3, 9-15), and where the neural phase jumps by 180° [25].

Results: These experiments found that twice the magnitude of the CA amplification gain was less than 10 [dB]. Thus we concluded that there is no significant CA cycle by cycle active gain.

In the years following[14], similar experiments were repeated in several laboratories. At least two of these experiments confirmed the observations of AF-92. Since negative results are typically not reported, these experiments were not published (as best I know).

Experiment II (FA-85)

Between 1983 and 1985[23] a quite different but related experiment was performed, which quantified the neural threshold tuning curve $P^*(f)$ as a function of a suppressor signal $P_s(f_s)$, where the suppressor frequency f_s is always below the best frequency of the neuron f_{cf}. An example is provided in the LEFT panel of Fig. 2, where the best frequency is $f_{cf} \approx 1.8$ [kHz] and the suppressor is $f_s = 0.5$ [kHz]. In these experiments the suppressor frequency is always below the best frequency of the neuron (Fig 2, LEFT). Similar data are common in the literature[23]. The two panels of Fig. 2 are modifications of figures taken from[18,23].

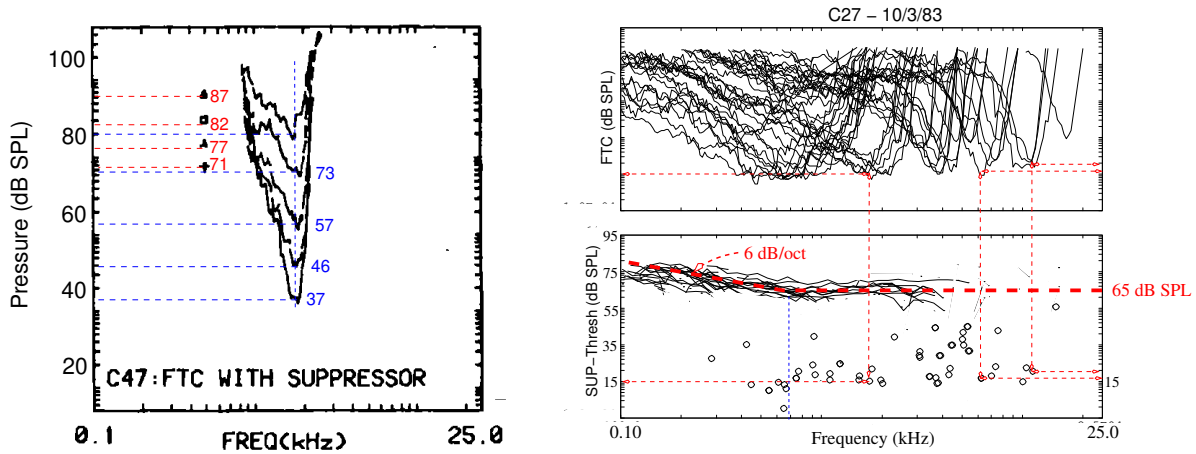


FIGURE 2. *Data from Experiment II, animals #27 and #47 from*[23].
LEFT: *Raw data from animal #47 of FA-85, Fig. 11. The neuron has a $f_{cf} = 1.8$ [kHz] and a threshold of 37 [dB-SPL]. The suppressor frequency was 0.5 [kHz] at intensities of {71, 77, 82, 87} [dB-SPL]. The suppressed threshold for these four suppressors was found to be {46, 57, 73, 80} [dB-SPL]. The suppression starts at 65 [dB-SPL], as shown in the LOWER–RIGHT panel, all of which have a threshold of 37 [dB-SPL] at f_{cf}.*
RIGHT: *This figure is a bit busy, but simple once explained. The tuning curves are from animal #27 taken from the experiment of Oct 3, 1983. On the top panel are all the tuning curves. This animal was chosen because tuning curves were observed over a wide range of frequencies, from 0.3 [kHz] to 14 [kHz]. Four examples are highlighted out of the 44 neurons tested. These four have best frequencies of {0.7, 1.9, 6.0, 10.9} [kHz]. Labeled with a red-dashed curve are the locus of suppression threshold which hover around 65 (±3) [dB-SPL]. Below 0.7 [kHz] the suppression threshold curves slowly rise, with a slope close to -6 [dB/Oct]. This effect is due to the middle ear transfer function, which acts as a high-pass filter (causing the threshold to decrease). If it weren't for the middle ear frequency response, the suppression thresholds would be independent of frequency at 65 [dB-SPL].*

The LEFT panel of Fig. 2 provides an example FTC, for the case of suppressors above 65 [dB]. Once the suppressor is greater than 65 [dB], the tuning curve is blunted around its best frequency, just below 2 [kHz]. For example, a suppressor level of 71 [dB-SPL] at 0.5 [kHz] raises the CF threshold at ≈1.8 [kHz] from 37 to 46 [dB-SPL], while a suppressor level

of 87 [dB-SPL] raises the threshold to 80 [dB-SPL], an increase of 80-37=43 [dB]. When the suppressor increases from 71 to 77 [dB] results in a threshold increase of 11=57-46 [dB], thus a slope of $11/6 = 1.833$ [dB/dB].

This NL compression effect is only seen around f_{cf}, even at the highest suppression levels, for all suppressors above 65 [dB-SPL], and more than 1 octave below f_{cf}[21].

Interpretation of the OHC nonlinear suppression: When originally published in 1985, we did not provide a physical interpretation of these suppression curves. Today, 37 years later, we have a plausible physical explanation. Figure 2 is taken from Allen and Sen [18], Fahey and Allen [23], and we repeat that story here, with an increased insight.

The RIGHT lower panel of Fig. 2 provides the big picture. When the suppressor is below 65 [dB-SPL] (bold-dashed red line in the RIGHT lower panel), there is no low-side neural suppression effect, thus the suppressed tuning curves are superimposed on the unmasked tuning curve. For suppressor levels above 65 [dB-SPL] all the neural threshold are dramatically elevated, with a slope greater than 1 [dB/dB], which means that a 1 [dB] increase in the suppressor results in more than 1 [dB] increase in its f_{cf} threshold.

Physical source of the NL-CA effect: There is substantial evidence in the literature that the source of the cochlear non-linearity is due to the acoustic stimulation of cochlear OHC cells. The obvious explanation is that above 65 [dB] the suppressor is exciting the outer hair cells (OHC), triggering the cochlear compressing non-linearity, thus reducing the quiescent sensitivity expressed by the tuning curve thresholds.

It seems highly significant that for the tails of the tuning curves having CFs greater than 2 [kHz], the IHC excitation thresholds are very similar to the the OHC supression threholds, since both are at 65 [dB-spl]. Namely the excitation and suppression thresholds are nearly the same for this large subset of neurons ($f_{cf} > 2$ [kHz]).

For those OHC that are only suppressed for tones above 65 dB, as characterized by the red-dashed line in the RIGHT-lower panel of Fig. 2, the BM would appear more linear below 65 dB since they are independent of the portion of the signal below f_{cf}. All the suppression thresholds are close to 65 [dB], as shown by red bold-dashed line in the lower right panel.

OK TO HERE

Each of the lines clustered about the dashed bold line of the lower RIGHT panel of Fig. 2 correspond to the suppression threshold for the circles below which label tuning curve thresholds. These relationships are explicitly shown as red dashed lines in the figure. The bold red dashed line below 0.7 [kHz] indicates where the suppression threshold slopes switches from 0 [dB/oct] to -6 [dB/oct]. This change in slope at 0.7 [kHz] follows the middle ear response commonly found in ear canal eardrum impedance measurements. Note that below 400 [Hz], at least half of these suppression threshold curves fall below the -6 [dB/oct] dashed bold-red line, while others lie close to it.

out of place?

When a 71 [dB-SPL] 0.5 [kHz] suppressor excites the OHCs at the 1.8 [kHz] place, the nonlinear (NL) suppression raises the FTC threshold from 37 to 46 [dB-SPL], causing the neural tuning curve to lose 9 [dB] threshold sensitive at its best frequency. When the suppressor is raised to 77 [dB-SPL], the CF threshold increases to 57 [dB-SPL], and for a 87 [dB] suppressor the CF threshold is 80 [dB] (not labeled) resulting in a compression slope of $(80-73)/(87-82) = 7/5 = 1.4$.

Based on the 44 examples in the LOWER–RIGHT panel (the number of small circles is 44), the threshold of the OHC nonlinear suppression is at 65 ± 5 [dB-SPL]. We did not see a similar suppression effect for suppressors above the best frequency.

In summary, the ramifications of these observations seem significant. Our neural measures presented here are highly consistent with different measurements from many other laboratories[6,9]. After thinking about these data for more than 37 years, I have come to a number of additional conclusions presented below, perhaps for the first time.

CONCLUSIONS

There are many:

1. The NL compression threshold at 65 [dB-SPL] has major implications for multi-band compression hearing aid signal-processing.

2. To explain the frequency independent 65 [dB-SPL] low-side suppression effect, there must be some sort of *second filter* effect that transforms the BM mechanical response to the inner hair cell (IHC) response. For example, the mathematical model of a radial resonant mode in the tectorial membrane, introduced in[17], is an example of such an process.

   Cite for example, this

   Fix footnotes!

   Related articles discussing BM to IHC transduction include[2,7,17,30]. The concept of a second-filter is quite old but was discarded after an unscientific discussion regarding an important paper[22]. This concept was later revisited by both[20] and[24], both of whom independently showed a similar elevated suppression threshold, but on the BM rather than in the neural FTC. The suppression threshold as a function of frequency (not constant and described in terms of BM velocity variables) indicate some form of second filter. However, they both avoid using the term "second filter," thereby obscuring the significance of their important observations.

   Their second-filter results confirm the elevated suppression threshold shown in Fig. 2. Their observations are quiet different because the suppression of the BM response below the CF differs in both the frequency slope and the suppressor intensity[20,24]. In summary, BM and neural thresholds are very different in their properties, but are similar in terms of the magnitude of the suppression effect. Both studies openly state the need for a "second filter."

This threshold response was previously discussed in[18] (2003). While a detailed summary of the suppression was presented, there was no attempt at a physical explanation of the source of the OHC suppression. The threshold of strong low-side suppression near 67 [dB] has been seen in many published data-sets as reviewed by[9]. This observation may date as early as the venerable loudness model of Fletcher and Munson (1933), or even Steinberg and Gardner (1924)[10].

I am hopeful, given the very strong suppression effect and the constant threshold of 65 [dB], that we can finally agree on the role of the OHC nonlinear response, given that[8,11]:

> *There is no (zero) nonlinear effect for suppressors below 65 [dB-SPL].*

It seems likely that we may continue to argue about this interpretation for years to come, but given the many observations of low-side suppression in so many labs and under many different conditions, the observed effects are not likely to change. We may (can) continue to argue over our terminology, but not the conclusions. We are all entitled to our personal opinions, but not the collective facts.

3. The observed suppression levels seen in these examples are similar, if not identical to many low-side suppression experiments in the literature, measured in various ways by[1,20,21,24,34], and possibly[28]. This view was discussed in detail by[18] (2004) as captured in the following quote:

   > A third important observation of both the Cooper and Geisler studies was that the displacement (of the OHC cilia) rather than the velocity must control the nonlinear response. This is simply proved by looking at how the iso-suppression response depends on the stimulus as a function of frequency. When BM displacement is used as the control, the suppression is independent of frequency. However when velocity is the control, the suppression threshold depends on frequency. One must conclude that the OHC stimulation is proportional to the BM displacement, not its velocity. When Ruggero [1992] uses BM velocity as the relevant variable in his characterization of suppression, he is viewing the response through a 6 dB/octave high-pass filter. This has important implications to the interpretation of his results, since, as we concluded above, that it takes a high-pass filter to bring the neural and BM measurements into alignment.

   This suppression effect is also clearly seen in the cochlear microphonic (CM) round-window voltage as shown in[3] (Fig. 16), which proves beyond question that the suppression is in the cochlea, as we have assumed in the previous paragraphs. The suppression effect also clearly shows up in the neural phase as shown in Fig. 19 of that publication.

   Add this figure to the MS

   As discussed by Allen[6], Fig. 8, p. 255 (1988), when the suppressor tone is presented at any level in the absence of an excitatory tone at the CF, the neuron has no signal output (there is only spontaneous activity). Of course, this depends on the source of spontaneous activity (SA). If the source of the SA is before the OHC cilia, then it can be suppressed. If it is after the OHC cilia it cannot be suppressed by a low-side suppressors. *Thus, the suppressor action is only seen if the excitatory tone is present.* This is a significant but perhaps underappreciated point.

4. If one were to define the cochlear amplifier in terms of the change in threshold sensitivity then it is seen to change from 37 (no masking) to 80 [dB] (87 [dB] masker). However, the effect of such a CA has nothing to do with the low intensity signals, as expected from a CA that amplifies the signal due to a negative resistance on the BM. It seems clear that we are observing what happens when the suppressor excites the OHC cilia that causes the passive gain (not the active amplification) to decrease. This follows from the fact that the cochlea shows no sign of a NL response when the suppressor is below 65 [dB-SPL].

   It is unreasonable to call this change in suppression an active amplifier power gain. It behaves more as a compressor, much like the wide band compression built into most modern hearing aids that were first introduced into the commercial market by ReSound. These commercially successful circuits were introduced to combat loudness recruitment[10].

5. An interesting parallel may be drawn between the disagreement in our field on cochlear function and the famous scientific "war" over positive feedback in the brain[29]. Scott's book is available in paperback, and online as a PDF.

   After more than 50 years (1930-1980) it was finally resolved that the brain has multi-layer positive feedback. It was a painful battle and hopefully important lesson, finally absorbed after more than 50 years.

   On pages 294-295 Scott also has a chapter on *Scientific Reductionism* that is a recognized theory of science, described as a never-ending cut-and-paste sequence of experimental overlays that converges to some fundamental answers about important questions raised mostly by scientists. For example, Reductionism is how Maxwell composed his famous equations of electro-magnetics in 1853, and how Einstein came to his theory of relativity in 1905.

   Science breaks when the research is caught up in some fundamental disagreement. In the case of a theory of the brain[29] (p. 11-12), this was a question of feed-forward vs. feed-back in brain circuits. By rejecting the now widely accepted view of neural feedback, neuroscience was retarded for at least 50 years.

   A dramatic well documented example is mentioned on page 10 of[29]:

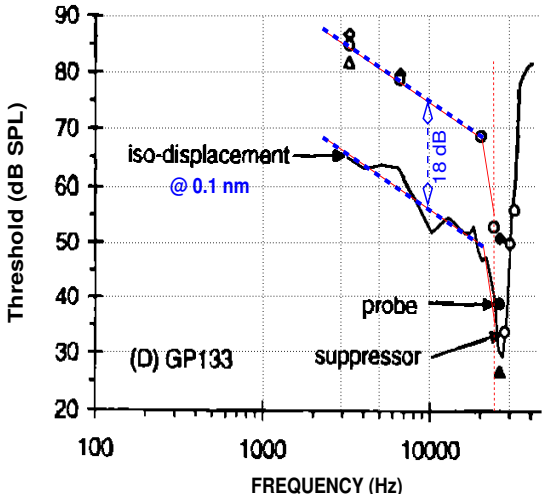
> Jerry Lettvin, a noted electro-physiologist in the Electrical and Computer Engineering Department at MIT, told me (in the spring of 1978) that when he and his colleagues reported blockages of nerve impulses in the optic nerves of cats and speculated on the possibilities of this phenomenon for visual information processing, his funding (from the National Institutes of Health) was cut off. "When you start doing science again," he was informed, "we are ready to resume your support." . . .
> An encouraging feature of science, however, is that the truth does eventually become recognized.

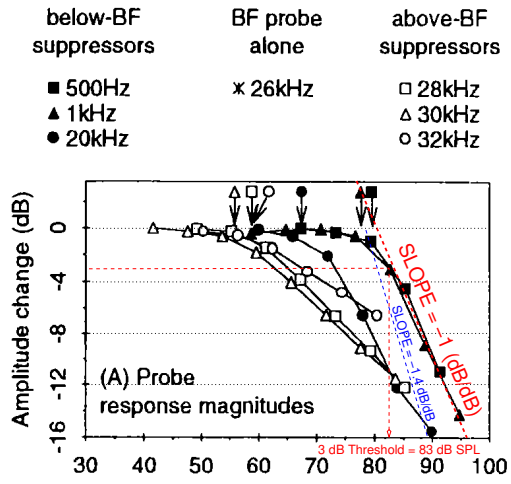
My interpretation of Lettvin's comment is that he viewed the processing in the retina prior to being put on the optic nerve. At that time this was a strongly contested point of view, not accepted then but now widely accepted.

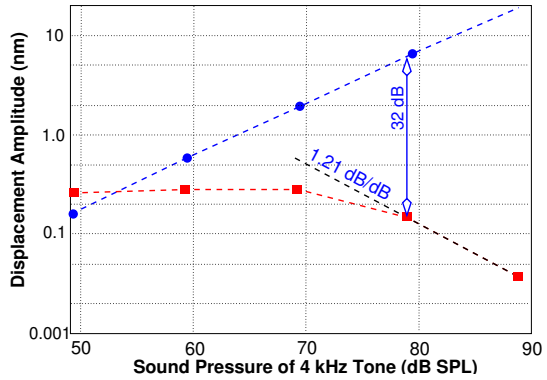
Due to the venerable principal of *Scientific Reductionism*, this last sentence does not bode well for the widely accepted cycle by cycle "cochlear-amplifier," and rejected "second-filter" concepts.

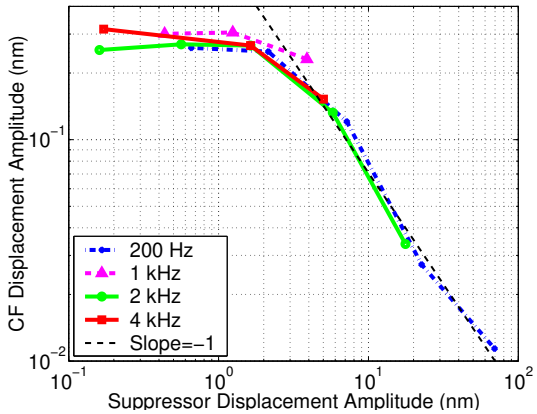
ACKNOWLEDGMENTS

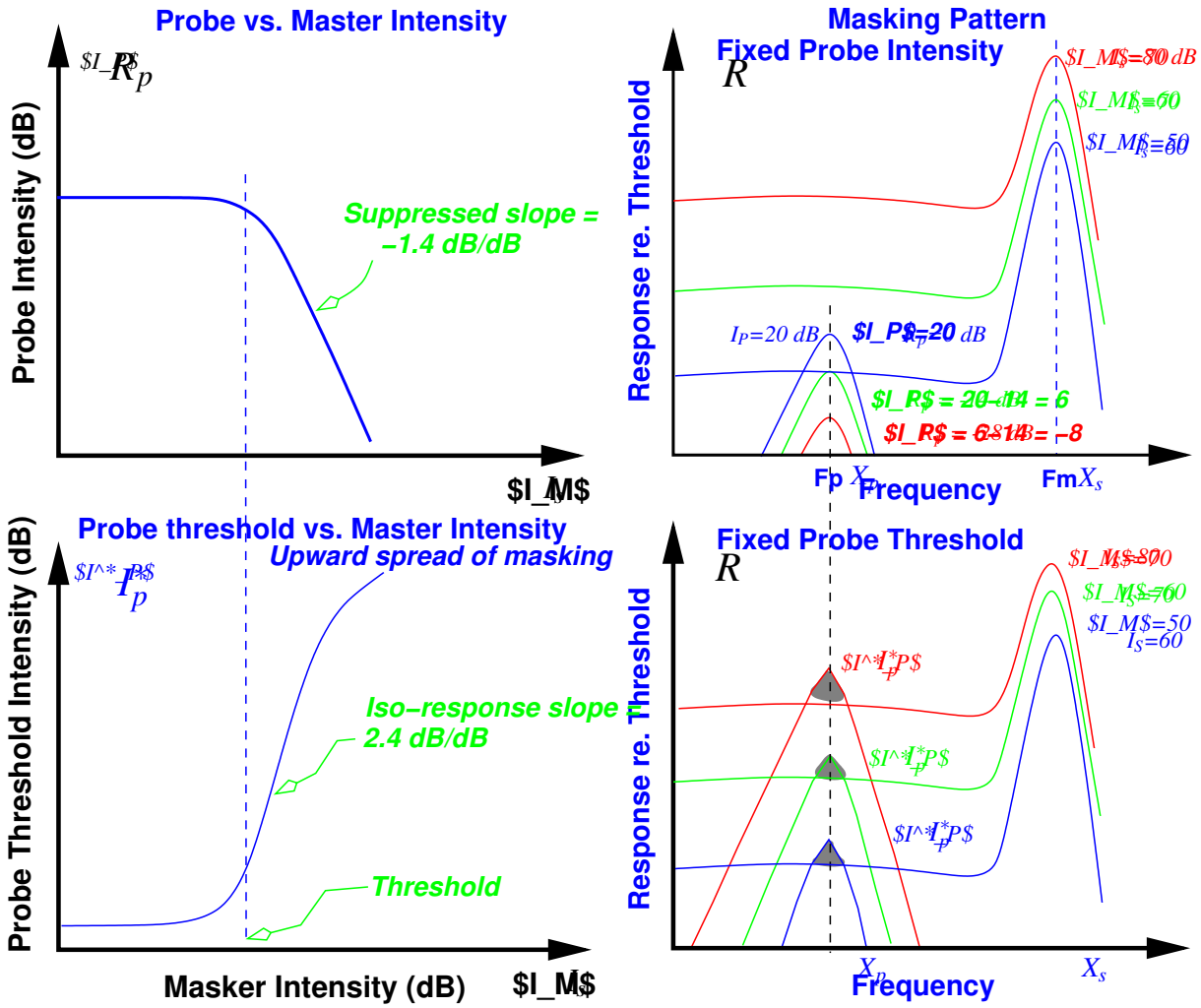
Finally I wish to acknowledge the important role played by my friend and colleague Joseph Zwislocki (1922-2018).


Threshold (dB SPL)
FREQUENCY (Hz)
iso-displacement
@ 0.1 nm
18 dB
probe
suppressor
(D) GP133


below-BF suppressors
BF probe alone
above-BF suppressors
500Hz
1kHz
20kHz
26kHz
28kHz
30kHz
32kHz
Amplitude change (dB)
SLOPE = -1 (dB/dB)
SLOPE=-1.4 dB/dB
(A) Probe response magnitudes
3 dB Threshold = 83 dB SPL


Displacement Amplitude (nm)
Sound Pressure of 4 kHz Tone (dB SPL)
32 dB
1.21 dB/dB


CF Displacement Amplitude (nm)
Suppressor Displacement Amplitude (nm)
200 Hz
1 kHz
2 kHz
4 kHz
Slope=−1


Probe vs. Master Intensity
Probe Intensity (dB)
Suppressed slope =
–1.4 dB/dB
Masking Pattern
Fixed Probe Intensity
Response re. Threshold
Frequency
Probe threshold vs. Master Intensity
Upward spread of masking
Iso–response slope
2.4 dB/dB
Threshold
Probe Threshold Intensity (dB)
Masker Intensity (dB)
Fixed Probe Threshold
Response re. Threshold
Frequency

REFERENCES

[1] Abbas, P. J. and Sachs, M. B. (1976). Two-tone suppression in auditory-nerve fibers: Extension of a stimulus-response relationship. *The Journal of the Acoustical Society of America*, 59(1):112–122.

[2] Allen, J. B. (1980). Cochlear micromechanics: A physical model of transduction. *Journal of the Acoustical Society of America*, 68(6):1660–1670.

[3] Allen, J. B. (1983a). A hair cell model of neural response. In deBoer, E. and Viergever, M. A., editors, *Mechanics of hearing*. Delft Univ. Press, Delft, The Netherlands.

[4] Allen, J. B. (1983b). Magnitude and phase-frequency response to single tones in the auditory nerve. *Journal of the Acoustical Society of America*, 73(6):2071–2092.

[5] Allen, J. B. (1986). Measurement of eardrum acoustic impedance. In Allen, J. B., Hall, J. L., Hubbard, A., Neely, S. T., and Tubis, A., editors, *Peripheral Auditory Mechanisms*, pages 44–51, New York. Springer Verlag.

[6] Allen, J. B. (1988). Cochlear signal processing. In Jahn, A. F. and Santos-Sacchi, J., editors, *Physiology of the ear*, pages 243–270. Raven Press.

[7] Allen, J. B. (1989). Is the basilar membrane tuning the same as neural tuning–where do we stand? In Wilson, J. P. and Kemp, D. T., editors, *Cochlear Mechanisms - structure, Function and Models*, volume Nato ASI Series. Plenum Press.

[8] Allen, J. B. (1997). DeRecruitment by multiband compression in hearing aids. In Jesteadt, W. and *et al.*, editors, *Modeling Sensorineural Hearing Loss*, chapter 6, pages 99–112. Lawrence Erlbaum Assoc., Mahwah, NJ.

[9] Allen, J. B. (2001). Nonlinear cochlear signal processing. In Jahn, A. and Santos-Sacchi, J., editors, *Physiology of the Ear, Second Edition*, chapter 19, pages 393–442. Singular Thomson Learning, 401 West A Street, Suite 325 San Diego, CA 92101.

[10] Allen, J. B. (2003a). Amplitude compression in hearing aids. In Kent, R., editor, *MIT Encyclopedia of Communication Disorders*, chapter Part IV, pages 413–423. MIT Press, MIT, Boston Ma.

[11] Allen, J. B. (2003b). Moderated discussion on selected topics in cochlear modeling. In Gummer, A., editor, *Biophysics of the Cochlea, from Molecules to Models*, pages 563–592. World Scientific Publishing Co., PO Box 128, Farrer Road, Singapore 912805.

[12] Allen, J. B. (2008). Nonlinear cochlear signal processing and masking in speech perception. In Benesty, J. and Sondhi, M., editors, *Springer Handbook on speech processing and speech communication*, chapter 3, pages 1–36. Springer, Heidelberg Germany.

[13] Allen, J. B. (2020). *An Invitation to Mathematical Physics and Its History*. Springer.

[14] Allen, J. B. and Fahey, P. F. (1992). Using acoustic distortion products to measure the cochlear amplifier gain on the basilar membrane. *Journal of the Acoustical Society of America*, 92(1):178–188.

[15] Allen, J. B. and Fahey, P. F. (1993). A second cochlear-frequency map that correlates distortion product, neural tuning measurements. *Journal of the Acoustical Society of America*, 94(2, Pt. 1):809–816.

[16] Allen, J. B. and Li, F. (2009). Speech perception and cochlear signal processing. *IEEE Signal Processing Magazine*, 26(4):73–77. **Invited**.

[17] Allen, J. B. and Sen, D. (1999). Is tectorial membrane filtering required to explain two tone suppression and the upward spread of masking? In Wada, H., Takasaka, T., Kieda, K., Ohyama, K., and Koike, T., editors, *Recent Developments in Auditory Mechanics*, pages 137–143. World Scientific Publishing Co., PO Box 128, Farrer Road, Singapore 912805.

[18] Allen, J. B. and Sen, D. (2003). The role of micromechanics in explaining two-tone suppression and the upward spread of masking. In *Biophysics Of The Cochlea: From Molecules to Models*, pages 383–392. World Scientific.

[19] Allen, J. B. and Sondhi, M. M. (1979). Cochlear macromechanics: Time-domain solutions. *Journal of the Acoustical Society of America*, 66(1):120–132.

[20] Cooper, N. (1996). Two-tone suppression in cochlear mechanics. *Journal of the Acoustical Society of America*, 99(5):3087–3098.

[21] Delgutte, B. (1990). Physiological mechanisms of psychophysical masking: Observations from auditory-nerve fibers. *Journal of the Acoustical Society of America*, 87:791–809.

[22] Evans, E. F. and Wilson, J. P. (1975). Cochlear tuning properties: Concurrent basilar membrane and single nerve fiber measurements. *Science*, 190.

[23] Fahey, P. F. and Allen, J. B. (1985). Nonlinear phenomena as observed in the ear canal, and at the auditory nerve. *Journal of the Acoustical Society of America*, 77(2):599–612.

[24] Geisler, C. D. and Nuttall, A. L. (1997). Two-tone suppression of basilar membrane vibrations in the base of the guinea pig cochlea using "low-side" suppressors. *Journal of the Acoustical Society of America*, 102(1):430–440.

[25] Kim, D., Siegel, J., and Molnar, C. (1979). Cochlear nonlinear phenomena in two-tone responses. In Hoke, M. and DeBoer, E., editors, *Scandinavian Audiology, Supplementum 9*, pages 63–82.

[26] Li, F., Menon, A., and Allen, J. B. (2010). A psychoacoustic method to find the perceptual cues of stop consonants in natural speech. *Journal of the Acoustical Society of America*, 127(4):2599–2610.

[27] Li, F., Trevino, A., Menon, A., and Allen, J. B. (2012). A psychoacoustic method for studying the necessary and sufficient perceptual cues of fricative consonants in noise. *Journal of the Acoustical Society of America*, 132(4):2663–2675.

[28] Ruggero, M., Robles, L., and Rich, N. (1992). Two-tone suppression in the basilar membrane of the cochlea: Mechanical basis of auditory-nerve rate suppression. *J. of Neurophysiol.*, 68(4):1087–1099.

[29] Scott, A. (2002). *Neuroscience: A Mathematical Primer*. Springer Science & Business Media.

[30] Sen, D. and Allen, J. B. (2006). Functionality of cochlear micromechanics–as elucidated by the upward spread ofmasking and two tone suppression. *Acoustics Australia*, 34(1):43–51.

[31] Sondhi, M. M. (1978). Method for computing motion in a two-dimensional cochlear model. *Journal of the Acoustical Society of America*, 63:1468–1477.

[32] Sondhi, M. M. (1981). Acoustical inverse problem for the cochlea. *Journal of the Acoustical Society of America*, 69:500–504.

[33] Sondhi, M. M. (1986). Resonances of a bent vocal tract. *Journal of the Acoustical Society of America*, 79:1113–1116.

[34] Wegel, R. and Lane, C. (1924). The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear. *Physical Review*, 23:266–285.

How exceptional is the ear?

Christopher Bergevin[a)]

Department of Physics & Astronomy,
York University
Toronto, ON Canada

[a)]*Corresponding author: cberge@yorku.ca*

Abstract. Many studies and reviews commonly refer to how “remarkable" and “exceptional" the ear is. Here we compile some of the key arguments raised in this regard. Several salient examples include: ▷ Ear is encased in hardest bone in body ▷ Ear contains most vascularized tissue in body ▷ Ear has highest resting potential in body ▷ Ear has a unique “fingerprint? ▷ Displacement of eardrum at the threshold of hearing is ∼1/100 of the diameter of a hydrogen atom ▷ Ear can detect signals below thermal noise floor and ▷ Ear is highly nonlinear (or highly linear, depending upon who you ask). For each, we provide biophysical arguments and references that allow their justification to be more critically assessed, aiming to assign a degree of validity. While of general interest to researchers and clinicians in the field, we also aim to frame these arguments so to draw attention and create enthusiasm towards the area of auditory neuroscience.

BACKGROUND

For many hearing scientists, it is usually not overly difficult to give a successful talk to a general audience. Most people *hear*, so it is natural to make a connection point to audience members. Especially older ones, who are finding hearing (typically in noisy environments) increasingly difficult. For audiences comprised of scientists and engineers, the task is even easier, as one approach is to impress upon the audience the “remarkable" [1, 2] functionality of the mammalian cochlea as a biological detector of sound. That is, the ear exhibits numerous characteristics that from a biomechanical point of view are, given limiting constraints (e.g., thermal noise, operating range of neural responses, fluid mechanics), impressive.

Consider for example that hearing intrinsically spans a very broad of spatial dimensions. The length scales (in meters) of the following hearing-related structures and phenomena span 16 orders of magnitude:

- $\diamond$ 10^{-12} m = displacement of the eardrum in response to sound at the threshold of hearing [3]
- $\diamond$ 10^{-9} = thermal-noise agitation of hair-cell bundles [4]
- $\diamond$ 10^{-8} = step size of a myosin motor
- $\diamond$ 10^{-7} = displacement of the eardrum in response to damage-inducing sounds
- $\diamond$ 10^{-5} = width of a hair cell
- $\diamond$ 10^{-2} = length of the cochlea
- $\diamond$ 10^{0} = “our" size
- $\diamond$ 10^{-2} to 10^{4} = the acoustic world we commonly perceive

On one hand, one might find it difficult to argue that this is not in fact remarkable. But on the other, there is nothing really special about hearing per se in terms of spatial scales of sensing. One could carry the argument to vision, claiming the spatial scale spanned there is 10^{26} (e.g., small: wavelength of visible light 0.5 μm, retinal cone diameter of 1 μm, large: distance between Earth and the far-away-but readily-visible star system Eta Carinae, $\approx 10^{20}$ m). An extra order of magnitude of magnitudes greater. So perhaps the ear is not so remarkable after all....

CLAIMS OF REMARKABLENESS

Nonetheless, numerous assertions regarding the ear are commonly mentioned. To set the stage, we state several here, without reference, justification, nor claims to veracity/refutation:

(a) Displacement of eardrum at the threshold of hearing is 10^{-12} m (as noted above)

(b) Ear is encased in hardest bone in body

(c) Ear contains the smallest bones in body

(d) Ear contains most vascularized tissue in body

(e) Ear has highest resting potential in body

(f) Ear is the fastest (mechanical) part of the body

(g) Ear can detect signals below thermal noise floor

(h) Ear has enormous dynamic range (in terms of incident sound energy)

(i) Input range of central nervous system (CNS) is much smaller than that of audible sound levels

(j) Ear not only detects sound, but emits it as well

(k) Ear has a built-in amplifier

(l) Ear has a unique "fingerprint"

(m) Cochlear waves do (or do not, depending upon who you ask) play an essential role

(n) Ear is (or is not) poised on the verge of instability

(o) Ear is (or is not) highly nonlinear

(p) Ear is (or is not) very fragile

(q) Human hearing exhibits relatively exceptional tuning sharpness

These claims range from provocative (e.g. a-g), to established orthodoxy in the field (e.g., h-l), to hotly debated (e.g., m-q). In many cases, these claims are not simply true or false, and the reality lies somewhere in-between. That is, direct evaluation of their validity is difficult in part due to the wide variety of animal models employed in auditory physiology. For example, how does one best compare in-vivo motions of the mouse basilar membrane to in-vitro bullfrog vestibular hair cells decoupled from the otoconial mass? Nonetheless, examination for rationale of these claims using published evidence serves to help illuminate many of the current (known and unknown) unknowns in auditory science. The goal of this manuscript is to provide a succinct and broad review of scientific literature relevant to the basis for these claims. As a preliminary version of this work, this manuscript will focus on a small subset of the claims (a-e) posed in the previous section.

A SUBSET OF JUSTIFICATIONS/REFUTATIONS

▷ *Displacement of eardrum at the threshold of hearing is* 10^{-12} *m*

True – A common methodology employed to measure auditory-related measures is laser Doppler vibrometry (LDV). This technique is capable of detecting motions whose displacements are much smaller than the wavelength of the light. Measurements indicate that at threshold, the human eardrum moves on the order of one picometer (pm; i.e., 10^{-12} m) [3]. Note that for comparison, the diameter of a hydrogen atom is 100 pm, two orders of magnitude larger. This value for threshold displacement is inferred by extrapolating down from the lowest level measured (20 dB SPL, which yielded a displacement amplitude of approximately 8 pm) to 0 dB SPL and rounding upwards. Such a value is consistent with other studies, such as [5], from which an umbo displacement of approximately 2 pm can be inferred from their Figure 3.

Physically, does this make sense? It is instructive to provide a simple back-of-the envelope calculation to argue that it does [6]. A common value that is often cited is that at our most sensitive frequencies (f, about 1-4 kHz), the

intensity of sound at threshold is approximately 10^{-12} W/m^2. Using the well-known formula relating sound intensity (I) and pressure (p) for an ideal gas

$$I = \frac{p^2}{\rho c} \tag{1}$$

and common values for ρ (density of air) and c (speed of sound in air; collectively, the product ρc is referred to as the *characteristic specific acoustic impedance*), one arrives at $p \approx 10^{-5}$ Pa, consistent with audiometric thresholds in normal hearing individuals. For sinusoidal oscillations, that air pressure depends upon several aspects of of the gas itself as follows:

$$p = \frac{2\pi BA}{\lambda} \tag{2}$$

where B is the bulk modulus of the gas, λ is the wavelength (i.e., c/f), and A is the amplitude of oscillation of the gas molecules. Rearranging and solving for A, one finds a value of approximately 10^{-11} m, roughly in the same ballpark as the LDV measurements. This means that the eardrum is moving a similar amount as the air itself in response to the acoustic pressure.

Carrying this analysis one step further, we can estimate the amount of energy incident at the ear at threshold. The human eardrum has roughly an area of 0.5 cm^2, or 5×10^{-5} m^2. Given the intensity value stated above, this would correspond to 5×10^{-17} W. Assume the ear requires 0.2 s to integrate a sound over near threshold for detection. Such would amount to one 10^{-17} J of energy, or about 2000 $k_B T$.

Several other considerations warrant further discussion. First, the eardrum is not a piston in that it does not exhibit uniform motion across its surface. In fact, sound-evoked motion of the eardrum is highly complex and not completely understood [7, 8]. One modeling study [9] suggested that the complex motion allows for the eardrum to behave in a multi-resonant fashion, thereby conveying a broadband response for transmission to the inner ear. Second, different animal species may exhibit different threshold displacements. Consider the gerbil for example, where it was reported that the umbo exhibited an approximately 1 nm amplitude for a 10 kHz stimulus at 80 dB SPL [10]. Auditory nerve thresholds at 10 kHz for gerbil at about -5 dB SPL [11]. Assuming linearity and extrapolating downwards [3, 12], that would lead to a umbo displacement of 0.1 pm, about an order of magnitude smaller than the values noted above. The reason for this discrepancy is not clear, but may stem from the spurious assumption about linearity (e.g., eardrum motion does not scale in direct proportion with incident sound amplitude at the lowest levels). As a non-mammalian example, [13] reported that for the Tokay gecko, eardrum displacements were 1 μm at 1 kHz for 100 dB re 2.0×10^{-4} dynes/cm^2 (this is equivalent to 100 dB SPL), which would translate (again, scaling linearly downward) to a 10 pm displacement at 0 dB SPL (consistent with their neural thresholds).

▷ Ear is encased in hardest bone in body

Likely True – The ear is encased in the petrous part of the temporal bone, also referred to as the otic capsule. The name stems from the Latin word *petrosus*, which translates to rocky, stone-like, or petrified [14]. Presumably this name derives from the notion that one primary purpose of the bone is for “offering protection from impacts, falls, and other trauma" [15], a helpful feature for the relatively fragile inner ear. Another viewpoint is that the denser the bone, the higher the bulk modulus that would thereby affect acoustic propagation via bone conduction.

It should be noted that with regard to density, the temporal bone is highly inhomegeneous and non-uniform. While the petrous portion, which sits on the floor of the cranial cavity, appears relatively dense in areas around the bony labyrinth, the mastoid portion has sections that are relatively sparse (e.g., holes to allow blood vessel passage) [16]. From a clinical perspective, this variation in density is significant, as it greatly improves surgical approaches (i.e., easier bone drilling) for procedures such as a mastoidectomy, an approach used commonly for implantation of cochlear implants. In some mammalian species such as gerbil, the cochlea bulges out from the base of the skull into the air-filled middle ear space (bulla). Thus, much of the murid cochlea only has a thin layer of bone around it.

The *hardness* of something, though empirically quantifiable, stems from other material properties such as Young’s modulus or density [17]. Furthermore, comparative studies of bone (e.g., [18]) note that mechanical properties such as “hardness" are not only tied to biochemical aspects such as the degree of mineralization, but functional aspects as well. Direct approaches to measuring hardness may include “work of fracture, bending strength, and modulus of elasticity in bending" [18], using the force required to produce a “crack" as a reference point. Indirect measures include “Dual

Energy Xray Absorptiometry" (DXA), ultrasound, and standard x-ray computerized tomography (CT). These methods basically boil down to how well bone absorbs or reflects incident energy. For example, CT is ultimately done by virtue of the Hounsfield scale, a transformation of the linear attenuation coefficient [19]. However, absolute calibration can be difficult. As a reference, hip bone (area) density measured via DXA is approximately 1 g/cm^2 (for 30 years of age, and decreasing from there on out) [20]. This density also varies across sex and race.

Temporal bone density has been reported in numerous clinical contexts, such as proxy measure for progression of otosclerosis [19]. To get a measure temporal bone density compares to the rest of the body, we anecdotally note that temporal bones have been used for extraction of DNA in archaeological studies (e.g., [21]). Consider this quote from [22]: “The only skeletal element for which a systematically higher endogenous DNA content compared to other skeletal elements has been shown is the petrous part of the temporal bone". These reports allude to the petrous part of the temporal bone as being the “densest" in the body, pointing towards a 1999 CT study [23] that claimed “The most dense element examined for each species is the petrous temporal, which is approximately 10 percent denser than the middle shaft portions of bones". That study however did not examine human bones, and was instead focused on “bovids, cervids, and equids".

It should be noted that there are other (non-bone) structures in the body that, at face value, common sense would dictate as likely the “hardest". Consider for example teeth (e.g., enamel density), which play an important functional role (i.e., mastication). Such are also used in forensic dentistry, for identifying persons by virtue of their teeth whose bodies have otherwise been destroyed in fire (e.g., [24]). [23] cited values as follows: “ an unsystematic sampling of enamel and dentine portions produced mineral density values of approximately 1.50 and 1.05 g/cm^3 respectively" whereas “When internal cavities are excluded from density calculations, cross-sections that consist primarily of cortical bone fall into a relatively narrow range of density (0.95-1.15 g/cm^3)."

▷ Ear contains the smallest bones in body

Likely True – Human anatomy textbooks commonly refer to the ossicles as relatively small, even the “smallest" (e.g., [16]) or “tiniest" (e.g., [25]). Research articles too also make similar claims (e.g., [26, 27]), as do books on auditory sensation (e.g., [28] states the stapes is“actually the smallest bone in the body"). Even general reference books make claims such as “The smallest bone in the human body is the stapes" [29]. But what is the rationale for such claims, as references for such are not provided? Of the three ossicles, the stapes is the smallest in terms of having the smallest maximum length. As an aside, the stapedius muscle is also commonly pointed to as the smallest muscle in the body.

There are 209 bones in the human body [30], though this number changes with ages (i.e., there are more bones at birth, but fewer as we age due to a subset that fuse together) and varies across different references. It has been observed that human ossicles are fully formed and ossified at birth, unlike other mammals [31]). In fact, the stapes has been pointed out as being the most “adult" bone in the human body (dimension-wise) even prior to birth [32]. So unlike many other bones in the body (e.g., skull) as well as ear canal geometry, the ossicles do not significantly change size after birth [31, 32]. Thus it may be possible that other bones are relatively smaller at birth.

The stapes closely resembles the stirrup of a horse saddle (*stapes* is Latin for stirrup). The footplate (which makes the final ossicular connection to the inner ear) is roughly elliptical and makes up the bulk of the stapes mass, in contrast to the overarching “superstructure" [33]. Roughly, typical dimensions are as follows [34]: footplate major and minor axes are 2.5 by 1.4 mm (yielding a footplate area of approximately 3.5 mm^2), a height of 2.9 mm (i.e., distance from footplate to top of superstructure arch), and a mass of 2.2 mg. Perhaps this value for the mass, which apparently comes from Wever and Lawrence’s 1954 book [35], might be the best quantitative basis for arguing “smallest"?

It is instructive to briefly mention that the middle ear ossicles play a significant role in bridging current theories of evolution with archeological record (e.g., [36, 37]). There is controversy over how the ossicles evolved, as well as how the relative size of the stapes has decreased. As such, there has been many comparative studies of the middle ear across mammals (e.g., [38]). Further, unlike mammals, non-mammal vertebrates only have a single ossicle. So would that still lead to the smallest bone in their bodies?

It is worth noting that older texts describing the morphology of the vertebrate skeleton sometimes indicate that there are four ossicles: the commonly mentioned three as well as the "lenticular" which was situated between the incus and stapes. This latter fourth ossicle was referred to as “smaller than the others and is sometimes not distinct" [39]. At face value, such would likely constitute the smallest bone. However, it appears that most anatomists have come to the conclusion that the lenticular is in fact the “lenticular process" of of the incus and not separate in of itself [40]. Nonetheless, some controversy has continued, with claims such as “the lenticular process is the smallest bone in the human anatomy" [40, 41].

To conclude this section, we quote the final paragraph from [32]: "*The stapes, then, is unique in its developmental history; it is the only annulet in the human body, and the most exquisitely fashioned bone in the entire supportive system. Beginning as an obscure cog in the respiratory machine, it becomes an almost indispensable part of a sensory apparatus. It is one of the most important elements in personal defense; for its services, there is no anatomical substitute. The stapes grows with dramatic speed, attaining its zenith twenty years before the humerus, or any other long bone, has achieved maturation. The stapes does not resort to the use of epiphyseal centers for growth; it need not be a giant to be important. Delicacy and mobility are required for the discharge of its physiological duties; to that end the stapes has thrown away half of its osseous self and its very marrow. It is the anatomical patriot par excellence."*

▷ *Ear contains most vascularized tissue in body*

Unlikely – Metabolic requirements for mechanical sensing systems are complex, and the vascular system is essential in this regard (see Secomb and Pries chapter in [42]). For the cochlea, the intuitive idea is that in order to generate the large resting potential that arises in the scala media due to potassium transport across the stria vascularis (see next Claim), energy is required for active transport to overcome the relatively large electro-chemical gradients [43]. Through that lens, a high degree of vascularization makes sense: Relatively high blood oxygenation would help fuel the underlying metabolic processes. However, several arguments can be made that do not support such a conclusion.

Superficially, at face value, the heart would be the most "vascularized" tissue. Or perhaps the lungs, given the highly branched nature of the pulmonary circulatory pathway. Thus, it is necessary to first define what it means to be "vascularized" via some basic quantitative details about human cardiac physiology. Focusing on the systemic circulation pathway, the general flow (i.e., heart $\rightarrow$ arteries $\rightarrow$ capillaries $\rightarrow$ veins $\rightarrow$ heart) suggests we confine our definition of "vascularized" to the capillary portion of the pathway. However, the *density* of such is a bit less straightforward from a functional viewpoint. As generally outlined by Secomb and Pries [42], there are three basic measures of vascular density: $\triangleright$ volume fraction, $\triangleright$ surface area per unit volume, $\triangleright$ length per unit volume.

For clarity, we briefly review the basic vasculature and dimensions of the cochlea. Note that the internal carotid artery, a major paired artery, passes through the the petrous portion of temporal bone via the carotid foramen [16]. So there is significant bloodflow occurring close to the cochlea. However, for actual blood supply to the inner ear, it appears there are two different principal pathways. First is the labyrinthine artery [44], which is part of the cerebellar pathway. Second are several vessels extending from the end branch of the external carotid artery (maxillary, auricular, and meningeal) supply blood to the bony labyrinth, where the cochlea resides [45]. As motivated above, the stria vascularis is a thin ribbon of tissue that runs along the lateral wall of the cochlea from base to apex, and has been noted as "a region where blood vessels are concentrated" [46] and "contains a rich supply of blood vessls" [47]. But how concentrated, how rich of a supply?

Let us now make a simple argument based upon the relative dimensions of the cochlea. An adult has approximately 5 L of blood, and at any given time: $\approx 7-10\%$ in the heart, $\approx 9-12\%$ in the lungs, $\approx 60-64\%$ in the venous, and the rest in the arteries and capillaries [16, 48]. We note that there can be considerable variation in the general dimensions of the cochlea across individuals [49], who noted "that the individual cochleae are not simply scaled versions of the same blueprint but have truly different morphologies". But to first order, the volume of the cochlea is on the order of 100 mm^3, or approximately 0.0001 L. Presumably only a small fraction (say 5%) would comprise vasculature, which is confined to the external wall and the spiral lamina). Thus at any given time, a relatively small fraction of an individual's blood (say 0.000005 %) is in a given ear, an estimate roughly consistent with [50]. From that point of view, there is relatively little blood flowing to the ear compared to other parts of the body.

However, such an analysis is likely the wrong way to consider "vascularization". Recent studies using cast corrosion and electron microscopy have provided detailed microanatomy of the cochlear vasculature [51, 52]. From these data, it becomes apparent that there is a high degree of vascular branching within the stria vascularis (see Fig.**??**A). One take away from these images is that while the cross-sectional area (in a plane slicing orthogonal through the stria) of the may be relatively small, the relative epithelial surface for transport is large. When viewed through this lens, the large degree of branching indicates that the cochlear vasculature is highly "vascularized", despite a relatively low blood volume. Now, does the cochlear vasculature bifurcate more relative to other tissues?

▷ Ear has highest resting potential in body

Likely True – First it is important to briefly define what precisely a resting potential is. A key functional aspect

of cell membranes is simply to separate *inside* from *outside*. Much of what is separated are various ionic species (e.g., sodium, potassium, calcium), whose concentrations can vary on either side, although solutions by and large obey electroneutrality [43]. As such, an electrical potential can be set up on either side due to charge separation and selective permeability. The resulting potential can be to first order most easily described as a weighted sum of Nernst potentials via the Goldman-Hodgkin-Katz equation. Changes in the permeability to the membrane can affect the instantaneous potential at any given time (e.g., generation of an action potential). But in its normal steady-state conditions, the potential is referred to as the *resting potential* [43] and effectively serves as a battery to help fuel cellular transport processes. Assuming ionic concentrations and membrane conductances are in steady-state (e.g., homeostatic conditions), this potential is relatively constant. A traditional approach to measuring the resting potential is to impale a cell with an electrode (ensuring a tight seal such that there is no leakage current) and comparing to a reference electrode nearby in the extracellular fluid. The size of this potential varies for different cell types and ionic conditions, but is typically negative (i.e., the inside of a cell appears to have a relatively negative charge compared to the outside) and takes on values ranging from 0 to -100 mV.

For the cochlea, hair cells have a relatively normal intracellular potential relative to ground ($\approx$-60 mV). But as a result of the active K^+ pumping actions of the SV, hair cells experience a much larger transmembrane potential of $\approx$130–170 mV across their apical surface relative to the scala media [53, 54]. This is called the endocochlear potential (EP) and it effectively helps drive mechano-electro transduction currents (i.e., hair bundle deflections causing N^+ and Ca^{++} currents) as well as (directly or indirectly) power active processes. But is this 130–170 mV the highest potential in the body?

There are numerous large trans-membrane potentials throughout the body in electrically excitable cells, such as a neural action potential (upwards of 140 mV; [55]), and cardiac cells. But these are transient de–/re–polarizations, unlike the EP which is relatively constant. A key facet to this relatively large transmembrane resting potential is the unique nature of the scala media environment, the endolymph composition in particular. [46] makes the point that endolymph, by virtue of having such a high potassium concentration, is unique in that it is an extracellular fluid that more closely resembles an intracellular fluid. It is also noted there that that scala media is at +80 mV relative to scala vestibuli and that the endolymph relative to the inside of the cell is “amounting to some 150 mV". Such a facet is used as a basis to motivate “the battery hypothesis", a biophysical mechanism underlying auditory transduction [46]. We also note that [47] claims the EP in mouse is +100 mV, where they define the EP as “potential difference measured in the endolymph of scala media with reference to earth". Note also that in his Fig.4 caption, [56] gives a value of 100-140 mV potential difference between SM and inside the hair cell.

SUMMARY

Here we have motivated a handful of claims regarding how “remarkable" various properties of the ear are. While inherently subjective at face value, detailed examination of such claims provides fertile ground for many key topics in auditory science. This manuscript only provides a preliminary account of a small subset of the claims, and a future version will be more comprehensive. One of particular interest is that the *Ear can detect signals below thermal noise floor*, as such integrates together many unsettled topics at the forefront of cochlear mechanics.

As a parting note, it should be noted that while beyond the scope of the present manuscript, it is illustrative to continue the conversation within a broader comparative scope [57], to examine how many of the stated claims above change when examining the multitude ear morphology in creatures inhabiting this planet. The associated evolutionary pressures have led to myriad biomechanical solutions to “hearing", a *remarkable* facet in its own right.

ACKNOWLEDGMENTS

CB was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Grant No. RGPIN-430761-2013. Instrumental to the efforts described here was Dennis Freeman, as was Allison Coffin. Comments/suggestions from Robert Harrison, Matthew Johnson, Timothy Secomb, and Sean Tulin are gratefully acknowledged.

REFERENCES

1. C. Geisler, *From sound to synapse* (Oxford University Press, New York, 1998).
2. J. Ashmore *et al.*, "The remarkable cochlear amplifier." Hear. Res. **266**, 1–17 (2010).
3. E. Dalhoff, D. Turcanu, H. Zenner, and A. Gummer, "Distortion product otoacoustic emissions measured as vibration on the eardrum of human subjects." Proc. Nat. Acad. Sci. USA **104**, 1546–1551 (2007).
4. W. Bialek, "Physical limits to sensation and perception," Ann. Rev. Biophys. Biophys. Chem. **16**, 455–478 (1987).
5. K. Whittemore, S. Merchant, B. Poon, and J. Rosowski, "A normative study of tympanic membrane motion in humans using a laser doppler vibrometer (ldv)." Hear. Res. **187**, 85–104 (2004).
6. C. Swartz, *Back-of-the-envelope Physics* (JHU Press, 2004).
7. S. Khanna and J. Tonndorf, "Tympanic membrane vibrations in cats studied by time?averaged holography." J. Acoust. Soc. Am. **51**, 1904–1920 (1972).
8. J. TONNDORF and S. Khanna, "Tympanic-Membrane Vibrations in Human Cadaver Ears Studied by Time-Averaged Holography," J. Acoust. Soc. Am. **52**, 1221–1233 (1972).
9. J. Fay, S. Puria, and C. Steele, "The discordant eardrum." Proc. Nat. Acad. Sci. USA **103**, 19743–19748 (2006).
10. O. de La Rochefoucauld and E. Olson, "A sum of simple and complex motions on the eardrum and manubrium in gerbil." Hear. Res. **263**, 9–15 (2010).
11. R. Schmeidt, "Spontaneous rates, thresholds and tuning of auditory-nerve fibers in the gerbil: Comparisons to cat data." Hear. Res. **42**, 23–36 (1989).
12. T. Buunen and M. Vlaming, "Laser–doppler velocity meter applied to tympanic membrane vibrations in cat," J. Acoust. Soc. Am. **69**, 744–750 (1981).
13. G. A. Manley, "Frequency response of the middle ear of geckos," J. Comp. Physiol. **81**, 251–258 (1972).
14. L. Aiello and C. Dean, *An introduction to human evolutionary anatomy* (Academic Press, 1990).
15. C. Etheir and C. Simmons, *Introductory biomechanics* (Cambridge University Press, Cambridge, 2007).
16. G. Tortora and B. Derrickson, *Principles of human anatomy.* (John Wiley & Sons, 2008).
17. M. Oyen, "Nanoindentation hardness of mineralized tissues." J. Biomech. **39**, 2699–2702 (2006).
18. J. Currey, "Mechanical properties of bone tissues with greatly differing functions." J. Biomech. **12**, 313–319 (1979).
19. A. Grayeli *et al.*, "Temporal bone density measurements using ct in otosclerosis." Acta Oto-Laryngologica **124**, 1136–1140 (2004).
20. A. Looker, H. Wahner, W. Dunn, M. Calvo, T. Harris, S. Heyse, C. Johnston Jr, and R. Lindsay, "Updated data on proximal femur bone mineral levels of us adults," Osteoporosis International **8**, 468–490 (1998).
21. C. Gamba, E. R. Jones, M. D. Teasdale, R. L. McLaughlin, G. Gonzalez-Fortes, V. Mattiangeli, L. Domboróczki, I. Kővári, I. Pap, A. Anders, *et al.*, "Genome flux and stasis in a five millennium transect of european prehistory," Nature communications **5**, 5257 (2014).
22. R. Pinhasi, D. Fernandes, K. Sirak, M. Novak, S. Connell, S. Alpaslan-Roodenberg, F. Gerritsen, V. Moiseyev, A. Gromov, P. Raczky, *et al.*, "Optimal ancient dna yields from the inner ear part of the human petrous bone," PloS one **10**, e0129102 (2015).
23. Y. M. Lam, X. Chen, and O. M. Pearson, "Intertaxonomic variability in patterns of bone density and the differential representation of bovid, cervid, and equid elements in the archaeological record," American Antiquity **64**, 343–362 (1999).
24. C. Savio *et al.*, "Radiographic evaluation of teeth subjected to high temperatures: Experimental study to aid identification processes." Forensics Sci. Int. **158**, 108–116 (2007).
25. E. B. Kelly, *The skeletal system* (Greenwood Press, 2004).
26. A. A. Hagr, W. R. J. Funnell, A. G. Zeitouni, and J. M. Rappaport, "High-resolution x-ray computed tomographic scanning of the human stapes footplate," J. Otolaryngol. **33**, 217–221 (2004).
27. P. Henrot, S. Iochum, T. Batch, L. Coffinet, A. Blum, and J. Roland, "Current multiplanar imaging of the stapes," Am. J. Neuroradiol. **26**, 2128–2133 (2005).
28. G. Mather, *Foundations of sensation and perception* (Psychology Press, 2016).
29. S. Jenkins, *Bones: Skeletons and how they work* (Scholastic Inc., 2016).
30. R. McNeill Alexander, *Human bones: A scientific and pictorial investigation* (Pearson Education, 2005).
31. J. Saunders, J. Kaltenbach, and E. Relkin, "Development of auditory and vestibular systems." (Academic Press, 1983) Chap. The structural and functional development of the outer and middle ear.
32. B. Anson and E. Cauldwell, "The developmental anatomy of the human stapes." Annals of Otology, Rhinology & Laryngology **51**, 891–904 (1942).
33. D. F. Àwengen, H. Kurokawa, S. Nishihara, and R. L. Goode, "Measurements of the stapes superstructure," Annals Otol. Rhinol. & Laryngol. **104**, 311–316 (1995).
34. R. Z. Gan, B. Feng, and Q. Sun, "Three-dimensional finite element modeling of human ear for sound transmission," Annals Biomed. Eng. **32**, 847–859 (2004).
35. E. G. Wever and M. Lawrence, "Physiological acoustics," in *Physiological Acoustics* (Princeton University Press, 2015).
36. N. Shubin, *Your inner fish: a journey into the 3.5-billion-year history of the human body* (Vintage, 2008).
37. G. A. Manley, "An evolutionary perspective on middle ears," Hearing research **263**, 3–8 (2010).
38. J. J. Rosowski, "The middle and external ears of terrestrial vertebrates as mechanical and acoustic transducers," in *Sensors and sensing in biology and engineering* (Springer, 2003) pp. 59–69.
39. S. H. Reynolds, *The vertebrate skeleton* (Cambridge University Press, 1913).
40. E. M. Graboyes, T. E. Hullar, and R. A. Chole, "The lenticular process of the incus," Otology & neurotology: official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology **32**, 1600 (2011).
41. E. Gaudin, "Biophysical considerations of some problems in tympanoplasties," The Journal of Laryngology & Otology **83**, 1135–1142 (1969).
42. F. G. Barth, J. A. Humphrey, and T. W. Secomb, *Sensors and sensing in biology and engineering* (Springer Science & Business Media, 2003).
43. T. Weiss, *Cellular biophysics*, Vol. 1 (MIT Press, Cambridge, MA, 1996).
44. R. T. Woodburne, *Essentials of human anatomy* (Oxford University Press, 1957).
45. R. Drake, W. Vogl, and A. Mitchell, *Gray?s anatomy for students.* (Elsevier, 2005).

46. R. Klinke, "Physiology of hearing," in *Fundamentals of Sensory Physiology* (Springer, 1986) pp. 199–223.
47. K. Steel and C. Barkway, "Another role for melanocytes: their importance for normal stria vascularis development in the mammalian inner ear." Development **107**, 453–463 (1989).
48. A. Davies, A. G. Blakeley, and C. Kidd, *Human physiology* (Churchill Livingstone, 2001).
49. E. Avci, T. Nauwelaers, T. Lenarz, V. Hamacher, and A. Kral, "Variations in microanatomy of the human cochlea." J. Comp. Neurol. **522**, 3245–3261 (2014).
50. T. Nakashima, S. Naganawa, M. Sone, M. Tominaga, H. Hayashi, H. Yamamoto, X. Liu, and A. L. Nuttall, "Disorders of cochlear blood flow," Brain research reviews **43**, 17–28 (2003).
51. M. Carraro and R. V. Harrison, "Degeneration of stria vascularis in age-related hearing loss; a corrosion cast study in a mouse model." Acta oto-laryngologica. **136**, 385–390 (2016).
52. M. Carraro, A. Almishaal, E. Hillas, M. Firpo, A. Park, and R. V. Harrison, "Cytomegalovirus (cmv) infection causes degeneration of cochlear vasculature and hearing loss in a mouse model." J. Assoc. Res. Otolaryngol. **18**, 263–273 (2017).
53. W. Sewell, "The effects of furosemide on the endocochlear potential and auditory-nerve fiber tuning curves in cats," Hear. Res. **14**, 305–314 (1984).
54. F. Offner, P. Dallos, and M. Cheatham, "Positive endocochlear potential: Mechanism of production by marginal cells of stria vascularis," Hear. Res. **29**, 117–124 (1987).
55. T. Weiss, *Cellular biophysics*, Vol. 2 (MIT Press, Cambridge, MA, 1996).
56. H. Davis, "Biophysics and physiology of the inner ear," Physiological reviews **37**, 1–49 (1957).
57. M. Vater and M. Kössl, "Comparative aspects of cochlear functional organization in mammals." Hear. Res. **273**, 89–99 (2011).

Sound-Evoked Vibrations at the Apex of the Chinchilla Cochlea

Alberto Recio-Spinoso

Instituto de Investigación en Discapacidades Neurológicas, Universidad de Castilla-La Mancha, Albacete, Spain

Corresponding author: reci0001@umn.edu

Abstract. Despite tremendous advances in the field of cochlear mechanics, the apex remains one of the least understood of the cochlear regions, particularly in laboratory animals whose low-frequency hearing shares similarities with humans. Direct measurements of sound-evoked vibrations of the organ of Corti at the apex were until recently fraught with technical difficulties and experimental artifacts due in part to the necessity of opening the otic capsule to allow such measurements. With the advent of optical coherence tomography (OCT) machines, some of those difficulties have been overcome. Recently published results showed that at the apex of the guinea pig cochlea, nonlinear vibration amplitudes as a function of frequency display low-pass characteristics, which differs from the band-pass behavior commonly seen in basal responses. Here we present an overview of apical recordings performed using an OCT system in the cochlea of the chinchilla, an animal with a frequency hearing range similar to humans'. Just as with apical recordings in the guinea pig, the vibration pattern in responses to single tones in the chinchilla apex exhibits low-pass characteristics at a site having a cutoff frequency of around 500 Hz.

INTRODUCTION

The development and application of new technologies, such as optical coherence tomography (OCT), to the field of cochlear mechanics has improved our understanding of the functioning of the cochlea [e.g., 1-6]. In many situations, measurement of vibrations in the organ of Corti (OoC) can be performed in intact cochleae using OCT and thus diminishing the effects of experimental artifacts. Whereas basilar membrane (BM) responses at the base of the cochlea exhibit a band-pass shape, mechanical responses at the apex of the guinea pig and gerbil show a low-pass profile.

Here we try to extend our knowledge of apical cochlear mechanics by recording sound-evoked responses in chinchillas, an animal with low-frequency hearing similar to humans' [7]. OoC recordings were performed in intact cochleae using a Spectral Domain OCT and vibrometry at a region with a characteristic frequency (CF) of around 500 Hz. Non-linear responses to acoustic stimuli were measured at various locations along the optical axis of the beam inside the OoC. Displacement amplitudes as a function of stimulus frequency reveal a low-pass characteristic with a cutoff frequency of 400-600 Hz, i.e., near CF.

METHODS

Adult chinchillas (N=12) of either sex weighing around 500 grams were used. Each chinchilla was anesthetized with ketamine (40 mg/Kg, i.m.) and xylazine (2 mg/Kg, i.m.) with additional smaller doses applied as necessary. The head of each chinchilla was fixed in a head holder using dental cement, a tracheotomy was performed, and the animal's body temperature was maintained at 37.5°C using a heating pad and a rectal probe with a servo-controlled heating system. The external right ear was removed and both dorsal and ventral sides of the bulla were opened to visualize the middle-ear bones and the apical end of the cochlea, respectively. The tensor tympani muscle was transected, and the stapedius muscle was detached from its anchoring.

Sound-evoked vibrations of the intact organ of Corti were recorded using a Telesto Spectral Domain OCT system (Thorlabs GmbH, Germany) and ThorImage® OCT version 5.4.2 (Thorlabs GmbH, Germany). Further analysis of the output of the ThorImage software was performed by ad-hoc programs written using Matlab R2020b (Mathworks®, Natick, Massachusetts, USA). During the initial experiments, sound stimuli were generated using a 14-bit Analog Discovery sound card (Digilent, Pullman, Washington, USA). For later experiments, sounds were digitally generated using a 24-bit RME Fireface UC sound card (IMM Elektronik GmbH, Mittweida, Germany). Software to control the sound cards was written using Matlab R2020b. For this work, stimuli consisted of sinusoidal tones of 1 sec of duration presented 2 or 3 times. Synchronization between the sound card and the Telesto system was achieved by enabling the external triggering option in ThorImage and a clock signal generated by the sound card. The frequency of the clock signal was modified using a custom-made frequency divider. In general, the output rate of the sound card was 32 K samples/sec and the sampling rate of the Telesto system was set by the clock signal frequency at 8 K samples/sec.

RESULTS

Vibration responses in the OoC were analyzed using the Fourier transform. The following section shows the results of this analysis, expressed by amplitudes and phases at several stimulus levels and frequencies. The CF of the cochlear region under study in this report was approximately 500 Hz, which is similar to the CF reported by Rhode and Cooper [8] in their recordings at the apex of the chinchilla. Also, the location of the cochleostomy in a previous study of auditory nerve fiber (ANF) responses to sound [9] had a CF similar to the current study.

Anatomic Imaging and Vibration Responses

At the beginning of each experimental session, the OCT system was used to create a cross-sectional image (or B-Scan) of the apex of the cochlea, which permitted visualization of different structures of an apical turn (Fig. 1A). It was possible to identify, in general, structures such as the otic capsule bone, the OoC, scala vestibuli (SV), scala tympani (ST), and Reissner's membrane (RM). The region of interest (ROI) was located inside the OoC (red dashed line in Fig. 1A). In most experiments the length of the ROI was around 250-300 μm.

Figure 1B displays sensitivity functions, which were defined as the ratio of the vibration amplitude to stimulus intensity, obtained at several depths as indicated by the horizontal axis of the figure. Because of the non-linear behavior of the vibrations, the two traces in Fig. 1B do not overlap each other. In general, the 60-dB SPL signal yielded higher sensitivities than the 70-dB SPL tone. Furthermore, sensitivity values are not constant as a function of depth; their values decrease toward the BM/ST region.

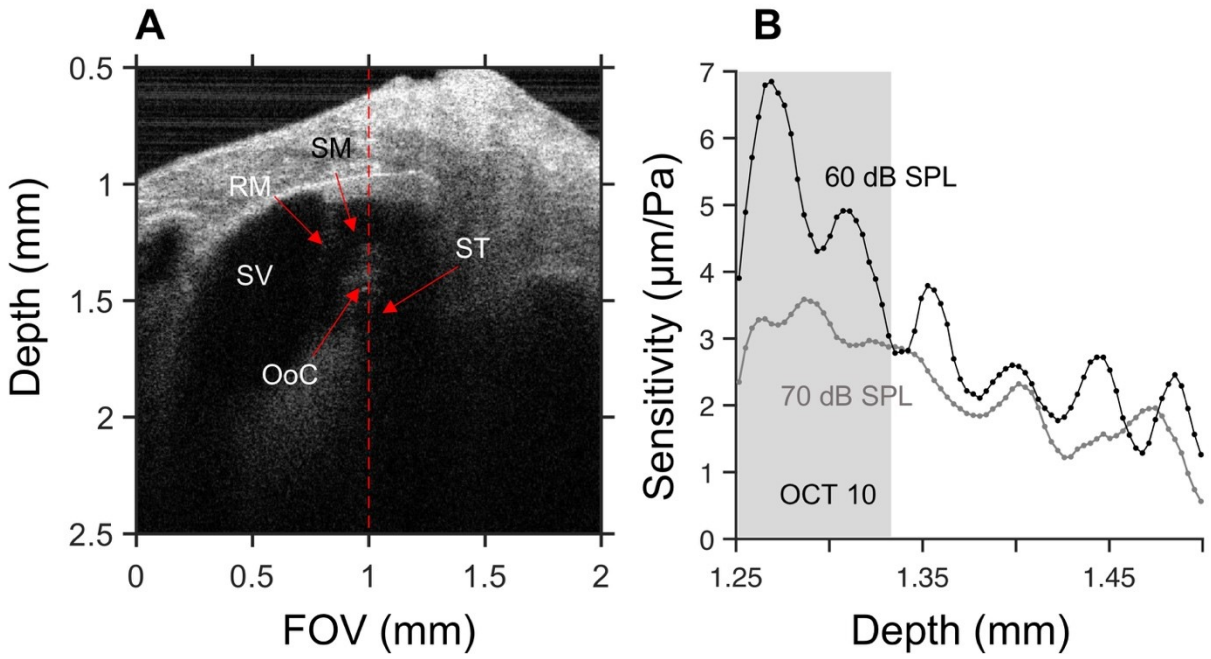


FIGURE 1. Display of a cross-section of an apical cochlear region in which sound-evoked vibrations were recorded at sites indicated by the red dashed line (panel A). The region of interest is at a depth range of 1.25-1.5 mm, which corresponds to locations in the organ of Corti (OoC). Reisssner's membrane (RM), scala vestibuli (SV), scala tympani (ST), and scala media (SM) are also indicated in A. Panel B shows sensitivities at 60- and 70-dB SPL as a function of depth. The stimulus frequency is 400 Hz.

Figure 2 shows response areas, which consist of a family of sensitivity-vs.-frequency curves at several stimulus levels. Sensitivity values in Fig. 2 correspond to maximum values inside one-third of the ROI (shaded area in Fig. 1B, between 1.25 and 1.33 mm). Values in Fig. 2A are relatively constant up to around 500 Hz (dashed lines) and then decrease — in accordance with the low-pass behavior of apical vibrations in the cochlea of the guinea pig [6]. Sensitivity functions in Fig. 2B were obtained using the velocity of the vibrations, instead of their displacement as in Fig. 2A. Results in Fig. 2B do show a band-pass profile with peak values at 500 Hz (dashed lines). Notice also that individual sensitivity-vs.-frequency functions across stimulus levels do not overlap each other, a consequence of OoC non-linear properties. Maximum cochlear gain values, defined as the ratio of peak to minimum sensitivities, at 500 Hz and 100 Hz were 50 and 32 dB, respectively.

A visual comparison of the results in Fig. 2A with those presented in a previous study of apical cochlear mechanics in the chinchilla [8] reveals significant differences. Whereas shapes of response areas in this study exhibit a low-pass characteristic, results in [8] reveal band-pass tuning.

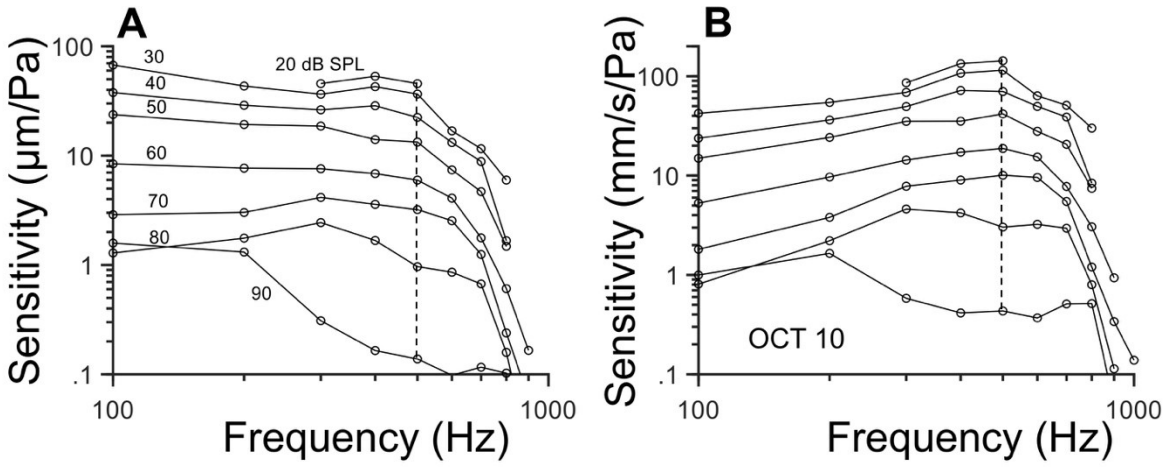


FIGURE 2. Sensitivity functions for sound-evoked displacements (panel A) and velocities (panel B) as a function of frequency. Intensity levels (20-90 dB SPL) are shown above the corresponding traces. Data correspond to peak values recorded at 1.25-1.33 mm depth (see Fig. 1). Dashed lines in Fig. 2 point to sensitivity values at 500 Hz.

Phase Responses

Sound-evoked vibrations across sites inside the ROI in Fig. 1A do not occur at the same time. This observation is shown in phase-vs.-depth functions obtained from the responses to a 400 Hz tone (Fig. 3A). Phases in Fig. 3A are relatively constant, regardless of the stimulus intensity, up to around 1.35 mm. Beyond that point, there is an overall increase in phase lead as a function of depth. The amount of phase lead in Fig. 3A appears approximately related to the stimulus level.

Phases obtained from the responses in Fig. 2 were also plotted relative to phases in responses to 90-dB SPL tones (Fig. 3B). Due to the non-linear properties of OoC vibrations, phase functions do not overlap each other. Below 600 Hz, increases stimulus level resulted in phase leads. For frequencies above 600 Hz, the opposite occurred: increases in stimulus level yielded phase lags.

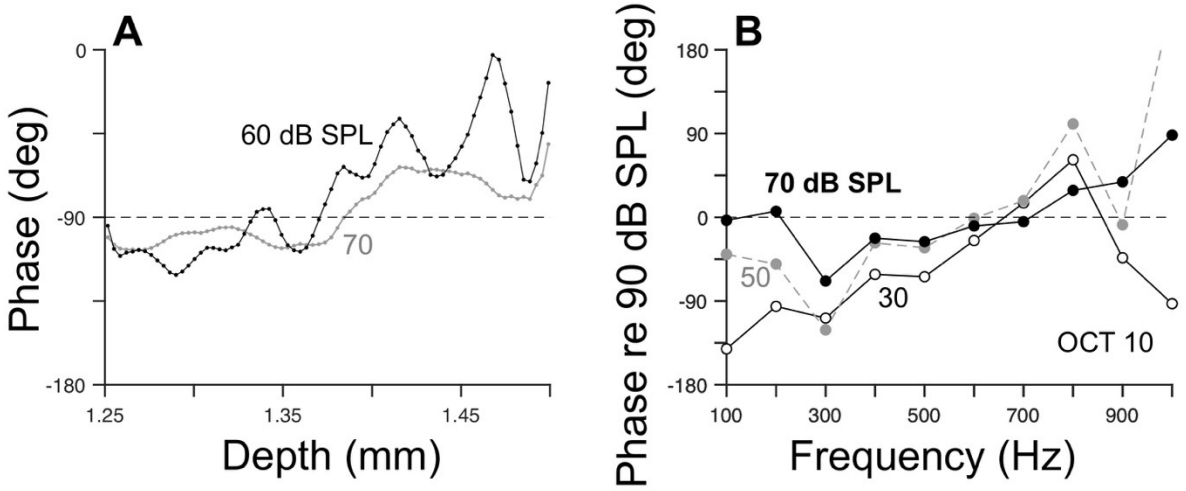


FIGURE 3. Panel A shows phase-vs.-depth corresponding to the amplitude data in Fig. 1B. Panel B displays phase values corresponding to the data in Fig. 2. Phases as a function of stimulus frequency are expressed relative to the response phases to 90-

dB SPL tones. Positive and negative phases indicate leads and lags, respectively. For clarity, phases for 20, 40, 60 and 80 dB SPL tones are not shown.

DISCUSSION

Data presented in this work disagree with certain aspects of previous measurements of vibrations in the apex of the chinchilla [8] that were performed at a site with similar CF. Whereas vibratory responses in our experiments do not significantly decrease at frequencies below CF, previous measurements do show a band-pass shape [8]. This discrepancy can be explained by the fact that the latter measurements were obtained following a cochleostomy, a procedure that has been shown to introduce artifacts in cochlear and neural responses to sound [6,9,10].

Level-dependent phase changes in Fig. 3B differ from those obtained from ANF responses [11]. Increases in stimulus levels in nerve responses yield phase lags for stimulus frequencies below CF. The reverse — phase leads with level increases — occurs for frequencies above CF. Interestingly, phase changes in Fig. 3B are similar to those found in a study of the effects of opening the otic capsule on ANF responses to sound [9].

It is important to take into consideration that results presented in this work originate from OoC regions that evoke the largest displacements (shaded area in Fig. 1B). It is thus conceivable that OoC regions that are closer to the BM yield amplitude- and phase-vs.-frequency curves that might be similar to those obtained from ANFs.

ACKNOWLEDGMENTS

Thanks to Dr. Martin Thunert of Thorlabs for his help in setting up the OCT system in Albacete, Spain. Work was supported by grant EQC2018-004820-P from the Ministerio de Ciencias, Innovación y Universidades (Spain) and the European Social Fund.

REFERENCES

1. S. Ramamoorthy, Y. Zhang, T. Petrie, A. Fridberger, R. Wang, S.L. Jacques and A.L. Nuttall, J. Biomed. Opt. **21**, 25003 (2016).
2. R.L. Warren, S. Ramamoorthy, N. Ciganović, Y. Zhang, T.M. Nilson, T. Petrie, R.K. Wang, S.L. Jacques, T. Reichenbach, A.L. Nuttall and A. Fridberger, Proc. Natl. Acad. Sci. USA **113**, E4304-E4310 (2016).
3. H.Y. Lee, P.D. Raphael, J. Park, A.K. Ellerbee, B.F. Applegate and J.S. Oghalai, Proc. Natl. Acad. Sci. USA **113**, 3128-3133 (2015).
4. T. Ren, W. He and D. Kemp, Proc. Natl. Acad. Sci. USA **113**, 9910-9915 (2016).
5. N.P. Cooper, A. Vavakou and M. van der Heijden, Nat. Commun. **9**, 3054 (2018).
6. A. Recio-Spinoso and J.S. Oghalai, J. Physiol. **595**, 4549-4561 (2017).
7. M. Trevino, E. Lobarinas, A.C. Maulden and M.G. Heinz, J. Acoust. Soc. Am. **146**, 3710-3732.
8. W.S. Rhode and N.P. Cooper, Audit. Neurosci. 3, 101-121 (1996).
9. A. Recio-Spinoso, A.N. Temchin and M.A. Ruggero, "Effects on Auditory-Nerve Fibers of Opening the Otic Capsule at the Apex of the Chinchilla Cochlea" in *Mechanics of Hearing: Protein to Perception*, AIP Conference Proceedings 1703, edited by K.D. Karavitaki and D.P. Corey (AIP Publishing, Melville NY, 2015), pp. 040007-1.
10. W. Dong and N.P. Cooper, J. R. Soc. Interface **3**, 561-571 (2006).
11. A.N. Temchin and M.A. Ruggero, J. Assoc. Res. Otolaryngol. **11**, 297-318.

OCT-based Method for the Combined Measurement of Structural Vibration and Fluid Pressure

Rahul Ramdas[1] , Naman Agarwal[1], Mangesh Atpadikar[1], and Sripriya Ramamoorthy[1, a)]

[1]*Department of Mechanical Engineering, Indian Institute of Technology Bombay, Mumbai, Maharashtra, India*

[a)]*Corresponding author: ramamoor@iitb.ac.in*

Abstract. Simultaneous measurement of both vibration and pressure inside the cochlea *in vivo* would enable combined tracking of slow as well as fast-wave propagation. Such a measurement method could provide new information on the backward propagation of otoacoustic emissions. Thus far, the *in vivo* measurements of the organ of Corti vibrations using Optical Coherence Tomography (OCT) and measurement of pressure in the scalae using optical microphone have been reported only separately and involving different sensors. In this article, a method for integrated pressure and vibration measurements using OCT is described. The microphone consists of a micro-scale cavity, with a thin diaphragm made from an optical adhesive attached to the cavity similar to Olson's [1] optical microphone. However, the key difference in this new method is that there is no optical fiber for the electrical recording of the pressure. Instead, the vibrations of the microphone diaphragm are detected along with the vibrations of any multi-layer structure, such as the organ of Corti, simultaneously using OCT. A calibration chart of velocity per pressure for the microphone is used to convert the OCT-measured microphone diaphragm velocity into pressure recordings. The demonstration of this method in an air-filled vibroacoustic microchannel and application to eardrum impedance measurement is reported. The eardrum impedance measured in a GRAS ear simulator agrees well with the measured human ear drum impedance from literature.

INTRODUCTION

Certain vibroacoustic applications benefit from or require measuring the vibration of a structure and the acoustic pressure in an associated fluid in the vicinity. While it is reasonable to use multiple devices to detect vibration and pressure separately in larger applications, microscale applications generally cannot afford the space requirements, or the uncertainties brought about by multiple sensors. The performance of these applications could be limited by discrepancies caused in the measurements using multiple sensors due to calibration errors or differences in sensitivities of the sensors.

These challenges are further extended in microscale channels like the hearing organ cochlea, where suitable commercial sensors may not be available [2]. Sound entering the cochlea causes the propagation of a fluid borne compression wave and a structurally coupled travelling wave [3]. The coupled travelling wave excites the microstructures within the cochlea and travels much slower than the compression wave. While the nature of the forward propagation of the waves inside the cochlea is reasonably understood [4], it is unclear whether the compression wave or the travelling wave is responsible for the reverse propagation of signals generated inside the cochlea. Understanding the reverse propagation of waves is essential to understanding otoacoustic emissions, which are sounds generated from the cochlea and important for auditory diagnoses. Pressure transducers like those developed by Olson [1] have been vital in providing information about fluid pressure in live rodent cochlea. There have also

been prior measurements of in-vivo vibration [5,6]. However, a minimally invasive combined measurement method to detect both vibration and pressure is not yet available.

Another area of interest for the combined measurement of fluid pressure and associated structural vibrations is the measurement of acoustic impedance at the eardrum looking into the ear in humans. This is necessary for understanding issues like conductive hearing loss. Tympanometry has been the technique that has been used extensively to measure the pressure reflected at the eardrum. However, the complex geometry of the ear canal, the shape of the eardrum and the lack of real time guidance makes measuring the impedance at the eardrum both difficult and potentially error prone.

This work presents an optical method for the simultaneous measurement of both fluid pressure and structural vibrations in vibroacoustic microchannels, as well as macroscale applications. The method expands on the vibration measurement capacity of phase-sensitive spectral domain optical coherence tomography (SDOCT), with the addition of a miniature sensor to perform fluid pressure measurements in vibroacoustic systems. The method is validated using a square shaped cavity with a flexible wall [7]. Further, a demonstration of the method to measure eardrum impedance is done by using an ear simulator as a test subject. While implementing a stand-alone device for impedance measurement has challenges, the proposed method can be integrated with OCT based middle ear otoscopes, like Park et al. [8].

METHODS

A commercial SDOCT imaging system (TELESTO III SDOCT from ThorLabs, Inc.) is custom adapted to perform complex vibration measurements. OCT interferometry uses a low coherence, broadband source of light in a Michelson interferometer setup to provide depth resolved images of samples. The central wavelength of the TELESTO III system is 1310 nm, with a bandwidth of 150 nm. The system has an axial resolution of 3.48 microns and lateral resolution of 13 microns in air. While there are multiple preset sampling rates provided by the manufacturer, the preset sampling rate used in this study is 76 kHz.

The OCT system and settings are controlled by a workstation which uses ThorImage software, a GUI that can be used to focus samples and record data. Using the external triggering feature available in the software for the system, time-synchronous vibrations measurements can be made. Vibrations are induced in a sample using a function generator. The OCT and the function generator are triggered simultaneously using a chain of TTL pulses (90 % duty cycle) generated by an NI PCIe-6351 multifunction I/O card, at a rate of approximately 76 kHz (exactly 75,643 Hz). Upon triggering, the function generator produces a sinusoidal voltage signal as input to the source which causes the source to vibrate, while the OCT system simultaneously starts image acquisition. The OCT light is focused on a single point on the sample by setting the field of view to zero. This disables the galvanometer actuated mirrors that control the probing beam and enables the M-mode scanning. The point is scanned 10,000 times (@76 kHz) in each M-scan and 10 such M-scans are obtained for time averaging. The raw data obtained from image acquisition is stored as intensity vs optical wavelength. The method for post processing the raw data to obtain the depth-resolved image is detailed in Lin et.al [9]. The image obtained is complex intensity as a function of depth and time. The phase of this complex intensity contains the vibration information. Extraction of this vibration information is performed in MATLAB using the method described in Wang et al. [10].

The custom vibration measurement capability of the SDOCT system is validated using a loudspeaker diaphragm as the sample. The central velocity of the diaphragm of a 10 W PUI-Audio loudspeaker (model AS20908AR-R) is measured using OCT vibrometry as well as using Laser Doppler Vibrometry (Polytec Micro System Analyzer 500 ppm available as central facility at IIT Bombay). The velocity response is normalized to the input voltage over a frequency range of 20 Hz to 20 kHz. The frequency response comparison obtained from both methods is shown in Figure 1 of our publication Ramdas et al. [7], thus providing a benchmark for OCT based vibrometry implemented in TELESTO-III system.

RESULTS

Pressure Measurement

A miniature sensor which uses OCT based vibrometry described in the previous section is developed to measure the local pressure at a point inside small vibroacoustic cavities. The sensor is a cavity-diaphragm assembly, as shown in the Fig. 1. The cavity, filled with air, is of cylindrical geometry with a diameter of 1 mm and depth of 1 mm. It is

fabricated by 3D printing using RC-90 resin (SLA printing technique). The diaphragm is made using a liquid photopolymer, NOA-68, which can be cured under ultraviolet (UV) light to obtain a transparent, thin, and flexible diaphragm. A small drop of about 10 µL of the liquid photopolymer is placed on distilled water in a petri dish. The drop expands upon contact with the distilled water and spreads to form a thin film. The film is cured using a UV lamp for 5 minutes. The cured film is scooped using the 3D printed part containing the cavity such that the top of the cavity is now covered by the cured film. The assembly is then cured for a further 15 minutes under UV light to ensure that the film adheres to the cavity well. NOA-68 has a Young's modulus of E = 138 MPa and density ρ = 1250 kg/m^3.

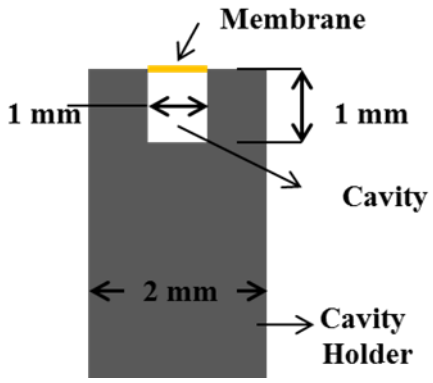


Figure 1. Model of the sensor used for pressure measurement

Calibration

The velocity of the sensor diaphragm to an applied external pressure is used to calibrate the sensor. To obtain this transfer function, the sensor is placed very close (center to center distance < 1 mm) to a pre-calibrated microphone (G.R.A.S 40 PP). A sinusoidal pressure excitation is provided from a 10 W PUI-Audio loudspeaker, kept ~7 cm away from the center of the sensor diaphragm. The vibrations of the sensor diaphragm due to this sinusoidal external pressure acting on it is measured using OCT based vibrometry. The pressure obtained by the pre-calibrated microphone is taken as the pressure acting at the center of the sensor diaphragm. The ratio of the velocity obtained from the OCT measurement to the pressure obtained from the pre-calibrated microphone is plotted across frequency to obtain the calibration chart of the sensor. The calibration is also predicted using a lumped model of the cavity-diaphragm assembly. Considering the diaphragm to be a rectangular simply supported, thin plate of dimensions a and b and thickness h, the transfer function is given by

$$\frac{v_0}{p_{ext}} = -\frac{j\omega\frac{16S}{\pi^2}}{[DS((\frac{\pi}{a})^2+(\frac{\pi}{b})^2)-\rho hS\omega^2]+\frac{64}{\pi^4}\frac{\rho_{air}c^2S^2}{V_{cav}}} \tag{1}$$

Where v_0 is the velocity of the diaphragm center and p_{ext} is the spatially constant external pressure acting at the diaphragm center. S is the area of the diaphragm, ρ is the density of the NOA 68 material, ρ_{air} is the density of air, c is the speed of sound in air, V_{cav} is the volume of the cavity and D is the bending rigidity of the diaphragm (having Poisson's ratio ν) given by

$$D = \frac{Eh^3}{12(1-\nu^2)} \tag{2}$$

The calibration chart obtained from experiment is compared with that from the lumped model as shown in Figure 2D of Ramdas et al. [7]. The lumped model is expected to agree with the calibration up to just below the first resonance frequency, which can be increased by reducing the size of the sensor cavity. Apart from this, the fluid mass loading is quite small when the experiment is performed in air owing to the small size of the sensor and diaphragm and is therefore neglected. In water or other fluids, this mass may become significant.

The use of OCT vibrometry along with the designed sensor for simultaneous measurement of vibration and pressure in a vibroacoustic cavity is the main goal of this study. To that end, the miniature sensor is inserted into a 3D printed cubical cavity of dimensions 5 mm x 5 mm x 1 mm such that the sensor diaphragm is flush with the bottom interior wall of the square cavity. The cavity is acoustically rigid apart from two openings: a small hole required to insert the sensor and the top face of the cavity which is covered by a flexible diaphragm cut out of plastic packaging material (HDPE). The cubical-cavity and sensor are exposed to sinusoidal pressure excitation from a 10 W PUI-Audio loudspeaker kept ~7cm away from the assembly. A pre-calibrated G.R.A.S 40 PP microphone is placed very close to the center of the square cavity diaphragm to obtain the external pressure acting on it. The setup is placed under the OCT sample arm and the vibrations of the square cavity diaphragm as well as the sensor diaphragm are measured simultaneously. An important consideration here is that the distance between the two diaphragms be less than the maximum imaging depth of the OCT system (3.5 mm), which is ensured here as the two diaphragms are only 1 mm apart. The velocity of the cubical cavity diaphragm is normalized to the external pressure acting on the diaphragm (measured by the G.R.A.S microphone) to obtain the velocity response of the diaphragm.

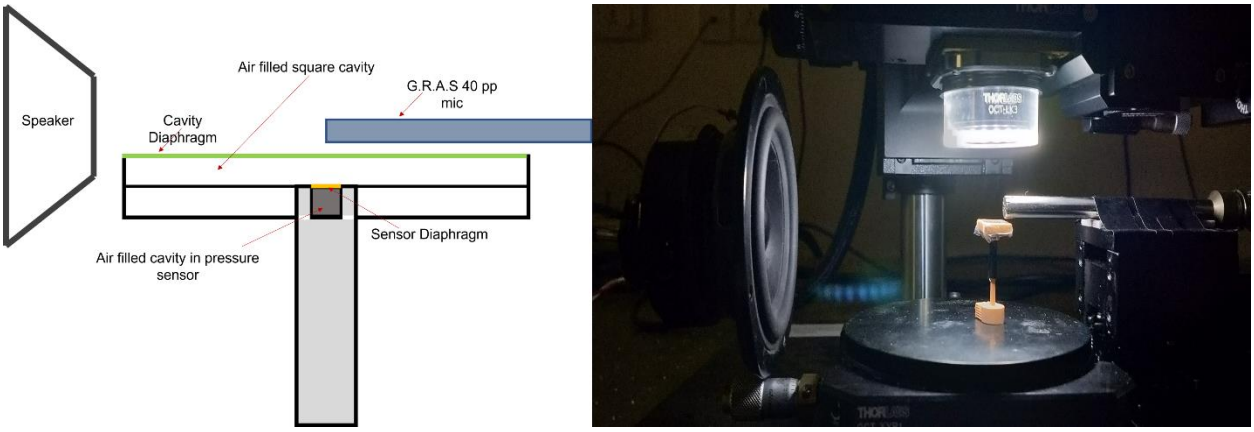


Figure 2. Experimental setup for pressure measurement

The velocity response of the sensor diaphragm is converted into the pressure inside the cavity using the calibration chart previously obtained. The velocity response of the cubical cavity diaphragm is obtained using the same lumped model consideration mentioned previously during sensor calibration. The internal pressure is modelled analytically using a lumped model where the air inside the cubical cavity is treated as a stiffness. The resulting pressure due to the diaphragm motion is given by

$$p_{int} = -j\frac{\omega\rho_{air}c^2}{V_{sqcav}}v_{0sq}\frac{4S}{\pi^2} \tag{3}$$

Where V_{sqcav} is the volume of the square cavity and v_{0sq} is the velocity of the centre of the square cavity diaphragm. The velocity response of the diaphragm and internal pressure inside the cavity are shown in Figures 4D and 4E respectively of Ramdas et al. [7].

Eardrum Impedance

The use of the proposed method to measure impedance at the eardrum is demonstrated in this section. The G.R.A.S 43AG7 ear simulator, which mimics the impedance at the human eardrum looking into the ear, is used as the test subject. Since there is no mechanical vibrating component in the ear simulator, a thin, flexible, polymer diaphragm is attached to the ear canal portion of the simulator. This would not be necessary in actual human ears as the eardrum itself acts as a vibrating component. The cavity of the sensor is once again fabricated using 3D printing, but with a translucent resin this time, to allow for the OCT light to pass through the back wall of the sensor cavity. The sensor is calibrated as explained in the calibration sub-section in the previous section. The calibrated sensor with the translucent

body is kept very close (~0.5 mm apart) to the polymer diaphragm attached to the 43AG7 ear simulator. The 10 W speaker is used to provide sinusoidal input sound to the assembly as shown in Figure 3.

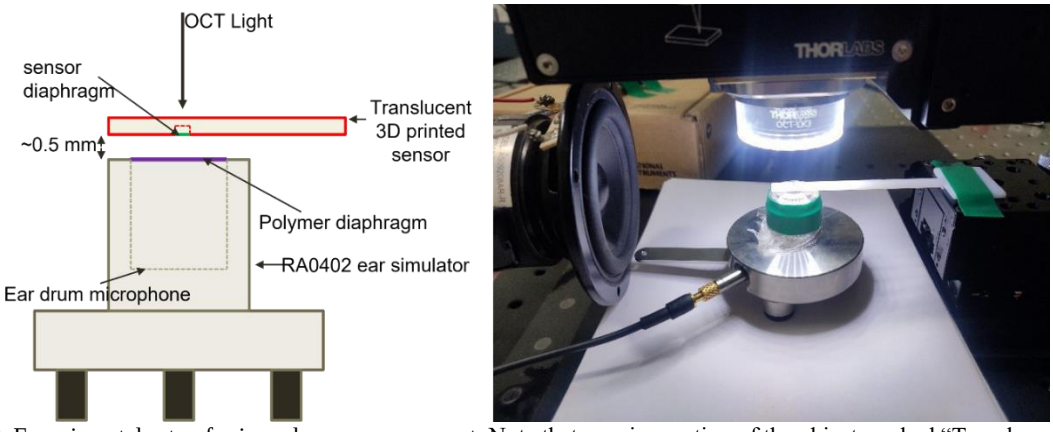


Figure 3. Experimental setup for impedance measurement. Note that a major portion of the object marked "Translucent 3D printed sensor" is for handling purposes. The actual sensor cavity is only about 2 mm diameter and 1 mm depth.

The vibration of the attached polymer diaphragm is measured using OCT vibrometry, while the vibrations of the sensor diaphragm is converted to the pressure acting at the polymer diaphragm using the calibration curve of the sensor. The acoustic impedance at the diaphragm is thus obtained by the ratio of the pressure at the diaphragm to the volume velocity of the diaphragm. Using the impedance translation theorem [11], this impedance is then transferred to the mechanical impedance at the location mimicking the 'eardrum' in the 43AG7 ear simulator, which is at the internal RA0402 microphone in that simulator.

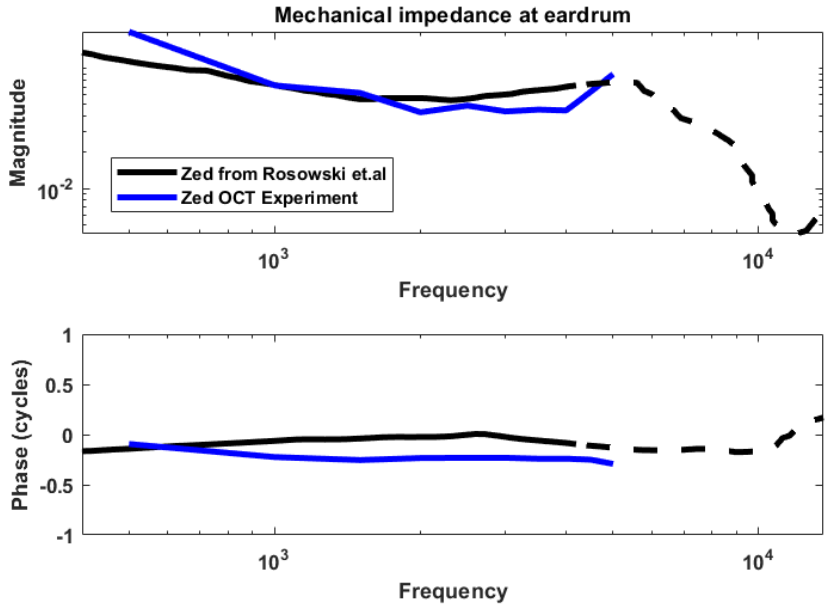


Figure 4. Mechanical impedance at the eardrum of the 43AG7 ear simulator

The mechanical impedance at the eardrum, thus obtained, is compared with the mechanical impedance reported in Rosowski [12] and is seen to agree well at low frequencies (between 700 Hz to 4 kHz)as shown in Figure 4. Additionally, the transfer impedance of the ear simulator is also obtained as a second benchmark. This quantity is derived by dividing the pressure at the eardrum microphone of the 43AG7 system by the volume velocity at the diaphragm. This is compared to the transfer impedance reported in the specifications of the 43AG7 system as shown in Figure 5.

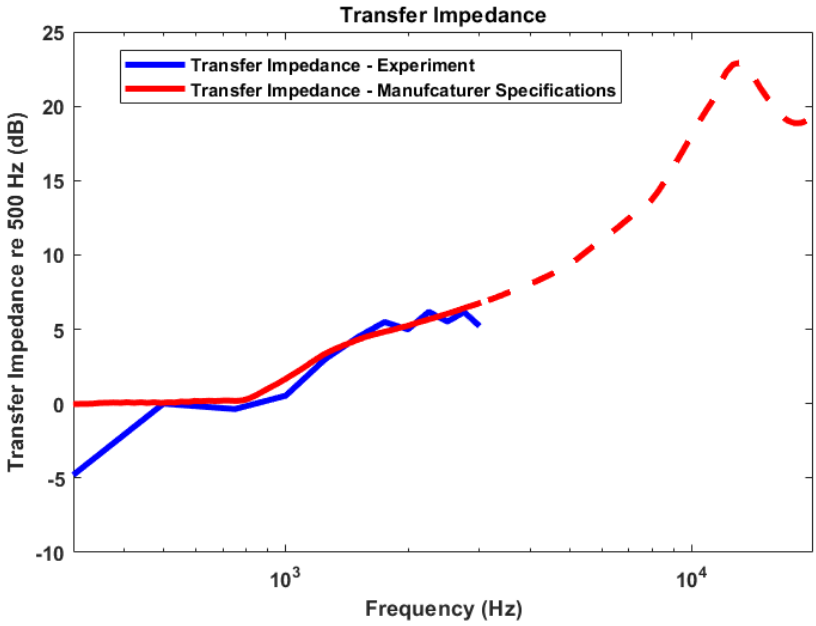


Figure 5. Transfer Impedance of the 43AG7 ear simulator

CONCLUSIONS

Custom-developed vibration measurement capability integrated into a commercial SDOCT system is used along with a designed miniature optical sensor to simultaneously measure pressure and vibration in vibroacoustic systems. The method is validated on a cubical vibroacoustic-cavity. If the sensor diaphragm and other vibrating surface(s) of interest in the system are within the maximum imaging depth of the SDOCT, the method will remain viable. Further, the method is applied to measure the impedance at the eardrum of an ear simulator at low frequencies. The range of frequencies within which the method is used can be further extended by suitably choosing the diaphragm material attached to the ear simulator. While the lack of a vibrating component required the use of a polymer diaphragm in the ear simulator, this would be unnecessary in the actual human ear. The sensor can also be integrated into OCT based middle ear otoscopes. The proposed method could also be used in the cochlea for pressure measurements in the scalae and simultaneous vibration measurements in the associated structures. This would require a potentially ten-fold reduction in the size of the sensor to perform measurements within the complete audible range.

ACKNOWLEDGMENTS

This work was supported by the grant ECR/2016/000514 from the Department of Science and Technology, India, and by a startup grant from the Industrial Research and Consultancy Center, IIT Bombay, India.

REFERENCES

1. Olson ES. “Observing middle and inner ear mechanics with novel intracochlear pressure sensors”. The Journal of the Acoustical Society of America. 1998 Jun;103(6):3445-63.
2. A. N. Popper, Springer Handbook of Auditory Research Series, Richard R. Fay and Arthur N. Popper (eds.). 1996.
3. Robles L, Ruggero MA. “Mechanics of the mammalian cochlea”. Physiological reviews. 2001 Jul 1;81(3):1305-52.
4. Olson ES. “Fast waves, slow waves, and cochlear excitation”. Proceedings of Meetings on Acoustics, ICA2013 2013 Jun 2 (Vol. 19, No. 1, p. 050134). Acoustical Society of America.

5. Ramamoorthy S, Zha D, Chen F, Jacques SL, Wang R, Choudhury N, Nuttall AL, Fridberger A. “Filtering of acoustic signals within the hearing organ”. Journal of Neuroscience. 2014 Jul 2;34(27):9051-8.
6. Gao SS, Raphael PD, Wang R, Park J, Xia A, Applegate BE, Oghalai JS. “In vivo vibrometry inside the apex of the mouse cochlea using spectral domain optical coherence tomography”. Biomedical optics express. 2013 Feb 1;4(2):230-40.
7. Ramdas R, Agarwal N, Atpadikar M, Ramamoorthy S. “Simultaneous measurement of vibration and pressure in vibroacoustic microchannels”. Applied Acoustics. 2020 Dec 1;169:107489.
8. Park K, Cho NH, Jeon M, Lee SH, Jang JH, Boppart SA, Jung W, Kim J. “Optical assessment of the in vivo tympanic membrane status using a handheld optical coherence tomography-based otoscope”. Acta oto-laryngologica. 2018 Apr 3;138(4):367-74.
9. Lin NC, Strimbu CE, Hendon CP, Olson ES. “Adapting a commercial spectral domain optical coherence tomography system for time-locked displacement and physiological measurements”. InAIP Conference Proceedings 2018 May 31 (Vol. 1965, No. 1, p. 080004). AIP Publishing LLC.
10. Wang RK, Nuttall AL. “Phase-sensitive optical coherence tomography imaging of the tissue motion within the organ of Corti at a subnanometer scale: a preliminary study”. Journal of biomedical optics. 2010 Sep;15(5):056005.
11. Kinsler LE, Frey AR, Coppens AB, Sanders JV. “Fundamentals of acoustics”. John wiley & sons; 2000 Jan 12:273-74.
12. Rosowski JJ. “The effects of external-and middle-ear filtering on auditory threshold and noise-induced hearing loss”. The Journal of the Acoustical Society of America. 1991 Jul;90(1):124-35.

Spatial Buildup of Cochlear Compression Revisited

Kostas Kondylidis[1], Anna Vavakou[1] and Marcel van der Heijden[1, a)]

[1]*Erasmus MC, Rotterdam, The Netherlands*
[a)] *marcel[dot]vdh[dot]work[at]gmail[dot]com*

Abstract. In 2013 we published an experimental study (Versteegh and Van der Heijden, doi.org/10.1007/s10162-013-0393-0) that introduced a traveling-wave perspective of cochlear nonlinearity, the so-called spatial buildup of compression and suppression. We recorded basilar membrane (BM) motion in the base of the gerbil cochlea, and obtained experimental evidence for the spatial buildup by independently varying the frequencies of the suppressor and the probe. The study concluded that the nonlinear response at any point in the cochlea can only be properly understood in terms of a spatial buildup of compression. Traveling waves start out as linear responses to the stapes drive and gradually accumulate compressive nonlinearity after entering their peak region, ~½ octave short of their best place. Likewise, the nonlinear effects of intensity on vibration phase can only be understood in terms of local changes in propagation velocity of the wave. Once viewed through this lens, a large body of published data on the effects of compression and suppression, often featuring complex patterns and seemingly contradicting trends, was shown to fit in a unifying framework. Despite the emphasis on spatial buildup, the data underling the 2013 study were obtained from single BM locations, due to the need for reflective microbeads with laser Doppler vibrometry. The spatial aspect was only explored by proxy, i.e., by varying the stimulus frequency and exploiting the tonotopy of the cochlear response ("scaling"). Optical coherence tomography (OCT) vibrometry has liberated us from this limitation, and in the current work we actually traced the spatial buildup of compression. In the BM and OHC region of the 13-24-kHz region of the gerbil cochlea, we recorded responses to equal-amplitude tone pairs centered at 15 kHz, spaced by 20 Hz. The resulting beating pattern affords a dynamic way of obtaining I/O curves. Spectral analysis of the responses yielded an accurate quantitative description of the compressive growth. We observed that the amount of compression (the difference in gain between the peak and dip episodes of the beat) increased gradually and systematically along the course of the slowly propagating wave, which confirms the spatial buildup claimed in our 2013 study. With increasing instantaneous intensity, the spatial rate of compression increased systematically over the nonlinear region. Overall, the conclusions of the 2013 study were confirmed, emphasizing the tight relation between cochlear compression and the wave character of cochlear vibrations.

INTRODUCTION

In the mammalian ear, the huge (>100-dB) dynamic range of audible sounds is compressed to less than 40 dB to fit the much smaller dynamic range of inner hair cell transduction. The exact mechanisms underlying this mechanical sensitivity control are unknown, but outer hair cells play a crucial role. The nature of cochlear compression is very complex, and capturing this complexity in terms of a limited number of unifying principles may be a first step toward understanding the physiological mechanisms. The main characteristics of cochlear compression at the level of the basilar membrane (BM) are:

1) It is very fast but not instantaneous. [1]
2) In essence it is a *multiband* gain control, meaning that sensitivity is independently regulated in different frequency bands [2,3].
3) At high intensities the independence of frequency bands breaks down asymmetrically. Low-frequency components start suppressing the sensitivity to sounds at much higher frequencies (upward spread of masking [4]).
4) At a given BM location the degree of compression, expressed as the slope in dB/dB of I/O functions, is strongly frequency dependent. Components more than ~1/2 octave below the best frequency (BF) show linear (1-dB/dB) growth. With increasing frequency the slopes become increasingly more compressive (< 1dB/dB), and this trend continues beyond the BF [3,5].

A useful methodological tool for the quantitative study of cochlear compression has been two-tone suppression (2TS), a paradigm in which the effect is evaluated of one tone ("suppressor") on the response to another tone ("probe"). In a landmark cochlear-mechanical 2TS study, Cooper reported [6] the suppression of BM responses to near-BF probes by suppressors of different intensity and frequency, and found that the amount of suppression (i.e. the reduction of probe-evoked vibrations) is determined by the RMS displacement of the BM evoked by the suppressor. Suppressors whose frequency varied over a wide range produced the same amount of suppression (and phase shifts) as long as they evoked an equally large their BM displacement.

Versteegh and van der Heijden [7] went on to systematically study the effect of probe frequency (in addition to the effects suppressor frequency and suppressor intensity addressed in previous work). The introduction of a new stimulus dimension may seem like a complication, but in this case led to a marked simplification and unification of the bewildering set of observations in the 2TS literature. The key to the simplification was the use of cochlear tonotopy, which enabled the authors to analyze the *spatial* distribution of compression along the length of the BM. The analysis underscored the essential role of traveling waves in the explanation of the complexities of cochlear compression, an ingredient that had not been systematically discussed before. Specifically, Versteegh and van der Heijden [7] showed that "many seemingly unrelated aspects of compression and suppression are in fact consequences of the gradual accumulation of [nonlinearity] along the travel direction. Likewise, the complex phase effects were found to reflect systematic changes in the local propagation speed of traveling waves." They called this the *spatial buildup* of cochlear compression.

One of the complexities of cochlear compression that finds a natural explanation in the conceptual framework of spatial buildup is the frequency dependence of the growth of compression (point 4 above). This can be seen as follows.

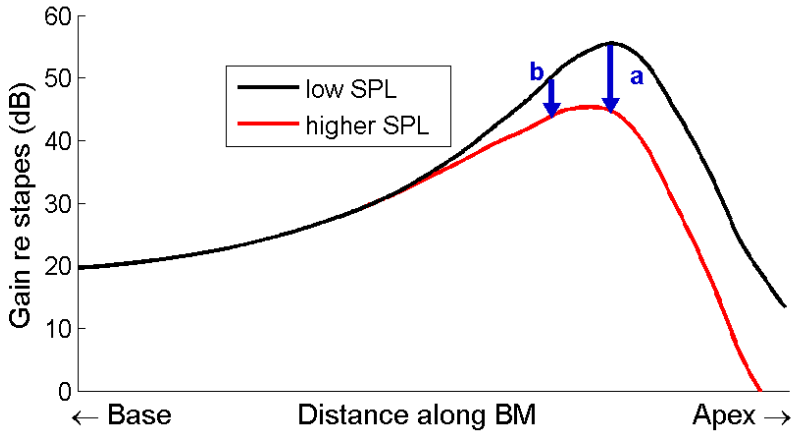


FIGURE 1. Schematic diagram illustrating how compressive growth of BM vibrations depends on longitudinal location. The curves represent the displacement patterns on the BM, normalized to middle-ear motion, evoked by a single tone presented at two intensities. The upper curve represents a low-intensity, linear, response. Increasing the intensity decreases the local forward gain of the traveling wave (e.g. by increased damping), resulting in a progressively smaller gain (lower curve). The arrows mark the amount of compression at two positions (b basal to a). The gradual divergence of the two curves causes the amount of compression to increase from base to apex.

If compression is indeed achieved by a regulation of local forward gain (e.g. through variable damping) the amount of compression for a single tone is seen to gradually accumulate from base to apex (Fig. 1). A simple "scaling" (tonotopy) argument then concludes the explanation of the frequency dependence of the slopes of I/O function at single BM positions.

The experimental data underpinning the spatial buildup of compression, however, exist of single-point vibration recordings on the BM. The spatial aspect was analyzed indirectly, by invoking "scaling", i.e., uniform changes in stimulus frequency are used as a proxy for changing the longitudinal recording position. The use of "scaling" to extrapolate single-point data to spatial excitation patterns is understandably popular among modelers, but its experimental basis is feeble. Neither its quantitative accuracy nor the scope of its validity has been subjected to

dedicated tests. For this reason, Versteegh and Van der Heijden [7] stressed the need for multipoint data to explore the spatial buildup of cochlear compression more directly. With the advent of Optical Coherence Tomography (OCT) as a tool for cochlear vibrometry, it has become much easier to record vibration along multiple locations in the cochlea.

METHODS

Basic idea

I/O functions were obtained by presenting a pair of equal-amplitude tones differing in frequency by 20 Hz. The resulting 20-Hz "beating" pattern affords a quasistatic way of varying the SPL. This dynamic method of obtaining I/O functions is efficient and robust, and 20 Hz is slow enough for the BM response to be free of hysteresis [1]. By recording BM responses along a longitudinal stretch (typically 500 to 1000 μm long), we assessed how the nonlinear compression developed along the propagation direction of the traveling wave.

OCT recording and data analysis

Sound-evoked vibrations were recorded from the basilar membrane of deeply anesthetized gerbils. The bulla was opened and the cochlear partition was viewed through the intact round window membrane. Images (B scans) and vibration measurements (M scans) were obtained using a Thorlabs Telesto III Optical Coherence Tomography (OCT) system at a sample rate of 111.6 kHz. The optical wavelength was 1300 nm and the effective lateral and axial resolution were ~10 μm and ~6 μm, respectively. For detailed methods, see [5]. Vibration responses were analyzed in the Fourier domain by extracting the stimulus-related spectral components and subjecting them to to a Rayleigh

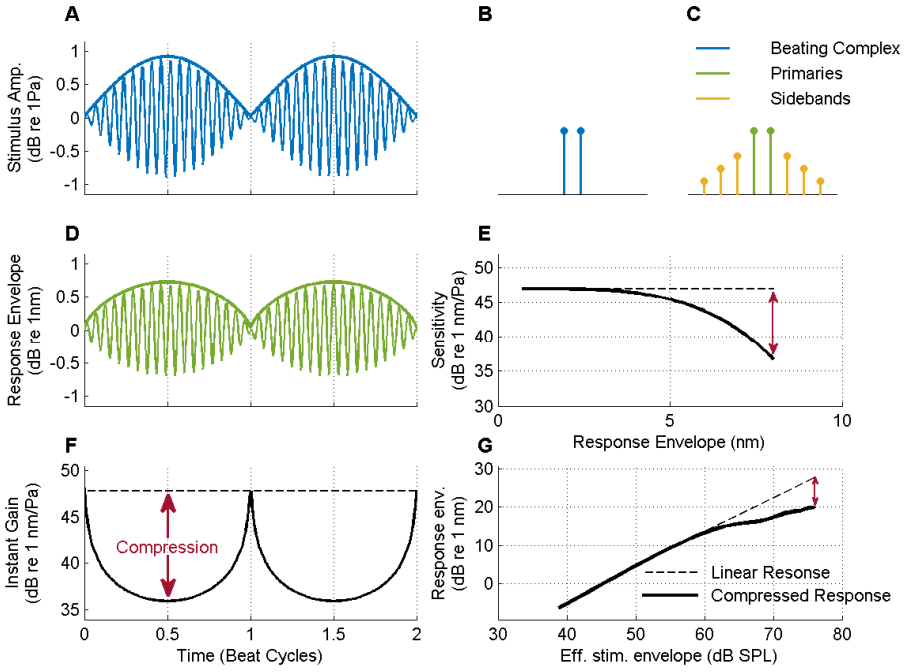


FIGURE 2. Extracting the compressive growth from BM responses to 20-Hz beating tone pairs. **A.** Stimulus waveform. **B.** BM response. **C**. Instantaneous gain determined by dividing the envelopes of the response and stimulus waveforms. Dashed line indicates low-SPL linear response. **D**. Stimulus spectrum. **E**. Response spectrum. The compression shows up as sidebands. **F**. Gain versus response envelope. **G**. I/O function plotted on a log-log scale. Dashed lines in F & G extrapolate the low-SPL linear response; red arrows mark the amount of compression.

test [8]. Only components with (p<0.001) were admitted to the analysis.

The analysis of the response in terms of compression is illustrated in Fig. 2. Acoustic stimuli, presented through an eartube sealed to the ear canal, consisted of beating tone pairs at frequencies f_1 and f_2 with f_2-f_1= 20 Hz (Fig. 2A). From the Fourier spectrum of the BM response the Rayleigh-significant spectral components were extracted at frequencies f_1+n(f_2-f_1), with integer n. The 20-Hz envelope of the resulting response waveform were extracted (Fig. 2B). The instantaneous sound- pressure-to-BM gain (Fig. 2C) was determined by comparing the response envelope (in nm) to the stimulus envelope (in Pa) at the corresponding instant within the beat cycle. As expected for a compressive response, the instantaneous gain decreased monotonically with response envelope (Fig. 2F). The I/O functions (Fig. 2G) show linear growth at low SPLs and compression at higher SPLs.

RESULTS

In Fig. 3 the responses to the same two-stimulus, centered at 15 kHz, is compared between two longitudinal locations on the BM spaced by 600 µm. The response at the basal location (BF=20.8 kHz; ochre curves) was linear, yielding a sinusoidal response envelope (Fig. 3A), constant gain (fig. 3B), absence of spectral sidebands (Fig. 3C), flat sensitivity curve (Fig. 3D) and a constant, 1-dB/dB slope of the I/O function (Fig. 3E). In contrast, the more apical location (BF = 15.2 kHz; blue curves) showed a more rounded response envelope, a periodically varying gain, many spectral sidebands, a downward sensitivity curve and compressive (< 1-dB/dB) I/O function.

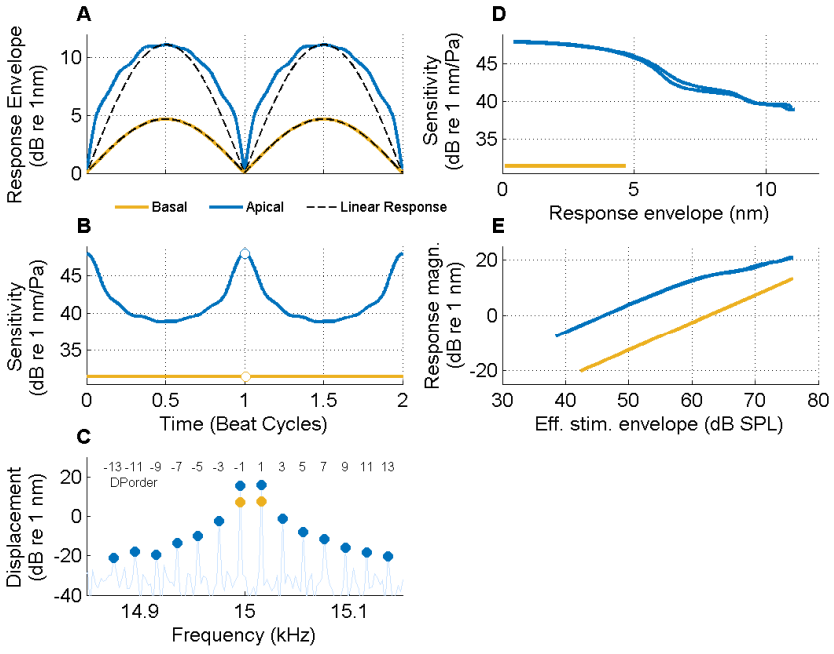


FIGURE 3. Comparing the BM response to the same stimulus between two longitudinal locations spaced by 601 µm. **A.** Envelope of the two responses (*colored lines*), with the normalized stimulus envelope (*dashed black lines*) shown for comparison. For the basal response (*ochre*) the envelopes of the stimulus and response have identical shapes, indicating linearity. The apical response (*blue*) is more flat-topped than the stimulus envelope, indicating compressive nonlinearity.. **B**. Instantaneous gain determined by dividing the envelopes of the response and stimulus waveforms. The linearity of the basal response is reflected by the constant gain, whereas the compressive nonlinear growth at the apical location causes the gain to systematically decrease with instantaneous intensity.ear response. **C**. Response spectrum. The normalized frequency axis is based on the (signed) order of the distortion products / sidebands. **E**. Gain versus response envelope. **F**. I/O function plotted on a log-log scale. The linearity of the basal I/O functions (*ochre*) is reflected by the constant 1-dB/dB slope.

This comparison of the responses evoked by the same single stimulus, recorded at two longitudinal positions, confirms that the amount of nonlinear compression changes along the course of the traveling wave. A more comprehensive analysis is shown in Fig. 4, showing data from 10 location spanning 720 µm of BM length. On its way to the peak region, the traveling wave acquired an increasing degree of nonlinear compression (Fig. 4A,B). In the spectral domain this buildup of compression is reflected by a systematically increasing number of sidebands (Fig. 4C) and a systematic increase of the sideband-to-carrier ratio (Fig. 4F), a spectral metric of the degree of compressive nonlinearity [9].

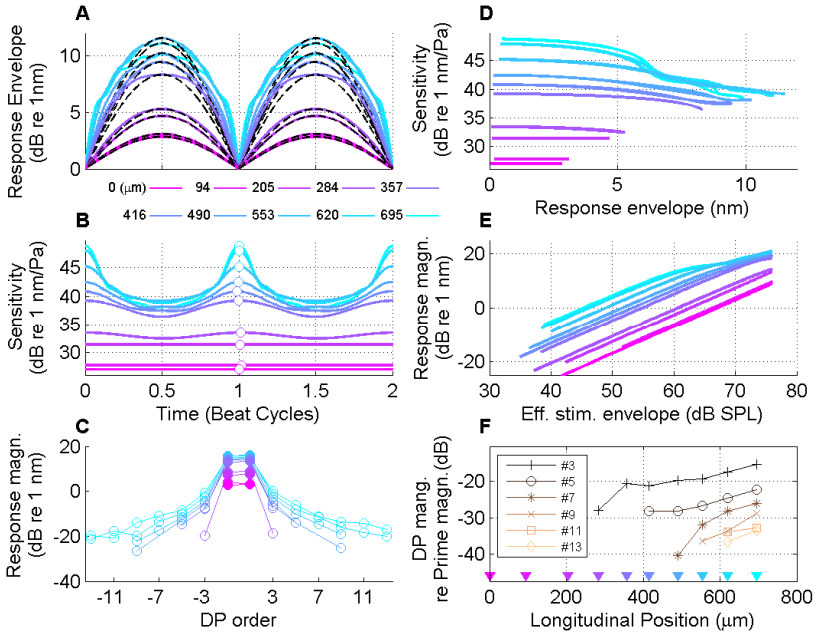


FIGURE 4. Compressive responses to a tone pair centered at 15 kHz compared across an array of longitudinal BM locations. **A-E**, same format as Fig. 2, but now showing data from 10 BM locations spanning 720 µm of BM length as indicated in panel A. **F**. Sideband-to-carrier ration as a function of longitudinal position (from base to apex). The order of the distortion products (i.e., sidebands generated in the cochlea) is indicated in the legend. Triangles mark the locations corresponding to the respective curves in panels A-E. [Dataset RG211043/k19_15k_70]

Figure 5 presents an entirely spatial perspective of the data of Fig. 4. The magnitude of the 15-kHz traveling wave is seen to grow over the 720-µm stretch of BM length in an SPL-dependent way (Fig. 5A). For the lowest SPLs, the total growth was 21 dB; for the highest SPL the total growth reduced to 10 dB. The amount of compression (Fig. 5B) is seen to gradually build up toward the peak of the wave over a longitudinal stretch of ~0.5 mm, turning from complete linearity (X=200 µm) into strong compression (X=720 µm, close to the best place of the 15-kHz stimulus). The amount of compression increased systematically with increasing SPL at all locations. From Fig. 5B it is clear that the dominant factor shaping the final compression (at best place) is the steepness of its accumulation. For instance, for the 76-dB-SPL case, the spatial rate of compression was 21 dB/mm, while for the 66-dB-SPL case the spatial compression rate was only 10 dB/mm. The data of Fig. 5 suggest that a secondary factor shaping the final compression is the spatial extent of the buildup. Specifically, compression appears to arise somewhat "earlier" (at a more basal position) for higher SPLs than for low SPLs. This apparent widening of the compressive region with increasing SPL requires further analysis based on a larger set of data.

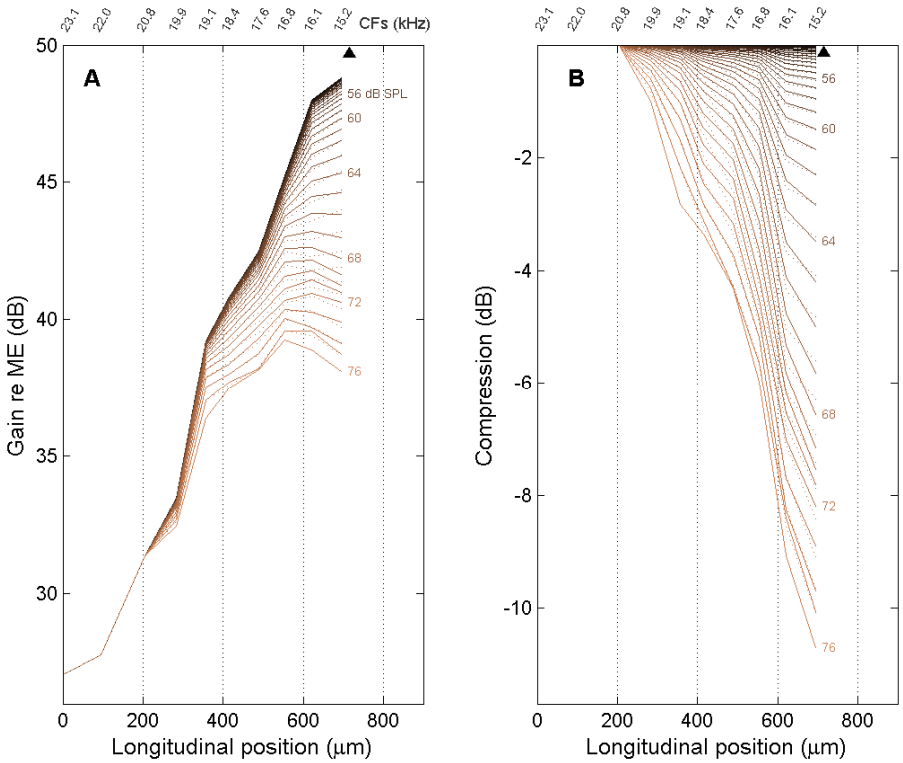


FIGURE 5. Spatial buildup of compression for a 15-kHz stimulus. **A**, Vibration magnitude at the BM, normalized to stapes motion, as a function of longitudinal position. BF of the locations indicated in upper X-axis. Each curve represents an instantaneous SPL as indicated in graph. *Triangle*, best site of the 15-kHz stimulus. **B**. Corresponding amount of compression (gain normalized to low-SPL gain). [Dataset RG211043/k19_15k_70]

DISCUSSION

Overall, the nonlinear longitudinal patterns of BM vibration confirm the spatial buildup of compression inferred from single-point recordings by Versteegh & van der Heijden [5]. Nonlinear growth at best place is indeed the culmination of a gradual buildup of compressive nonlinearity during wave propagation. In the base of the gerbil the nonlinear region starts ~0.5 mm basal to the best place. This coincides with the slow part of the traveling wave, and is consistent with a framework in which the slow wave propagation is exploited as a means to regulate sensitivity, most likely by controlling the amount of local damping in an SPL-dependent way. As discussed more comprehensively in [5,10,11], an attractive feature of this framework is the clear functional role of traveling waves in the cochlea. In an alternative scheme, a parallel or "filterbank" design, each local filter would need to be strongly compressive by itself, whereas in a series or "cascade" design the workload of dynamic compression is neatly shared among a spatially distributed collection of regulators. The cumulative operation of a cascade ensures that each local link has only a modest role in the final result. Mathematically speaking, the effect of a cascade is characterized by exponential behavior (as a function of place), and small variations of the exponential coefficient (local damping) suffice to realize a considerable regulation of net gain. This is both simple and efficient, but the cascade design comes with a price, namely a potential interference across frequency bands in the form of suppression (physiology) and masking (psychophysics). This is the breakdown of the independence of sensitivity control across frequency bands mentioned in the Introduction: high-intensity, lower-frequency components can strongly reduce the audibility of sound at (much) higher frequencies, an effect known as "upward spread of masking" [4].

The gradual change of forward gain (and thus, rate of compression, Fig. 5B) with increasing SPL is consistent with a scheme in which local damping is adjusted on the basis of local vibration magnitude. A plausible

implementation of such negative feedback is the regulation of local impedance controlled by outer hair cell (OHC) length changes. In such an automatic gain control scheme, the rectification of the OHC receptor current plays an essential role, as it generates an envelope-following receptor potential that survives the low-pass filtering by the cell membrane, whose corner frequency is very low (~3 kHz) compared to the frequencies being processed in the basal turn of the gerbil (13-25 kHz).

The presence of low-pass filtering poses no problem to an automatic gain scheme, and in fact is a necessary ingredient of such schemes. On the other hand, schemes based on high-frequency cycle-by-cycle OHC feedback are problematic because the filtering reduces high-frequency variations of the receptor potential to insignificant amplitudes [11,12,13]. The cascaded nature of cochlear compression aggravates the problems for cycle-by-cycle feedback at high frequencies: a serial arrangement of amplifiers leads to an exponential deterioration of signal to noise ratio, because the noise introduced by every link of the chain is amplified by the next link. Near hearing threshold the high-frequency AC receptor potential of OHCs is minute (<1 μV) [11], so the signal-to-noise ratio of each OHC is poor. Therefore, if cycle-by-cycle feedback would provide a major driving force, the serial arrangement would cause an exponential growth of internal noise. Rather than improving the detection of soft sounds, the cycle-by-cycle feedback would worsen it.

In order to put the spatial buildup of compression on firmer footing, more data are needed than reported here. In particular it is important to test the effect of stimulus frequency. The spatial buildup framework predicts a tonotopic shift of spatial compression patterns (as shown in Fig. 5) when varying stimulus frequency: the onset of nonlinearity is also expected to shift. These effects can also be inferred from the large literature of BM recordings, but it is important to provide data obtained with a single, consistent methodological framework, and to test the effects of experimental parameters (stimulus frequency; spatial location) within single ears as much as possible. With regards to computational models of the cochlea, it would be interesting to explicitly test their predictions of the type of spatial response patterns reported here. Most of the modeling literature still presents “classical” single-point predictions, and has not tuned into the much richer world of OCT data.

Work supported by Netherlands Organization for Scientific Research, ALW 823.02.018, and an EU Horizon 2020 Marie Sklodowska-Curie Action Innovative Training network, H2020-MSCA-ITN-2016, LISTEN – 722098.

REFERENCES

1. N. P. Cooper and M. van der Heijden, ”Dynamics of Cochlear nonlinearity.” In *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing*. Springer, Cham, pp. 267–273 (2016).
2. N. P. Cooper, “Compression in the peripheral auditory system.” In: Bacon SP, Fay RR, Popper AN (eds) *Compression: from cochlea to cochlear implants*. Springer, New York, pp 18–61 (2004) .
3. L. Robles and M. A. Ruggero, “Mechanics of the mammalian cochlea,” Physiol Rev **81**, 1305–1352 (2001).
4. R. C. Bilger and I. J. Hirsh, “Masking of tones by bands of noise,” J. Acoust. Soc. Am. **28**, 623-630 (1956).
5. N. P. Cooper, A. Vavakou, and M. van der Heijden, “Vibrationa hotspots reveal longitudinal funneling of sound-evoked motion in the mammalian cochlea,”, Nat. Commun. **9**, 1–12 (2018).
6. N. P. Cooper, “Two-tone suppression in cochlear mechanics,” J Acoust Soc Am **99**, 3087–3098 (1996).
7. C. P. Versteegh and M. van der Heijden, “The spatial buildup of compression and suppression in the mammalian cochlea,” J. Assoc. Res. Otolaryngol. **14**, 523–545 (2013).
8. C. P. C. Versteegh and M. van der Heijden, “Basilar membrane responses to tones and tone complexes: nonlinear effects of stimulus intensity,” J Assoc Res Otolaryngol **13**:785–798 (2012).
9. M. van der Heijden, “Cochlear gain control,” J. Acoust. Soc. Am. **117**, 1223-1233 (2005).
10. M. van der Heijden, C. P. C. Versteegh,“ Energy flux in the cochlea: evidence against power amplification of the traveling wave. J. Assoc. Res. Otolaryngol. **16**, 581–597 (2015).
11. M. van der Heijden and A. Vavakou, “Rectifying and sluggish: Outer hair cells as regulators rather than amplifiers,” Hear. Res. in press, https://doi.org/10.1016/j.heares.2021.108367 (2021).
12. A. R. Cody and I. J. Russell, “The response of hair cells in the basal turn of the guinea-pig cochlea to tones,” J. Physiol. **383**, 551–569 (1987).
13. A. Vavakou, N. P. Cooper, and M. van der Heijden, “The frequency limit of outer hair cell motility measured in vivo,” Elife **8**, e47667 (2019).

The Timing of the Cochlear Wave Propagation: A Comparative Study on How Computational Models Reproduce the Phase Response of Excitation Patterns

Amin Saremi[1,a)] and Davood Khodadad[1]

[1] *Department of Applied Physics and Electronics, Umeå Uiversity, 901 87 Umeå, Sweden.*

[a)] *Corresponding author: amin.saremi@umu.se*

Abstract. It is believed that the phase of the cochlear wave propagation might play a crucial role in binaural perception and sound localization by generating cochlear disparities. Experimental data demonstrate that, while a tone is being played, an excitation pattern is formed extending from the base to the apex of the cochlea. The phase of the excitation pattern decays along the cochlear length until the slope of the phase curve (i.e. group delay) reaches its maximum at a location with the characteristic frequency (CF) that matches the frequency of the input tone. Thereafter, the phase stays almost constant (group delay equals zero) until the apex. Computational models have been aiming to simulate the cochlear responses and thereby illuminate the underlying electromechanics of the human inner ear. These computational models can be divided, according to their topology, into two groups: Parallel filterbanks that model the cochlea as several independent filters versus cascade filterbanks (including transmission lines) that assume that the filters are coupled in series. Due to their modeling principles, cascade filterbanks intrinsically include the longitudinal traveling wave propagation whereas the parallel filterbank models lack this feature by default since there is no longitudinal relation between the filter stages in these models. The objective of this study was to verify if cascade filterbanks are actually more successful in simulating the phase responses than parallel filterbanks. The excitation patterns generated by seven cochlear models (four parallel filterbanks, two cascade filterbanks and a transmission-line model) in response to 4 and 9 kHz tones were estimated using an impulse, and the results were compared with corresponding experimental data recorded at comparable lengths in living mammalian cochleae. The accuracy of the model predictions was reported in mean absolute error (MAE) relative to the experimental data. The results show that cascade filterbank models are remarkably more successful with CARFAC and VERHULST models predicting the experimental data most closely. However, DRNL model (a parallel filterbank) also produces outcomes that are comparable in accuracy with those generated by cascade filterbank models. Further investigations showed that this specific parallel filterbank model externally incorporated correct phase delays in the impulse responses of its filter stages. The results indicate that if parallel filterbanks incorporate the phase delays in their impulse responses according to the proposed method, they could successfully simulate the timing of the longitudinal wave propagation along the cochlea within the same accuracy range as cascade filterbank models do.

INTRODUCTION

The human auditory system is capable of distinguishing sounds with a just-noticeable frequency resolution of 1/230 octave, approximately ±3Hz at 1 kHz. (Patterson and Moore, 1986; Glasberg and Moore, 1990). This remarkable frequency sensitivity is believed to arise in the early stages of the auditory system, from the biomechanics of the cochlea. When a mammalian cochlea is subject to a sound pressure field, a displacement wave (i.e. traveling wave) is formed along the cochlear structures propagating from base to apex. The envelope of the traveling wave (i.e. excitation pattern) peaks at a certain location depending on the frequency of the input sound (von Bekesy, 1960; Ruggero et al., 1997). Consequently, every location on the cochlea is attributed to a certain frequency, known as the characteristic frequency (CF) (von Bekesy, 1960; Greenwood, 1990). This location-frequency selectivity of the cochlea is believed

to explain the remarkable capability of the human auditory system in distinguishing sound frequencies (Patterson and Moore, 1986; Glasberg and Moore, 1990).

Another observation from experimental measurements is that the phase of the traveling wave envelope (excitation pattern) decays slowly before it reaches the CF site where the phase decay becomes much sharper (de Boer and Nuttall, 2000; Ren, 2002). The phase decay of the traveling wave determines the time delay of the cochlear wave propagation since the time delay of a quasi-linear system is characterized by the first derivative of its phase response (i.e. group delay) (Hammond and White, 2008).

It had been generally unknown how the timing of the cochlear traveling wave propagation contributes to the auditory perception until new evidences emerged that indicated its crucial role in binaural hearing. A remarkable feature of the human auditory system is its sensitivity to differences between the neuronal spikes received from the two ears based on interaural time difference (ITD). For each medial superior olive (MSO) cell connected to a pair of specific locations on left and right cochleae, only neuronal spike trains that are shifted by a certain interaural time difference (known as best delay, BD) activate the binaural cues maximally (Shroeder, 1977; Joris et al. 2006). This gives an ITD tuning that is frequency dependent and therefore cannot be explained by a fixed delay line (Van der Heijden et al., 2014). A theory called 'cochlear disparity', introduced by Shroeder (1977) and later backed by measurements (Joris et al. 2006), suggests that BD ITDs are created between the two ears because each monaural input is taken from a slightly different location on the cochlea (left versus right) and thereby producing an intrinsic ITD as a function of CF.

Computational models of cochlear mechanics could be traced back to Helmholtz works on independent resonator theory (Helmholtz, 1895). Several models have been built on this concept which regards the cochlea as multiple independent filters that mimic a second-order harmonic resonator with a specific resonance frequency for each one (e.g. Lopez-Poveda and Meddis, 2001; Hohmann, 2002; Irino and Patterson, 2006; Zilany et al. 2014). Since the filters are independent of each other and are separately excited by the input stimuli, this configuration is called 'parallel filterbank'.

Besides, several models have been developed that regard the cochlea as a signal transmission line (e.g. Zweig et al., 1976; Verhulst et al., 2012) or multiple filters in series (e.g. Lyon, 2011; Saremi and Stenfelt, 2013). Accordingly, the input stimuli excites firstly the filter which represents the most basal region of the cochlea and the produced response travels to subsequent filters until it reaches the most apical filter. Since the filters are organized in series from base to apex, this configuration is called 'cascade filterbank'.

As the sound-induced signal travels across the filter stages of a cascade filterbank, each filter contributes to the timing of the overall response by delaying the signal components based on its phase response. In other words, the time delays are intrinsically embedded in a cascade filterbank whereas in a parallel filterbank such delays do not exist by default because there is no longitudinal coupling between different filter stages. As a result, it is expected that cascade filterbank models reproduce the phase and timing of the traveling wave more accurately than parallel filterbank models.

To address this question, the excitation patterns were estimated in 7 publicly-available models of the human cochlear mechanics consisting of 4 parallel filterbanks (DRNL, Gammatone, Gammachirp, and Zilany) versus 3 cascade filterbanks (CARFAC, Verhulst, and Saremi) and compared with the corresponding available experimental data.

METHODS

Seven models of human cochlear mechanics were retrieved comprising of four parallel filterbanks versus three cascade filterbank models. The parallel filterbanks are: **DRNL** (by Lopez-Poveda and Meddis, 2001), **Gammatone** (by Hohmann, 2002), **Gammachirp** (by Irino and Petterson, 2005), and **Zilany** (by Zilany et al., 2014). The cascade filterbanks are: **Verhulst** (by Verhulst et al., 2012), **CARFAC** (by Lyon, 2011), and **Saremi** (by Saremi and Stenfelt, 2013). Appendix I of Saremi et al. (2016) introduce these models in details and also provides information on where to find the source codes.

Since the human cochlea is inaccessible in vivo, data from other mammalian cochleae have been extrapolated to humans based on the fundamental assumption that all mammalian cochleae share analogous mechanics at similar locations on the cochlear duct with respect to relative distance from base and apex [see Robles and Ruggero (2001) for a review]. To our knowledge, the only available in vivo measurements of the longitudinal excitation patterns have been reported by Ren (2002) and He et al. (2018) in the base of the Gerbil cochleae. These data have been used as the experimental reference in this study to be compared with the model predictions.

Ren (2002) and He et al (2018) stimulated Gerbil cochleae in vivo by 16 and 26 KHz sinusoids, respectively, and reported longitudinal excitation patterns across the cochlear length. These frequencies maximally excite 33% and 17% of the cochlear length from the base, respectively. In humans, these locations are attributed to CFs of 4 and 9 kHz (Greenwood, 1990), as marked in Fig. 1. Therefore, according to the fundamental assumption, to evoke excitation patterns in humans that could be compared to those measured in Gerbils, sinusoids at 4 and 9 kHz should be used as stimuli.

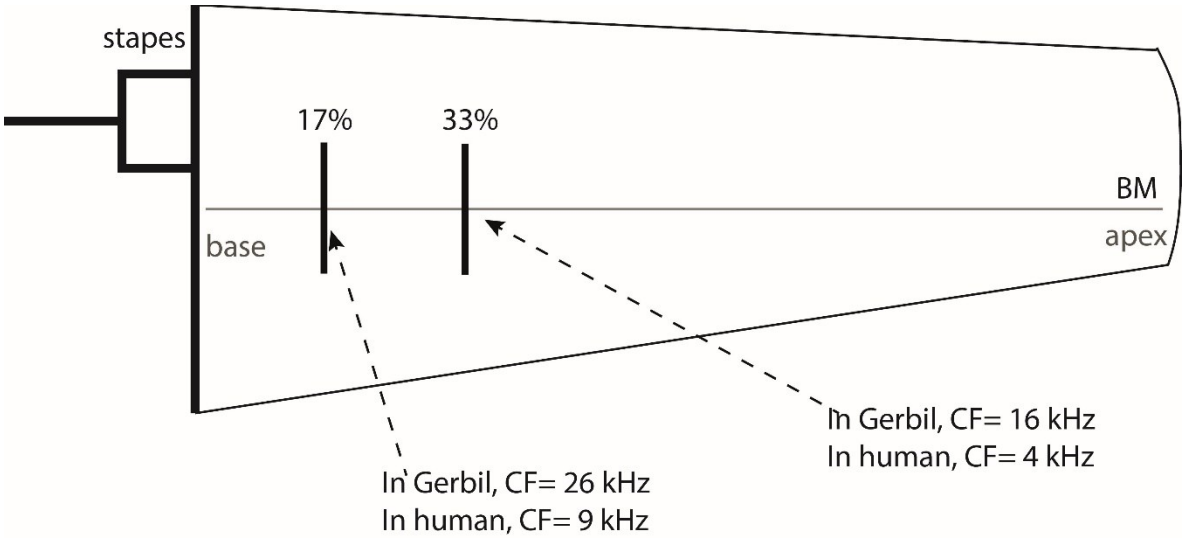


FIGURE 1. The locations where the stimuli used in this study maximally excite across in the cochlea, compared between human and gerbil cochleae.

The phases of the excitation patterns were estimated by feeding a 80-µs condensation impulse at 30 dB SPL into the models. The impulse responses of the models at all filter stages were transformed to the frequency domain using the fast Fourier transform (FFT) instruction in MATLAB 2017b. The phases of the excitation patterns were estimated for each model by calculating the phase delay of each filter stage along the cochlea at 4 and 9 kHz.

The phases of the excitation patterns in response to 4 and 9 kHz sinusoid were compared with the corresponding data reported in Fig. 1(B) of Ren (2002) and Fig. 5 of He et al. (2018) respectively where the phase delays were measured across the cochlea in response to 16 and 26 kHz sinusoids using a laser interferometer and a locked-in amplifier. To quantify how closely each model reproduces the phase delays, we used the mean absolute error (MAE) according to Eq. (1) below where M_i and R_i are respectively the model prediction and the reference phase delay measured at point i, and N is the total number of measurement points.

$$MAE = \frac{\sum_{i=1}^{N} |M_i - R_i|}{N} \quad (1)$$

RESULTS

Figure 2 shows the phase of the excitation pattern generated by the models in response to a 4 kHz, 30 dB SPL sinusoid. The vertical dashed line marks the corresponding CF cite i.e., 33% of the cochlear length from the base and the crosses represent the experimental data reported in Fig. 1 of Ren (2002) shown as a function of the relative distance from the base. Figure 3 shows the phase of the excitation pattern generated by the models in response to a 9 kHz, 30 dB SPL sinusoid. The vertical dashed line marks the corresponding CF site i.e., 17% of the cochlear length from the base and the crosses represent the experimental data shown as a function of the relative distance from the base. The experimental data (Fig. 5 of He et al. (2018)) were originally reported in degrees which were converted to radian and reproduced in Fig. 3.

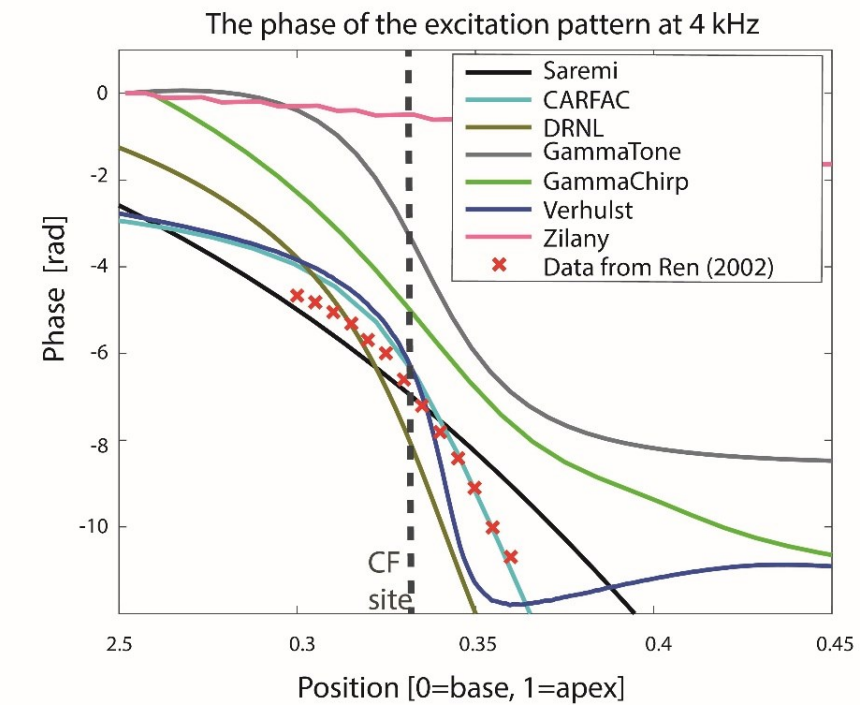


FIGURE 2. (Colored) The phase of the excitation pattern in response to a 4 kHz sinusoid estimated by models versus the corresponding experimental data (Ren (2002)).

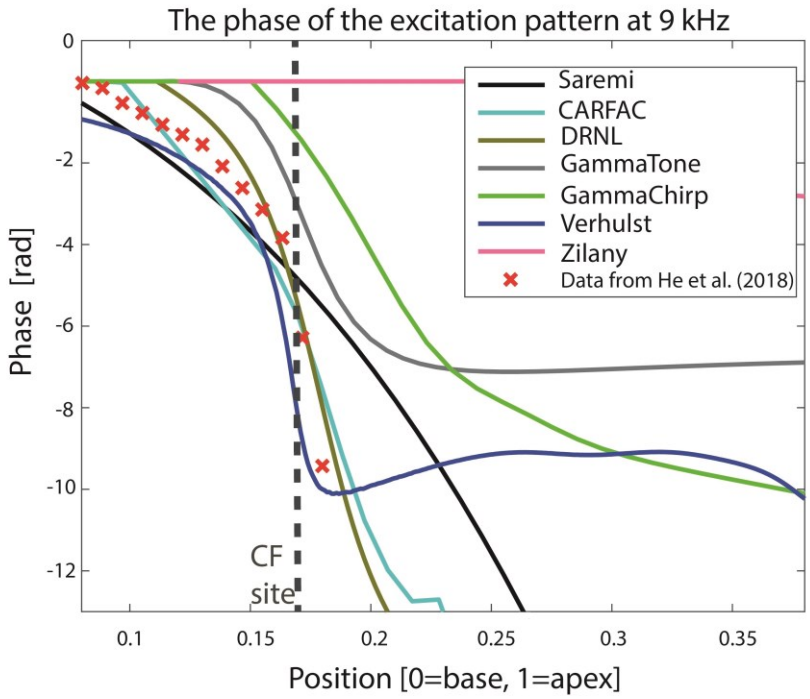


FIGURE 3. (Colored) The phase of the excitation pattern in response to a 9 kHz sinusoid estimated by models versus the corresponding experimental data (He et al. (2018)).

Figure 2 and Fig. 3 illustrate that all models show a decaying phase curve as a function of distance from base except for Zilany. This model appears to be insensitive of phase differences across the filters. The MAEs, as a measure of the deviation between model prediction and the experimental reference data, are reported in Table I. The values shown in Fig. 2 and Fig. 3 were inserted in Eq. (1) to estimate the MAEs.

TABLE I: The deviations of the model predictions with respect to the corresponding experimental data in terms of MAEs. All values are in radians.

Model	CARFAC	DRNL	Gammachirp	Gammatone	Saremi	Verhulst	Zilany
4 kHz	0.15	1.16	1.36	4.12	0.91	0.85	6.02
9 kHz	0.73	0.74	2.24	1.4	0.79	0.71	2.66

DISCUSSIONS

Table I also reports that the DRNL model reproduces the experimental data with 1.16 and 0.73 MAE radians at 4 and 9 kHz, respectively. These numbers are close to the average MAE scores by cascade filterbank models. In other words, although DRNL is a parallel filterbank model it can reproduce the phase of the excitation pattern with similar accuracy as the cascade models. Besides, Gammatone and Gammachirp models also depict a sloping phase decay similar to the trend by the experimental data although their accuracy is much worse than that of the DRNL model. The output from a filter stage in the DRNL model is the sum of the signals from the output of two separate filter paths i.e. a linear and a nonlinear path. As seen in Fig. 3 of Lopez-Poveda and Meddis (2001), both of these paths include Gammatone filter components.

In a parallel filterbank model, for a single filter to accurately simulate the phase delay at a given location in the cochlea, the correct time delay value should be externally incorporated into its impulse response according to Eq. 2(a) below. In Eq. 2(a), $\widehat{h_i}[n]$ and $h_i[n]$ are respectively the discrete impulse responses of the i^{th} filter stage of the mode before and after incorporating the time delay, D_i, which gives the time it takes for the excitation pattern to travel from stapes to the location on the cochlea that corresponds to the i^{th} filter stage.

$$\widehat{h_i}[n] = h_i[n - D_i] \quad 2(a)$$

$$\widehat{H_{j\omega}}[\mathrm{n}] = H_{j\omega}[\mathrm{n}]\, e^{-j\omega D_i} \quad 2(b)$$

$$\varphi_i(\omega) = -\omega D_i \quad 2(c)$$

Equation 2(b) is the frequency-domain counterpart of Eq. 2(a) where ω is the angular frequency (i.e. ω=2πf). Equation 2(c) describes the relation between the time delay (D_i) and the corresponding phase delay ($\varphi_i(\omega)$) of the i^{th}filter stage in the model. Gammatone model is a linear parallel filterbank model (i.e. insensitive to the intensity of the stimuli) whose filters were first defined by Patterson et al. (1992) according to the following Eq. 3 where *k* is gain, *n* is the order of the filter, *B* is its bandwidth and f_c is its center frequency.

$$h(t) = kt^{(n-1)}\, e^{-2\pi Bt}\cos(2\pi f_c t + \varphi) \quad (3)$$

Equation (3) also includes a phase delay (denoted by φ) in the impulse response of the filter which acts similarly to the phase delays introduced in Eq. (2). However, if this term was the only source of the phase delay, it would have resulted in a decaying line versus frequency with a slope of (dφ/dω=-D), which is not the case considering the solid gray line in Fig. 2 and Fig. 3. Another contributing factor to the observed shape of the phase decay is that Eq. (3) creates an asymmetric band pass filter (i.e. with asymmetric amplitude and phase responses around f_c which modifies the phase response of the filter from a simple decaying line leading to the observed curves.

Both Gammachirp and DRNL models were devised based on Gammatone with the purpose of introducing the cochlear nonlinearity to that model. This is the reason why Gammachirp and DRNL also depict decaying phase delays in Fig. 2 and Fig. 3. In contrast, the Zilany model which is not based on the Gammatone model, doesn't show sensitivity with regard to the phase of the excitation pattern. This suggests that this model has not incorporated an

explicit phase term into its impulse response (unlike suggested in Eq. (2)) and thus is not suitable for applications that require simulating the phase and timing of the cochlear wave propagation along the cochlea.

CONCLUSION

The timing of the wave propagation along the human cochlea might play a key role in binaural hearing. It is expected that cascade filterbank models reproduce the timing of the wave propagation more accurately than the parallel filterbank models. To estimate the timing of the cochlear wave propagation, the phase of the excitation pattern along the cochlear duct in response to 4 and 9 kHz sinusoids at 30 dB SPL was assessed in three cascade filterbank models versus four parallel filterbank models. The results were compared with corresponding measurements recorded from comparable locations in the living Gerbil cochleae.

Our findings suggest that cascade filterbank models simulate the phase of the excitation pattern and thereby the timing of the wave propagation more accurately with CARFAC and Verhulst models producing the most accurate results. However, we found out that a parallel filterbank model could also simulate the timing of the wave propagation if correct phase delays are externally incorporated in every filter stage of the model according to the proposed Eq. (2). We observed that DRNL, a parallel filterbank model, produces phase delays that are comparable in accuracy with those generated by cascade filterbank models.

REFERENCES

1. Békésy, G. (**1960**). "Part 4: Cochlear Mechanics" in Experiments in the Hearing (McGraw-Hill, New York, USA), pp 403-703.
2. de Boer, E., and Nuttall, A. L. (**2000**). "The mechanical waveform of the basilar membrane. III. Intensity effects," J. Acous. Soc. Am. **107**, 1497-1507.
3. Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hear. Res. **47**. 130-138.
4. Greenwood, D. D. (**1990**). "A cochlear frequency-position function for several species - 29 years later," J. Acoust. Soc. Am. **87**, 2592-2605.
5. Hammond, J., and White, P. (**2008**). "Signals and systems," in Handbook of signal processing in acoustics, 1st ed. (Springer, USA), pp. 3-16.
6. He, W., Kemp, D., and Ren, T. (**2018**). "Timing of the reticular lamina and basilar membrane vibrations in living Gerbil cochleae," Res. Aricles, 1-17.
7. Helmholtz, H. L. F. (**1885**). "On the sensations of tone as a physiological basis for the theory of music," Second English Edition, translated by Alexander J. Ellis. London: Longmans, Green, and Co., UK., pp. 1-44.
8. Hohmann, V. (**2002**). "Frequency analysis and synthesis using a Gammatone filterbank," Acta Acustica. **88**, 433-442.
9. Irino, T., and Patterson, R. D. (**2006**). "A dynamic compressive Gammachirp auditory filterbank," IEEE Trans Audio Speech Lang Processing. **14**(6), 2222–2232.
10. Joris, P. X., Van de Sande, B., Louage, D. H., and van der Heijden, M. (**2006**). "Binaural and cochlear disparities, " PNAS. **13**. 12917–12922.
11. Lopez-Poveda, E. A., and Meddis, R. (**2001**). "A human nonlinear cochlear filterbank," J. Acoust. Soc. Am. **110** (6), 3107-3118.
12. Lyon, R., F. (**2011**). "Cascades of two-pole–two-zero asymmetric resonators are good models of peripheral auditory function," J. Acoust. Soc. Am. **130**(6), 3893–3904.
13. Patterson R. D., and Moore, B. C. J. (**1986**). "Auditory filters and excitation patterns as representations of frequency resolution," in: Moore, B.C.J. (Ed.), Frequency Selectivity in Hearing. Academic, London, UK, pp. 123-177.
14. Patterson, R., Robinson, K., Holdsworth, J., McKeown, Zhang, C., and Allerhand, M. (**1992**). "Complex sounds and auditory imaging," in Auditory phsyiology and perception, (Oxford, UK). pp. 429-443.
15. Ren, T. (**2002**). "Longitudinal pattern of basilar membrane vibration in the sensitive cochlea," Proc. Natl. Acad. Sci. USA. **99**(26), 17101–17106.
16. Robles, L., and Ruggero, M. A. (**2001**). "Mechanics of the mammalian cochlea," Physiol. Rev. **81**, 1305-1352.

17. Saremi, A., and Stenfelt, S. (**2013**). "Effects of metabolic presbyacusis on cochlear responses: A simulation approach using a physiologically-based model," J. Acoust. Soc. Am. **134** (4), 2833-2852.
18. Saremi, A., Beutelmann, R., Dietz, M., Ashida, G., Kretzberg, J., and Verhulst, S. (**2016**). "A comparative study of seven human cochlear filter models," J. Acoust. Soc. Am. **140**, 1618-1634.
19. Schroeder, M. R. (**1977**) in Psychophysics and Physiology of Hearing, (Academic, New York), pp. 455–467.
20. Van der Heijden, M., Lorteije, J. A. M., Plauska, A., Roberts, M. T., Golding, N. L., and Borst, J. G. G. (**2013**). "Directional hearing by linear summation of binaural inputs at the medial superior olive," Neuron. 78(5). 936-948.
21. Verhulst, S., Dau, T., and Shera, C. A. (**2012**). "Nonlinear time-domain cochlear model for transient stimulation and human otoacoustic emission," J. Acoust. Soc. Am. **132** (6), 3842-3848.
22. Zilany, M. S., Bruce, I. C., and Carney, L. H. (**2014**). "Updated parameters and expanded simulation options for a model of the auditory periphery," J. Acoust. Soc. Am. **135**, 283-286.
23. Zweig, G., Lipes, R., and Pierce, J. R. (**1976**). "The cochlear compromise," J. Acoust. Soc. Am. **59**, 975–982.

Cochlear Wave Propagation and Dynamics in the Human Base and Apex

Model-Based Estimates from Noninvasive Measurements

Samiya A Alkhairy[a)]

(Dated: 4 April 2022)

Abstract.
We estimate wavenumbers and impedances in the human cochlea using noninvasive measurements to better understand cochlear wave propagation and dynamics. The wavenumber and impedance encode information regarding how the cochlea works including wavelengths, gain and dissipation, and amplification and absorption. They provide a window into the cochlear amplifier and effective dynamic representations of the Organ of Corti.
Motivated by our interest in understanding how the cochlear works, we develop methods to estimate mechanistic variables (wavenumber and impedance), using noninvasive response characteristics (such as the quality factors of psychophysical tuning curves or group delays of stimulus frequency otoacoustic emissions) and an existing analytic shortwave model of the mammalian cochlea. The model is valid at low stimulus levels and was derived using a physical-phenomenological approach, and tested using a variety of datasets from multiple locations and species.
The model's small number of parameters and simple closed-form expressions enable us to develop methods for estimating mechanistic variables from noninvasive response characteristics. Developing these methods involves (1) deriving expressions for model constants, which parameterize the model expressions for wavenumber and impedance, in terms of characteristics of response variables - e.g. bandwidths and group delays of pressure across the Organ of Corti; followed by (2) deriving expresssions for the model constants in terms of noninvasive response characterisics. Using these derived expressions we provide general methods for estimating the wavenumber and impedance from noninvasive response characteristics.
We then use these expressions along with reported values for quality factors from psychophysical tuning curves measured in humans. Our resultant estimates for human wavenumbers and impedances show that the minimum wavelength (which occurs at the peak of the traveling wave) is smaller in base than the apex. The Organ of Corti is stiffness dominated rather than mass dominated, and there is negative damping prior to the peak followed by positive damping, indicating a region of amplification followed by a region of absorption. The stiffness, and positive and negative damping are greater in the base than the apex.
Future work involves studying the closed-form expressions for wavenumber and impedance for qualitative mechanistic interpretations across mammalian species as well as studying derived mechanisms such as power flux into the traveling wave and features of the cochlear amplifier. The methods introduced here for estimating mechanistic variables from characteristics of invasive or noninvasive responses enable us to derive such estimates across species and locations where the responses are describable by sharp filters. In addition to studying cochlear wave propagation and dynamics, the model properties (such as the ability to determine model constants based on desired response characteristics) make it particularly appropriate for auditory filter design.

MOTIVATION AND OBJECTIVES

The cochlea has fascinating signal processing features which motivates our interest in understanding how it works. For instance, stapes vibration results in traveling waves that propagate along the length of the cochlea and are subject to dispersion and amplification. Relative to the stapes stimulus, the response at each location along the length of the cochlea relative to the stapes stimulus peaks at a particular frequency. Such features are of particular interest to auditory physicists as well as those interested in bio-mimetic design, which motivates our goal of determining what underlies these interesting features.

For box representations of the cochlea, information regarding how it works is entirely encoded in two mechanistic variables: differential pressure wavenumber, and the organ of Corti effective impedance. The wavenumber and impedance are a window into properties such as effective stiffness, positive and negative damping or amplifier profile, incremental wavelengths, gain and decay, phase and group velocities, travel times, and dispersivity. These two mechanistic variables therefore provide us with a better understanding of what underlies cochlear features of high gain and place-based wavelet analyzers. Of further interest, are comparisons between how the cochlea works in different species or at different locations.

[a)]Corresponding author: samiya@mit.edu; samiya@alum.mit.edu

As the wavenumber and impedance cannot be directly observed, we must instead estimate them (a) via utilizing model assumptions, and (b) from observable variables. As for (a), we use a previously developed model of the healthy mammalian cochlea [1]. The model builds on classical box representations of the cochlea and is based on physics and phenomenon. It is valid near the peak region for low stimulus levels due to assumptions of linearity and short wave approximations. As for (b), we use noninvasive response characteristics, such as the quality factors of psychophysical tuning curves.

In this paper, we first briefly summarize the model expressions relevant for our mechanistic study. We then use our model towards our goal of understanding how the cochlea by achieving the following objectives:

1. Develop methods to determine estimates for the mechanistic variables - wavenumber and impedance, from *noninvasive* response characteristics - specifically quality factors of psychophysical tuning curves. The methods are schematized in figure 1.

2. Provide estimates for mechanistic variables for humans which have not been previously estimated due to the invasive surgical nature of experiments to obtain mechanical measurements from the cochlea

3. Compare the mechanistic variables between the human base and the human apex.

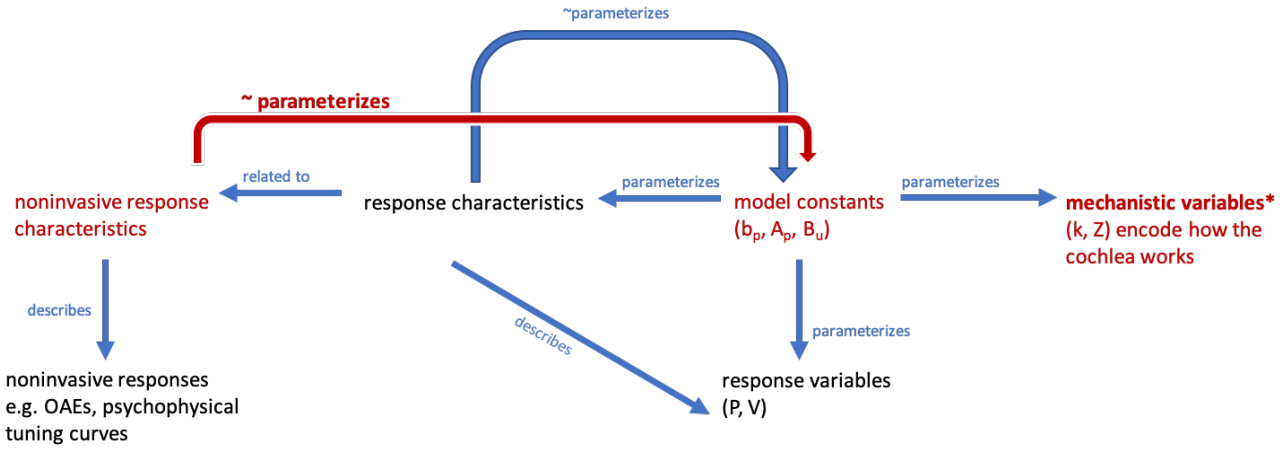


FIGURE 1. We may estimate the mechanistic variables which encode how the cochlea works in various species (particularly humans) and at various locations along the length of the cochlea (starred), by utilizing a model parameterized by three model constants, and inverting based on reported values of noninvasive response characteristics such as the quality factors of psychophysical tuning curves - i.e. following the path of the red text and arrows. In order to make use of this scheme to estimate k, Z for humans, we must first derive a parameterization of model constants in terms of noninvasive response characteristics. For species where invasive measurements are reported, we may alternatively parameterize the model constants in terms of (invasive) response characteristics such as the quality factors and group delays of experimental mechanical (or neural) measurements with both magnitude and phase information.

SUMMARY OF MODEL EXPRESSIONS

We use a previously developed interpretable analytic model of the mammalian cochlea, to contribute to our understanding of how the cochlea works [1]. The analytic model we previously developed was tested against existing data from various regions and species generated by experimental labs - e.g. [2]. It has interpretable closed-form expressions for all variables (pressure, velocity, wavenumber, and impedance), and is parameterized by a small number of model constants. These model constants can be computed based on values of response characteristics such as bandwidths and maximum group delays [1, 3, 4] as will be elaborated upon later in this paper. The aforementioned model properties of the model make it ideal for using it to understand how the cochlea works, as well as for utilizing characteristics of noninvasive responses (such as psychophysical tuning curves and otoacoustic emissions) to invert for mechanistic variables.

The frequency domain representation of the model is in terms of x, ω. For the wavenumber, we collapse the dependence onto a single independent variable which is a normalized frequency or transformed space, β,

$$\beta(x,\omega) \triangleq \frac{f}{\mathrm{CF}(x)} , \tag{1}$$

where $\mathrm{CF}(x) = \mathrm{CF}_{max} e^{-x/l}$ is the known characteristic frequency map of the species.

Here we summarize our model expressions for the mechanistic variables, wavenumber $k(\beta)$, and effective impedance, $Z(x,\omega)$ which encode how the cochlea works. Details of the model derivation based on physics and phenomenon, model testing, and model expressions for mechanistic variables (wavenumber and effective impedance) and response variables (pressure and velocity) were developed previously [1]. In developing the model, we have not assumed any resonances or any other forms of impedance. We utilized qualitative information from Weiner-Kernel based estimates of the wavenumber from chinchillas [5]. We have further imposed purely forward traveling waves, and that there is no gain in the differential pressure traveling wave beyond its peak. The model expression we constructed for the wavenumber, k, is

$$k(\beta)\frac{l}{\beta} = 2B_u \frac{i\beta + A_p}{(i\beta - p)(i\beta - \overline{p})} . \tag{2}$$

The expression for k is closed form and easily interpretable. Its dependence on x, ω is encoded in its dependence on β alone, which couples the dependence on x with the dependence on ω. Function-wise, this intriguingly couples the inhomogeneity (spatial variation of material properties) and dispersivity (separation of different frequency components of traveling waves as they propagate) in the cochlea. The assumption of scaling symmetry of k is valid locally as the parameter values vary relatively slowly along the length of the cochlea [1, 4].

The expression for k is a rational transfer function that has a pair of complex conjugate poles, $p = ib_p - A_p$ and $\overline{p} = -ib_p - A_p$, as well as a real zero at $i\beta = -A_p$. The three model constants, A_p, b_p, B_u take on positive real values. The constant, l is the space constant of the cochlear map, $\mathrm{CF}(x)$, that is empirically known for a variety of species, including humans.

The closed-form model expression for effective impedance, Z is derived from the above expression for k, as well as the relationship between k and Z for the short-wave approximation of box model representations of the cochlea,

$$Z(x,\omega)k(\beta(x,\omega)) = -2i\rho\omega . \tag{3}$$

Much can be inferred regarding how the cochlea works, *qualitatively*, based on the form of the above expressions for k and Z and their variation as a function of β or x, ω [4]. However, in this paper, we focus on developing methods for *quantitative* analysis and for comparative studies for species and location-specific interpretations.

METHODS FOR DERIVING MODEL ESTIMATES FROM MACROMECHANICAL RESPONSE CHARACTERISTICS

In this section, we provide expressions for the model constants in terms of (invasive) response characteristics, which is necessary towards estimating model constants (and hence mechanistic variables) from noninvasive response characteristics.

We previously provided expressions for our model macromechanical responses, pressure and velocity, as parameterized by our three model constants [1]. The responses (pressure or velocity), relative to the input at the stapes, can be characterized by a variety of frequency domain measures that describe their behavior - such as peak frequency, bandwidth, quality factor, group delay, and phase accumulation. We refer to these collectively as response characteristics.

It is desirable to parameterize the model constants in terms of response characteristics rather than determining them using responses themselves which suffer from incompleteness and noisiness issues. Furthermore, response characteristics are more readily available from the literature. Note that, to, first approximation, the response characteristics of pressure and velocity are similar for the parameter values of interest, and hence we do not distinguish between them.

In order to obtain values of model constants from values of reported response characteristics, we first derive expressions for response characteristics in terms of our three model constants [3]. We deal with the following set of response

characteristics: peak normalized frequency, β_{peak}, group delay at the peak in the normalized frequency domain, N, and n dB bandwidth in the normalized frequency domain, $\mathrm{BW}_{\mathrm{n\,dB}}$. Note that $\mathrm{BW}_{\mathrm{n\,dB}}$ can easily be converted to the f domain bandwidth by multiplication by $\mathrm{CF}(x)$, and N can be converted to the f domain by division by $\mathrm{CF}(x)$.

We arrive at the following simple closed-form expressions for response characteristics in terms of model constants which we derived for sharp-filters as is descriptive of the human base and human apex as well as the chinchilla base.

$$\beta_{peak} = b_p \tag{4}$$

$$N = \frac{B_u}{2\pi A_p} \tag{5}$$

$$\mathrm{BW}_{\mathrm{n\,dB}} = 2A_p\sqrt{e^{\frac{\log(10)n}{10B_u}} - 1} \tag{6}$$

We then invert these expressions to arrive at formulas (below) for the model constants in terms of the response characteristics. For animals in which invasive mechanical or proxy neural measurements can be obtained, we are provided with values for response characteristics. From these, using the method below, can determine the values for the model constants b_p, A_p, B_u, and consequently the model wavenumber and impedance in order to study how the cochlea works in those animals.

1. As $f_{peak} = \mathrm{CF}(x)$, we generally set $b_p = 1$
2. Solve for B_u from N x $\mathrm{BW}_{\mathrm{n\,dB}} = \frac{B_u}{\pi}\sqrt{e^{\frac{\log(10)n}{10B_u}} - 1}$
3. Compute A_p from B_u determined above and either (a) the equation for N (as $A_p = \frac{B_u}{2\pi N}$) or (b) the equation for $\mathrm{BW}_{\mathrm{n\,dB}}$ (as $A_p = \frac{\mathrm{BW}_{\mathrm{n\,dB}}}{2\sqrt{e^{\frac{\log(10)n}{10B_u}} - 1}}$)
4. Plug the values for model constants obtained above into the model expressions for wavenumber and impedance to study how the cochlea works

METHODS FOR DERIVING MODEL ESTIMATES FROM NONINVASIVE RESPONSE CHARACTERISTICS

In this section, we develop methods to determine A_p, B_u (and hence k, Z which encode how the cochlea works) from *noninvasive* response characteristics. The method for determining the model mechanistic variables described above requires response characteristics from invasive measurements, and can readily be applied to study mechanisms in animals for which we have invasive response characteristics. However, this is not feasible in humans with today's technology. Consequently, we must develop a way to determine the model constants and mechanistic variables from characteristics of noninvasive responses. Specifically, we may use psychophysical tuning curves quality factors and/or stimulus frequency otoacoustic emission (SFOAE) group delays.

In order to achieve our goal and use these noninvasive response characteristics to determine the model constant values necessary to determine k, Z, we must first determine the relationship between the noninvasive response characteristics and the (invasive) response characteristics described above. This allows us to then use the methods outlined in the previous section towards our goal. Here we describe one such mapping which utilizes quality factors of psychophysical tuning curves and an assumed species invariant ratio.

Reported values of quality factors, Q_{erb}, are based on equivalent rectangular bandwidths. In contrast, our simple closed-form expression for bandwidth, and hence associated quality factor, Q_n, is for n dB bandwidths. To convert between Q_{erb} and Q_n, we use an empirical relationship between these two measures for $n = 10$, obtained from ANF tuning curves and found to be largely species and CF invariant [6]: $Q_{erb} = \alpha Q_{10}$, where α is in the range $1.7 - 1.8$. Note that Q_{10} is the same whether defined in β or f domains, and that $\frac{1}{Q_{10}} = BW_{10\,\mathrm{dB}}$ (bandwidth defined in the β domain) because $\beta_{peak} = 1$. This allows us to replace the second step in the previous section (determining B_u) with,

$$\frac{N}{Q_{erb}} = \frac{B_u}{\alpha\pi}\sqrt{e^{\frac{\log(10)n}{10B_u}} - 1} \tag{7}$$

The ratio $g = \frac{Q_{erb}}{N}$ was empirically found to be a constant in chinchilla ($g \approx 1.25$) and we assume that this ratio is species-independent for most of the length of the cochlea - note that the tuning ratio $r = \frac{Q_{erb}}{N_{sfoae}}$ has been shown to be species invariant [7]. Using these values for g, α and with $n = 10$, we solve for B_u. This gives $B_u \approx 7$ for various species across the length of the cochlea where the sharp-filter approximation holds.

Finally, we determine $A_p(\text{CF(x)})$ from reported values of quality factors of psychophysical tuning curves determined using a non-simultaneous masking paradigm with low stimulus levels [8]. Specifically, following step 3-(b) of the previous section, we use reported $Q_{psych-erb}(\text{CF(x)})$ and the B_u determined above. We assume that the psychophysical tuning curves quality factors approximate quality factors of macromechanical responses, $Q_{psych-erb} \approx Q_{erb}$, and use the aforementioned factor α to substitute $\text{BW}_{10\,\text{dB}}$ of step 3-(b) with $\text{BW}_{10\,\text{dB}} = \frac{\alpha}{Q_{erb}}$.

$$A_p(\text{CF}) = \frac{\alpha}{2Q_{psych-erb}(\text{CF})\sqrt{e^{\frac{log(10)}{B_u}} - 1}} \tag{8}$$

Another way to estimate $A_p(\text{CF})$ from quality factors of psychophysical tuning curves, is by using step 3-(a), and substituting $N = \frac{Q_{erb}}{g}$,

$$A_p(\text{CF}) = \frac{gB_u}{2\pi Q_{psych-erb}(\text{CF})} \tag{9}$$

Using either expression above, we use reported quality factors of psychophysical tuning curves from humans to determine A_p. Both equations yield very similar estimates for $A_p(\text{CF})$ from $Q_{psych-erb}(\text{CF})$. Note that we may alternatively have estimated $A_p(\text{CF(x)})$ using another noninvasive response characteristic which can be measured in humans. Specifically, we may determine A_p using step 3-(a) with group delays of stimulus frequency otoacoustic emissions if we make assumptions regarding the relationship between N and N_{sfoae}. This relationship has been the subject of previous studies [9].

For humans the reported psychophysical tuning curve quality factors are as follows: at 10 kHz, $Q_{psych-erb} \approx 25.34$ and at 1 kHz, $Q_{psych-erb} \approx 12.7$ [8]. Using the equations above, these result in $A_p \approx 0.055$ and $A_p \approx 0.11$ respectively. With this we have achieved our first objective of developing methods for estimating model constant values (and thereby mechanistic variables) from noninvasive response characteristics.

ESTIMATES FOR HUMAN WAVENUMBER AND IMPEDANCE

With the method introduced above for determining model constant values from noninvasive response characteristics, we can now (2) determine the wavenumber and impedance from the human noninvasive response characteristics, and also (3) differentiate between their profiles in the human base and apex (which provides us with information regarding how these two regions of the cochlea function differently). We discuss both these aspects here by demonstrating the wavenumber and impedance of two points along the length of the cochlea - one from the apex with a low CF, 1 kHz, and another from the base with a higher CF, 10 kHz. From the two sets of model constants estimated in the previous section, we can determine the mechanistic variables - the wavenumber and impedance, and their differences between the base and the apex. In figure 2, we have plotted the model wavenumber and impedance for the two points in the human apex and base.

The real part of the wavenumber has a peak, the value of which determines the minimum incremental wavelength of the pressure traveling wave as it propagates along the length of the cochlea. The peak is greater in the base than the apex, indicating that the incremental wavelength is much smaller in the base than in the apex near the peak of the wave (though this is not the case outside the peak region). The imaginary part of the wavenumber has a peak and trough that are larger in the base than the apex which indicates that the gain and dissipation accumulate more quickly about the peak of the differential pressure traveling wave in the base than in the apex. This can alternatively be observed in the corresponding pressure response.

The imaginary part of the impedance at a particular location, x, is negative (indicating an effective stiffness rather than mass). The real part of the impedance is negative prior to the peak frequency then positive indicative of negative effective damping or amplification followed by positive effective damping or absorption. Previous estimates of the impedance at a particular location [10, 11, 12] suggest that $\text{Im}\{Z\} < 0$ prior to the peak in velocity and at least shortly before that, which qualitatively agrees with our model. Beyond quantitative differences, there are also qualitative

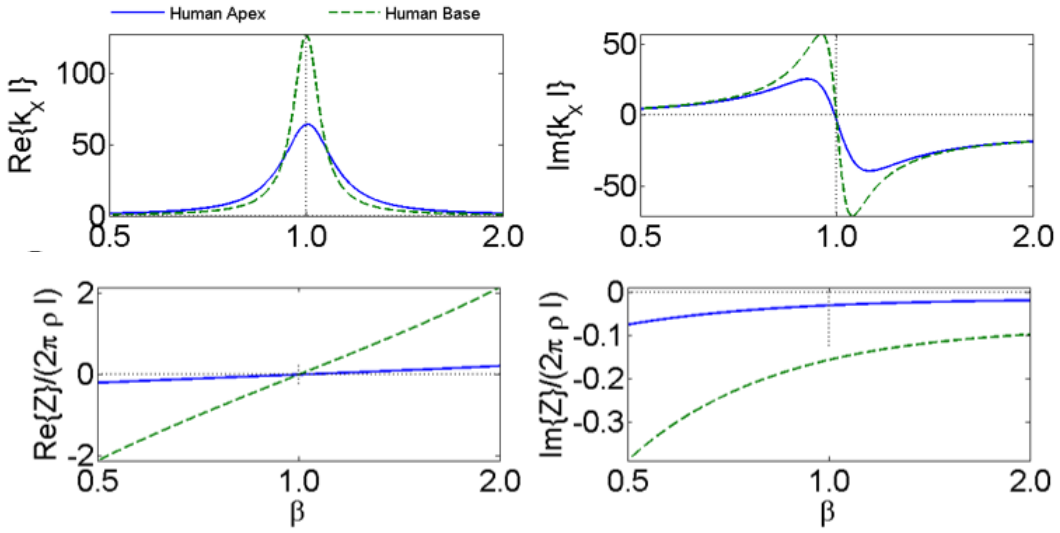


FIGURE 2. Estimated human wavenumber and impedance: The top panel shows the real and imaginary parts of the model generated human wavenumber, and the bottom panel shows the real and imaginary parts of the impedance. The blue solid lines are for a point in the apex (CF = 1 kHz, where $Q_{psych-erb} \approx 12.7$) which has model constants $A_p = 0.11, B_u = 7$, and the dashed green lines are for a point in the base (CF = 10 kHz, where $Q_{psych-erb} \approx 25.34$) which has model constants $A_p = 0.055, B_u = 7$. The values for A_p, B_u are estimated from noninvasive response characteristics in the previous section.

differences in terms of the dependence of their impedances on frequency and/or location. The three estimates have a $\mathrm{Re}\{Z\}$ that is negative then positive and crosses zero near the peak, which is qualitatively consistent with our model - though the frequency of the zero crossing, and the behavior away from it varies across these estimates.

Our effective stiffness is greater in the base than in the apex. If we interpret the effective impedance, Z, as a *local* impedance due to properties at a single point, we may infer that the stiffness of the Organ of Corti is greater in the base than in the apex. The effective positive and negative damping are greater in the base than the apex. It is of interest to determine how this quantitatively translates to potential amplification and absorption differences, for which must determine the power flux into the traveling wave as it propagates along the length of the cochlea.

CONTRIBUTIONS AND FUTURE WORK

The mechanistic variable estimation methods developed in this paper enable us to study how the cochlea operates in various species, regions, or pathologies that have different response characteristics but the same underlying model structure, as well as determine differences between these various cases. The model utilized is for healthy mammalian cochleas at low levels, and the estimation methods, are best suited for cases where the tuning is relatively sharp such as is the case in the human base and apex and chinchilla base.

In this paper, we have used the estimation methods to determine human wavenumber and impedance, for which there have been no previous estimates. We also discussed differences in these mechanistic variables between the base and the apex. From these estimates, it is possible to derive other mechanistic information, such as phase velocity, dispersivity, power flux, effective stiffness, as well as their differences between the base and the apex to develop a deeper understanding of function.

We may also extend the methods to derive mechanistic information from another noninvasive response characteristic - specifically stimulus frequency emission group delays. In addition to comparative analysis built on estimating constants from characteristics, we can study the model *expressions* for wavenumber and impedance more deeply (without specified model constant values) as they provides us with a general understanding of how the cochlea functions across mammalian species. In addition to their use for studying auditory physics, the model and methods are also useful for developing easily tunable auditory filters for engineering applications and perceptual studies.

REFERENCES

1. S. A. Alkhairy and C. A. Shera, "An analytic physically motivated model of the mammalian cochlea," The Journal of the Acoustical Society of America **145**, 45–60 (2019).
2. A. N. Temchin, A. Recio-Spinoso, P. Van Dijk, and M. A. Ruggero, "Wiener kernels of chinchilla auditory-nerve fibers: verification using responses to tones, clicks, and noise and comparison with basilar-membrane vibrations," Journal of neurophysiology **93**, 3635–3648 (2005).
3. S. A. Alkhairy, "Easily tunable auditory filters with flexible implementations," (in prep.).
4. S. A. Alkhairy, *An analytic model of the cochlea and functional interpretations*, Ph.D. thesis, Massachusetts Institute of Technology (2017).
5. C. A. Shera, "Laser amplification with a twist: traveling-wave propagation and gain functions from throughout the cochlea," The Journal of the Acoustical Society of America **122**, 2738–2758 (2007).
6. C. A. Shera and J. J. Guinan Jr, "Stimulus-frequency-emission group delay: A test of coherent reflection filtering and a window on cochlear tuning," The Journal of the Acoustical Society of America **113**, 2762–2772 (2003).
7. C. A. Shera, J. J. Guinan, and A. J. Oxenham, "Otoacoustic estimation of cochlear tuning: validation in the chinchilla," Journal of the Association for Research in Otolaryngology **11**, 343–365 (2010).
8. C. A. Shera, J. J. Guinan, and A. J. Oxenham, "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements," Proceedings of the National Academy of Sciences **99**, 3318–3323 (2002).
9. C. A. Shera, A. Tubis, and C. L. Talmadge, "Testing coherent reflection in chinchilla: Auditory-nerve responses predict stimulus-frequency emissions," The Journal of the Acoustical Society of America **124**, 381–395 (2008).
10. W. Dong and E. S. Olson, "Detection of cochlear amplification and its activation," Biophysical journal **105**, 1067–1078 (2013).
11. E. De Boer and A. L. Nuttall, "The "inverse problem" solved for a three-dimensional model of the cochlea. iii. brushing-up the solution method," The Journal of the Acoustical Society of America **105**, 3410–3420 (1999).
12. G. Zweig, "Linear cochlear mechanics," The Journal of the Acoustical Society of America **138**, 1102–1121 (2015).

Comparison of Synchronized Spontaneous Otoacoustic Emission Dynamics in Human Data and Cochlear Mechanics Simulation: Effects of Roughness

Namita Sengar[1, a)], Hao-Ping Lin[1, b)] and Yi-Wen Liu[1,c)]

[1]*Department of Electrical Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan*

[a)] s106061859@m106.nthu.edu.tw
[b)] jeremy.0515@gapp.nthu.edu.tw
[c)] ywliu@ee.nthu.edu.tw

Abstract. The synchronized spontaneous (SS) otoacoustic emission (OAE) refers to the long-lasting oscillation evoked by clicks after transient-evoked OAEs (TEOAEs) attenuate. This research aims to study the SSOAE spectro-temporal dynamics by comparing human data and simulation. The simulation of SSOAE is based on a computer model of cochlear mechanics. In the model, the roughness of the basilar membrane (BM) can be adjusted and its effects on SSOAE patterns are observed. It is noted that the change in roughness level (5%, 10%, 15%, and 20%) affects the likelihood distribution between single-frequency, two-frequency, and a band-of–frequencies components. There is a general trend that band-of-frequencies occur more as the roughness level increases. The simulation result shows that, unlike the roughness level of 5% and 10%, roughness levels of 15% and 20% can generate SOAE oscillation below 2 kHz. Signal patterns of simulation are similar to what is observed in human SSOAE data, and thus support the standing wave theory of SOAE.

INTRODUCTION

OAEs can be regarded as a tool to assess power-amplifying function in the cochlea [1]. Among different types of OAEs, the synchronized spontaneous (SS) otoacoustic emission (OAE) refers to the long-lasting oscillation evoked by a series of clicks. As the name suggests, SSOAEs are closely related to spontaneous OAEs (SOAEs) but can be synchronously triggered by external stimuli. SSOAEs are present at isolated frequencies and typically decay slowly with time on the order of several hundred milliseconds [2]. There have been past studies on the theory of SOAE and stimulus-frequency otoacoustic emissions (SFOAE) generation using computer simulation. One of the studies tested the prediction of the multiple-reflection theory of SOAE generation using state-space equations for humans and concluded that instabilities across a wide bandwidth of frequencies arise when a smooth spatial variation of BM impedance is disturbed [3]. In a recent work [4], cochlear roughness was added to test the ability of their developed model to simulate SFOAE, and the link between fine structure and standing wave resonance of SFOAEs were further examined. The model was poised close to but below the limit of linear stability, so that it did not emit SOAEs. Then, the model was compared with and without roughness. Given that the afore-mentioned studies were focused on the effect of roughness on SFOAE generation, in this research, we aim to explore the roughness effect on SSOAE generation using a computer model of cochlear mechanics [5] and compare the simulation results with human data.

In particular, we study the SSOAEs spectro-temporal dynamics. In the model, the roughness of the basilar membrane (BM) can be adjusted and its effects on SSOAE patterns are observed. We shall describe how the roughness level affects the number of SSOAE spectral components and their temporal tendencies (decaying or sustaining) in simulation. The paper is organized as follows. In the method section, SSOAE human data collection is described and the related signal processing procedures are introduced. The result section consists of the comparison of simulation and human data in terms of SSOAE components and signal tendency. Discussions are given in the last section.

METHOD

TEOAEs were recorded by an ER-10X microphone via an RME Fireface UFX II soundcard in a sound-proof room at National Tsing Hua University, Taiwan. Nine subjects between the age of 21 to 25 years were recruited. All ears were measured across different days for a total of 3 times. In total, 18 ears were measured; 16 measurements were successful while 2 were omitted due to artifact rejection. Eight subjects had SSOAEs in the right ear but only five subjects had SSOAEs in the left ear. The method used for measurement is known as the nonlinear protocol [1]. The nonlinear protocol helps to eliminate the initial linear response, which did not pertain to cochlear conditions. Three clicks were included in a unit of stimulus, where the first two clicks are equal to each other in amplitude and the third click is opposite in magnitude with an amplitude two times larger than the first 2 clicks. The interval between the clicks is approximately 99.8 ms (4400 sample points in 44.1 kHz sampling rate), which is long enough for SSOAEs to be observed in recordings [6].

After recording, a zero-phase band-pass filter was applied to preserve the range from 0.8 kHz to 10 kHz on the recorded signal. Then, the signal was divided into non-overlapping units of length 3T, where T represents the interval between clicks. Recorded signals inevitably contained environmental and human-generated noise (such as the sound of swallowing and wire friction with clothes). For accurate analysis, signal quality has to be maintained. Therefore, a unit of a signal consisting of 3 clicks would be rejected if the amplitude was higher than 50 dB SPL at any instant.

To search for SSOAEs visually, the response later than 20 ms after the click was analyzed in the frequency domain. After identifying spectral peaks, single-frequency (1F) components, two-frequencies (2F) components, and band-of-frequencies (BoF) components were extracted by custom-adjusted band-pass filters (BPF). Figure 1 shows typical examples of the 2F and BoF SSOAE components. To select BoF components, the search range was controlled to be within half of an octave [7].

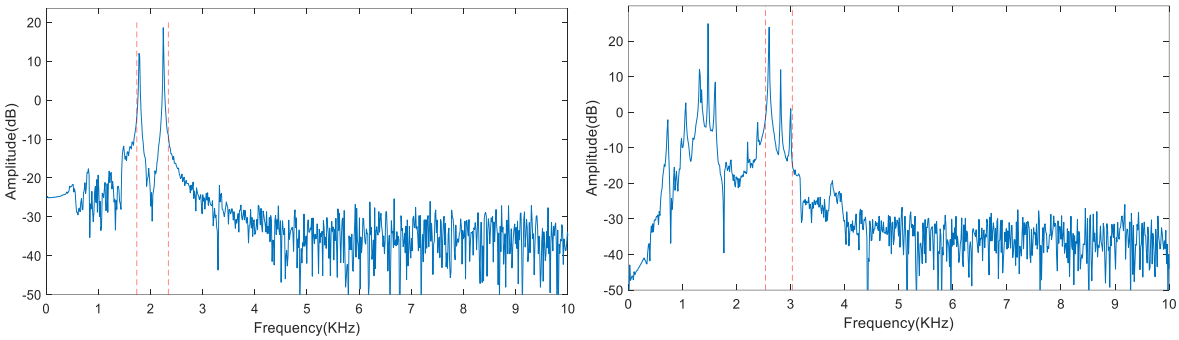


FIGURE 1. Example of a 2F component (Left) and a BoF component (Right) in the spectrum of SSOAE.

The simulation model used in this work has been shown to produce many features typical of the cochlea, such as the compressive response growth function on BM and reticular-lamina (RL) displacement, the DPOAE level dependency on primary tone levels and frequencies [6], and the fluctuation of TEOAE instantaneous frequency as a function of time [8]. To analyze the roughness effect on RL displacement in this study, the roughness level is changed to 5%, 10%, 15%, and 20%. The roughness is spatially smoothed by a symmetric low pass kernel of length 0.6 mm [8, 9]. Then, RL displacement is plotted as a function of time at 12 equally spaced locations, each 0.2 cm apart. Subsequently, these 12 plots are categorized into “1F”, “2F”, or “BoF”. The signal tendency is also noted, in terms of whether the signal is growing, sustaining, or decaying with time. Thus, signals from human data and simulation can be compared with each other based on the number of frequency components, signal tendency, and signal pattern over time.

RESULT

In the simulation model, it is expected that different oscillatory components will be observable at different BM locations due to roughness. The characteristic frequency at these 12 locations is grouped into three frequency bands: namely the high-frequency band (13.7, 10.4, 7.9, 6.0 kHz), the mid-frequency band (4.5, 3.4, 2.5, 1.9 kHz), and the low-frequency band (1.4, 1.0, 0.76, and 0.56 kHz). Figure 2 shows an example of the simulation for a roughness level of 10%, where in the high frequency and middle-frequency band, the SOAE oscillations are long-lasting over 75ms, but in the low-frequency band, no SSOAE components are observed.

Figure 3 shows an example of 2F and BoF components from one human ear. For the 2F component, note that the signal pattern and amplitude both seem to be stationary in time. In contrast, the BoF component exhibits a signal pattern that repeats itself until 50 ms; subsequently, the amplitude is decaying with time. After analysis, it is observed that the same kind of signal pattern is noticed in simulation data and human data.

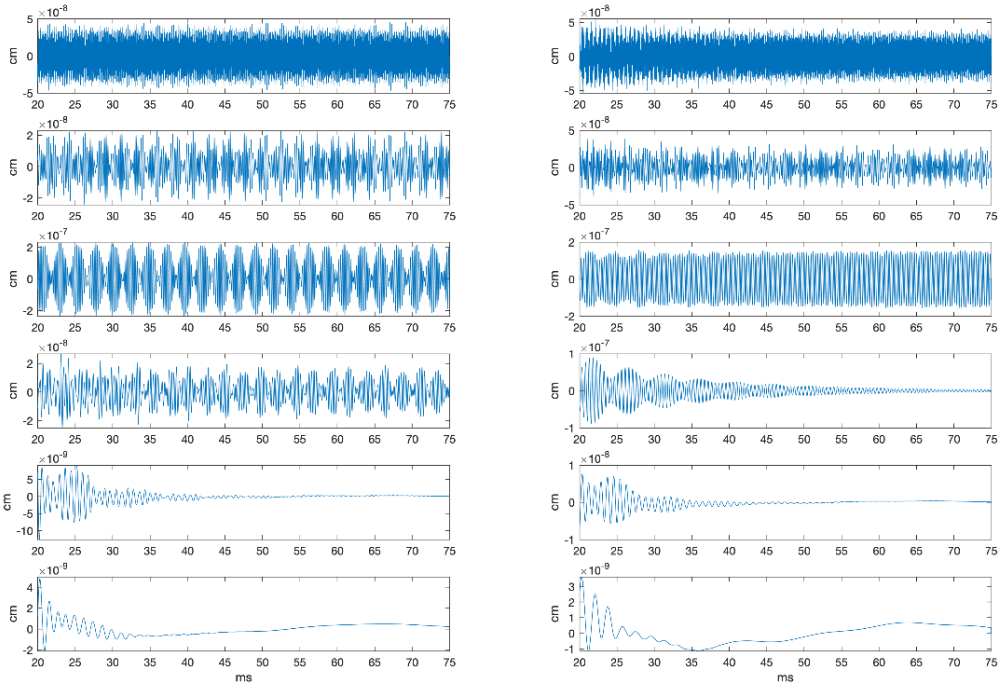


FIGURE 2. In this example (simulation), a roughness of 10% is applied to the basilar membrane mass density, and then the roughness is spatially smoothed by a symmetric low pass kernel of length 0.6 mm. Reticular lamina displacement is plotted as a function of time at 12 equally spaced locations, 0.2cm apart. (Top left: the most basal (13.7 kHz), Top right: the second most basal (10.4 kHz), and Bottom Right: the most apical (0.56 kHz) location.)

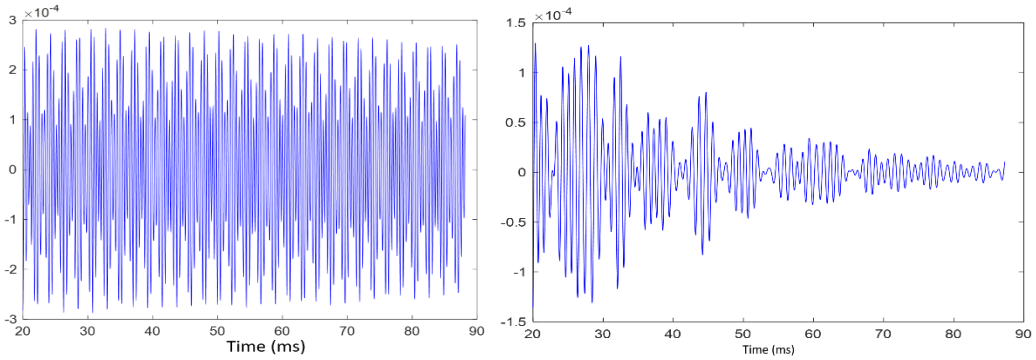


FIGURE 3. Typical examples of 2F (left) and BoF (right) SSOAE components from human ears, viewed in the time domain.

To determine how the roughness level affects the pattern of SSOAE, we count the number of SSOAE components and observe the signal tendency in the simulation over 75 ms. Table 1 shows the number of SSOAE components observed for different levels of roughness. Note that the change in roughness level (5%, 10%, 15%, and 20%) affects the likelihood to observe 1F, 2F, and BoF components. At high roughness levels (15% and 20%), we observe fewer 1F components and more 2F components or BoF components.

Table 2 shows the number of SSOAE components for human data. It is observed that simulation data for roughness levels of 5% and 10% are more similar to human data in terms of a ratio between the number of 1F components, 2F

components, and BoF components, while at 15% or 20% the simulated RL displacement waveforms are predominantly categorized as BoF.

TABLE 1. The number of spontaneous oscillation components with different roughness levels (%). Simulation is run for 30 pseudo-random instances of roughness, and observations of RL displacement were conducted at 12 different locations, resulting in 360 observations at each roughness level. 1F: one-frequency; 2F: two-frequencies; BoF: band-of-frequencies.

Roughness	No SSOAE	1F	2F	BoF
5%	215	60	83	2
10%	112	10	123	115
15%	64	8	87	201
20%	54	2	104	200

TABLE 2. Number of SSOAE components for human data

1F	2F	BoF
16	20	26

Table 3 shows SSOAE signal tendency with time for the simulation data. For the roughness levels of 15% and 20%, the number of sustained and decayed components is large in comparison to the roughness levels of 5% and 10%.

TABLE 3. SSOAE signal tendency with time

Roughness	Grow	Sustain	Decay
5%	31	39	75
10%	12	165	71
15%	40	152	104
20%	42	162	102

TABLE 4. SSOAE signal tendency with time for human data

Grow	Sustain	Decay
1	27	34

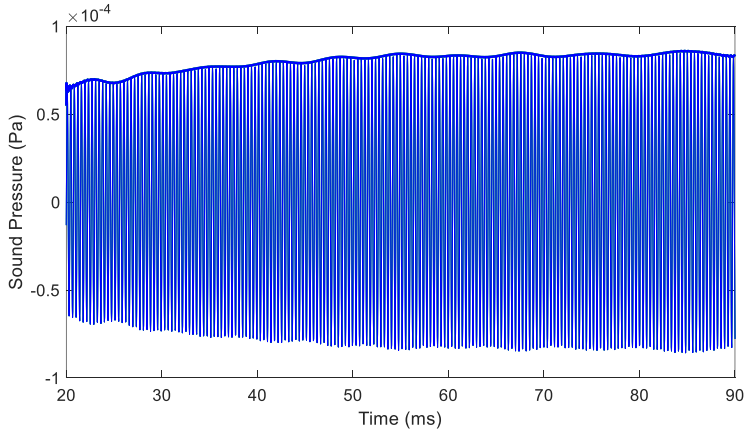


FIGURE 4. In this example (a unique case of human data), SSOAE component grows with time

Table 4 shows SSOAE signal tendency with time for human data, where similarly to simulation the sustained and decayed components are higher in number. Figure 4 shows a rare case of human data, where the signal grew with time. It is observed in this example that signal was growing till 55 ms and after that it sustained with time.

DISCUSSION

Nonlinearity and distributed roughness play significant roles in the function of the cochlea. A past study qualitatively discussed the combined effects of both nonlinearity and distributed roughness in the context of TEOAEs [10], which led to the hypothesis that roughness increases in the aging cochlea due to natural tissue degradation. More recently, the levels of DPOAE and SFOAE were found to decrease with age, but the rate of decline in DPOAE levels was greater than SFOAE, which implies that SFOAE is relatively preserved with advancing age [11].

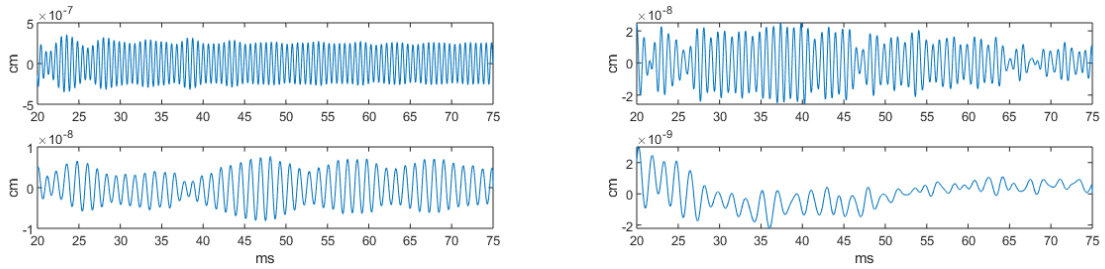


FIGURE 5. In this example (simulation), a roughness of 15% is applied to the basilar membrane mass density

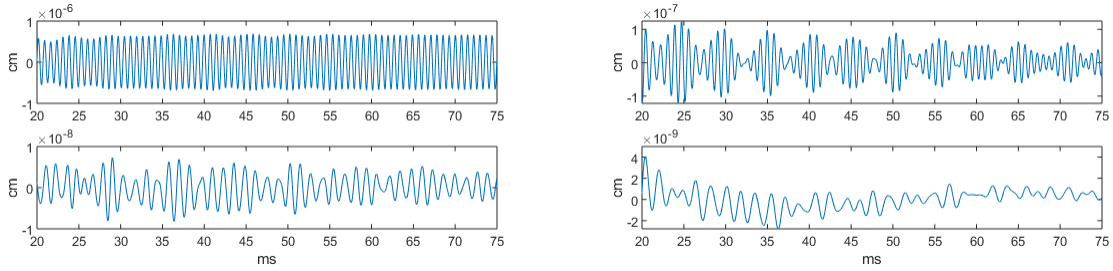


FIGURE 6. In this example (simulation), a roughness of 20% is applied to the basilar membrane mass density

We are aware that the present simulation model could not produce clear spontaneous long-lasting oscillations below 1.5 kHz for 5% roughness level, and few SSOAE components are observed in the low-frequency band for 10% roughness level. As we proceed with 15% and 20% roughness levels, clear long-lasting oscillation could be observed below 1.5 kHz (corresponding to the 4 most apical locations in Fig. 2) as shown in Fig. 5 and Fig. 6.

Figure 7 shows a comparison of the number of SSOAE components observed in simulation vs. human data at different frequency bands. Note that for a roughness level of 5%, no SSOAE component is found in the low-frequency band in any of the simulations. In contrast, a large number of SSOAE components are found for the high-frequency band and mid-frequency band at all roughness levels, with a significant jump when the roughness level increases from 5% to 10 %. For the human data, a large number of SSOAE components are found in the mid-frequency band, and also a significant number of SSOAE components are found in the low-frequency band. However, few SSOAE components were found in the high-frequency band i.e above 6 kHz. In literature also, for human data analysis, prominent SSOAE components are found somewhere in between low-frequency and mid-frequency band, i.e around 1 kHz to 4 kHz [7] [12].

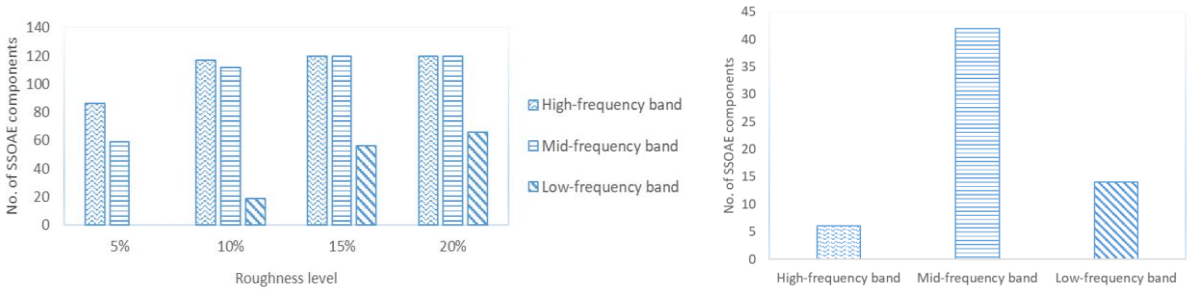


FIGURE 7. Number of SSOAE components (simulation) for different roughness levels and their frequency dependency (Left). Number of SSOAE components (human data) and their frequency dependency (Right).

These results indicate that change in roughness level affects spontaneous oscillations in simulation. To summarize, the present findings agree with the theory that increased roughness would boost the backscattering of traveling waves [11]. However, there is a discrepancy between our data acquisition and simulation protocols. Note that the human data in this research were acquired by the non-linear protocol mentioned in the method section, but the simulation was conducted under the linear protocol; the click level was not adjusted to see how the simulation result varies. In the future, we would suggest to investigate further about the nonlinearity effect [13] on SSOAE generation through simulation, and systematically study the change in magnitude and phase with roughness level.

ACKNOWLEDGMENTS

This research is supported by the Ministry of Science and Technology of Taiwan (Grant No. 110-2221-E-007-047-MY2).

REFERENCES

1. Kemp, D.T., "Otoacoustic emissions, travelling waves and cochlear mechanisms," *Hearing Research*, Volume 22, Issues 1–3, 95-104, 1986, https://doi.org/10.1016/0378-5955(86)90087-0.
2. Sisto, R., A. Moleti, and M. Lucertini. "Spontaneous otoacoustic emissions and relaxation dynamics of long decay time OAEs in audiometrically normal and impaired subjects," *J. Acoust. Soc. Am.* Volume 109, Issue 2, 638-647, 2001, https://doi.org/10.1121/1.1336502.
3. Ku, E. M., Elliott, S. J., and Lineton, B., "Statistics of instabilities in a state space model of the human cochlea," *J. Acoust. Soc. Am.*, Volume 124, Issue 2, 1068-1079, 2008.
4. Wen, H., and Meaud, J., "Link between stimulus otoacoustic emissions fine structure peaks and standing wave resonances in a cochlear model," *J. Acoust. Soc. Am.* Volume 151. Issue 3, 1875-1894, 2022.
5. Liu, Y.-W., and Stephen T. Neely. "Distortion product emissions from a cochlear model with nonlinear mechanoelectrical transduction in outer hair cells," *J. Acoust. Soc. Am.*, Volume 127, Issue 4, 2420-2432, 2010, https://doi.org/10.1121/1.3337233.
6. Jedrzejczak, W. W., Blinowska, K. J., Kochanek, K., and Skarzynski, H., "Synchronized spontaneous otoacoustic emissions analyzed in a time-frequency domain," *J. Acoust. Soc. Am.* Volume 124, Issue 6, 3720–3729, 2008, https://doi.org/10.1121/1.2999556.
7. Lewis, J. D., Amy Mashburn, and Donguk Lee., "Jittering stimulus onset attenuates short-latency, synchronized-spontaneous otoacoustic emission energy," *J. Acoust. Soc. Am.* Volume 147, Issue 3, 1504-1515, 2020, https://doi.org/10.1121/10.0000848.
8. Wu, H.-T., and Liu, Y.-W., "Analyzing transient-evoked otoacoustic emissions by concentration of frequency and time," *J. Acoust. Soc. Am.* Volume 144, Issue 1, 448-466, 2018.
9. Choi, Y.-S., Lee, S.-Y., Parham, K., Neely, S. T., and Kim, D. O., "Stimulus-frequency otoacoustic emission: Measurements in humans and simulations with an active cochlear model," *J. Acoust. Soc. Am.* Volume 123, Issue 5, 2651–2669, 2008.
10. Yates, G. K., and Withnell, R. H. ''The role of intermodulation distortion in transient-evoked otoacoustic emissions,'' *Hearing Research*, Volume 136, Issues 1–2, 1999, 49-64, https://doi.org/10.1016/S0378-5955(99)00108-2.

11. Abdala, C., Ortmann A.J., and Shera C.A., "Reflection-and distortion-source otoacoustic emissions: Evidence for increased irregularity in the human cochlea during aging," *Journal of the Association for Research in Otolaryngology*, Volume 19, Issue 5, 493-510, 2018, https://doi.org/10.1007/s10162-018-0680-x.
12. Mertes, I. B., "Medial olivocochlear reflex effects on synchronized spontaneous otoacoustic emissions," *J. Acoust. Soc. Am.*, volume147, Issue 3, EL235-EL240, 2020
13. Talmadge, C. L., Tubis, A., Long, G. R., and Tong, C., "Modeling the combined effects of basilar membrane nonlinearity and roughness on stimulus frequency otoacoustic emission fine structure," *J. Acoust. Soc. Am.* Volume 108, Issue 6, 2911-2932, 2000, https://doi.org/10.1121/1.1321012

Understanding Bone Conduction In The Human Cochlea With Optical Coherence Tomography Vibrometry

Lore Kerkhofs[1,2, a)], Anastasiya Starovoyt[1,2], Tristan Putzeys[1,2,3] and Nicolas Verhaert[1,2 4]

[1]*Research Group Experimental Oto-Rhino-Laryngology, Department of Neurosciences, KU Leuven, University of Leuven, Leuven, Belgium*
[2]*Leuven Brain Institute, Department of Neurosciences, KU Leuven, 3000 Leuven, Belgium*
[3]*Laboratory for Soft Matter and Biophysics, Department of Physics and Astronomy, KU Leuven, University of Leuven, Leuven, Belgium*
[4]*Department of Otorhinolaryngology, Head and Neck Surgery, UZ Leuven, University Hospitals of Leuven, Leuven, Belgium*

[a)]Corresponding author: lore.kerkhofs@kuleuven.be

Abstract. To date, bone conduction (BC) has been extensively investigated using laser doppler vibrometry (LDV), intracochlear pressure measurements (ICP) and lumped element models (LEM). However, these techniques can disrupt the cochlear homeostasis, or are only applicable extracochlear. In this study, we used Optical Coherence Tomography (OCT) to overcome these limitations by in-depth, longitudinal, and contactless imaging of the human cochlear hook region through the round window membrane. Hereby, we were able to visualize the vestibular and tympanic plate of the osseous spiral lamina, basilar membrane, spiral limbus, Reissner's membrane epithelium, stapes anterior crus and head, scala tympani and scala vestibuli. Using OCT vibrometry, BC mechanisms are investigated in combination with subjective loudness balancing tests (LB), equalizing a 60 dB HL AC sound with an adaptable BC sound. This way, the displacement differences between AC and BC sound, assuming the same psychoacoustic loudness level, were mapped on intracochlear structures. Different force levels where needed to equalize the AC and BC sound in the loudness balancing test, and different velocity amplitudes and phases compared to the input signal were found on the basilar membrane and osseous spiral lamina using OCT vibrometry. Additionally, the stapes, RWM and promontory moved substantially different both in the human temporal bones and human cadaveric head. To the best of the authors' knowledge, this is the first study, investigating the differences between AC and BC in the human cochlea by means of OCT vibrometry. The obtained results were validated against the existing literature on AC and BC hearing, demonstrating that OCT vibrometry is a reliable method for studying the mechanics of the human cochlea.

INTRODUCTION

Bone conducted (BC) sound is the perception of sound travelling through the skull bones and soft tissue towards the inner ear, exciting the basilar membrane (BM) and hair cells causing a hearing sensation [1]. Contrary to air conduction (AC), which only involves one pathway, BC vibrates the BM through multiple frequency-dependent pathways (Fig. 1) [2]. Nobel prize winner von Békésy was the first to question the mechanisms of BC, revealing a similar stimulation of the BM through AC and BC sound [3]. Ever since, research to unravel the BC mechanisms extensively increased, using e.g., laser doppler vibrometry (LDV), intracochlear pressure measurements (ICP) or lumped element models (LEM). Consequently, evidence regarding similarities and differences of BM motion through AC and BC is rising. Additionally, the motion of the bony osseous spiral lamina and non-bony cochlear partition bridge are of high interest to date to predict their influence on the

BM motion [4–7]. However, when using LDV, retroreflective material needs to be placed on the structure of interest to measure the velocity and phase, and when performing ICP, sensors must be placed in the scalae via a cochleostomy, resulting in invasive experimental approaches and interruption of the cochlear homeostasis [17–19]. Additionally, Lumped element models used to understand BC mechanics, are mainly based on the externally visible structures of the cochlea, such as the middle ear ossicles or the round window membrane (RWM) [20–22].

Understanding BC mechanisms is important in view of clinical practice, e.g., bone anchored hearing aids are used to bypass the malfunctioning outer or middle ear in patients with conductive hearing loss. Additionally, BC is the most reliable way to determine the inner ear status [8]. Despite the large-scale applications of BC, the exact mechanisms are to date poorly understood, which is partly due to the embedding of the cochlea in one of the hardest bones of the human body, making it challenging to research and sparking the need for additional experimental approaches [9].

One way to investigate the differences between AC and BC is using subjective LB, where participants must equalize an AC and BC stimulus. A different loudness growth between AC and BC stimuli may be due to different frequency-dependent pathways [1,10]. While in AC mid-frequencies (1.5 – 4 kHz) are mainly amplified, the five BC pathways each have their specific amplification characteristics. The first pathway expands and compresses the ear canal, amplifying frequencies up to 0.7 kHz [11,12]. In the second pathway, the middle ear ossicles move with a relative motion concerning the inner ear due to inertia and pass sound towards the inner ear with greater amplification between 1.5 – 3.5 kHz [1,13–15]. In the third pathway, the cochlear fluid inertia causes a greater amplification for frequencies below 4 kHz [16,17]. Fourthly, compression and expansion of the cochlear bone have the biggest influence on frequencies above 4 kHz [13,18]. Finally, in the fifth pathway, pressures originating from the interior skull vibrate the cochlea through the cochlear and vestibular aqueduct, however exact characteristics remain unknown[19].

In addition to LB, we used Optical coherence tomography (OCT) to overcome the limitations of LDV, ICP and LEM, by enabling in-depth and longitudinal imaging and vibrometry of intracochlear structures through the intact RWM. OCT makes it possible to visualize the osseous spiral lamina (OSL), BM, scala tympani (ST), and scala vestibuli (SV) in the hook region of the

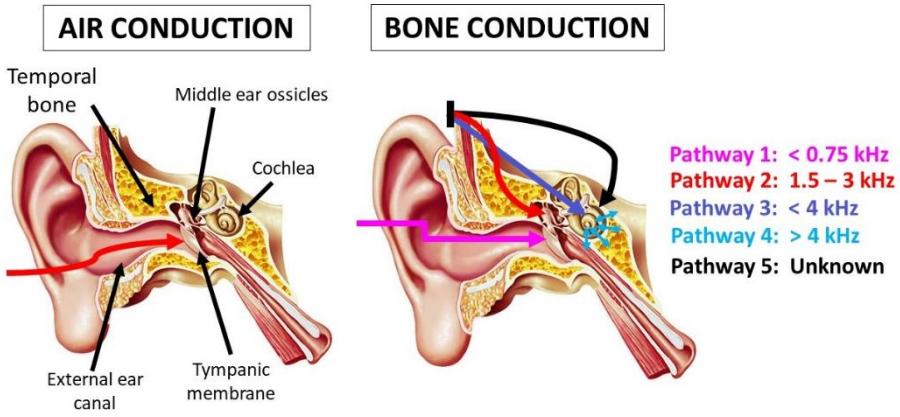


FIGURE 1: Illustration of air and bone conduction pathways. For BC the pink arrow indicated the lowest frequency pathway in the external ear canal, the red arrow indicates the mid-frequency pathway, passing the middle ear ossicles. The green arrow indicates the fluid inertia in the cochlea and the black arrow the compression and expansions of the cochlea. The blue arrow indicates the BC pathway through the interior skull. AC sound follows one pathway, indicated by the red arrow. Fig. based on Stenfelt et al. 2016 [39].

human cochlea non-invasively [20]. Besides applying OCT vibrometry (vOCT) to the middle ear structures, multiple studies illustrated the value of intracochlear vOCT through the RWM in rodents [21–26]. We tested the feasibility of vOCT on human temporal bones (TB) and cadaver heads (CH) to investigate the differences between AC and BC sound on human intracochlear structures based on LB tests, assuming a similar psychoacoustic perception of AC and BC stimuli.

MATERIALS AND METHODS

Psychoacoustics

During the LB test, 15 subjects aged between 18 and 25 years old had to match a fixed AC sound at 60 dB HL to an adaptable BC sound. Subjects were included if they showed no sign of any ear diseases, obstructing ear wax and had no history of otologic problems, noise exposure, genetic hearing loss or ototoxic drug use. Subjects were screened using pure tone audiometry (Madsen Orbiter 922-2, TDH-39 headphones) and were included if their Fletcher Index (FI) was < 20 dB HL (ISO 8253-1:2010).

The LB test was administered using a graphical interface in MATLAB (version R2018b), a laptop was connected to an RME Fireface UC sound card to stimulate the ER-2 insert phones (50 Ohm) for AC stimuli, and the RadioEar B71, placed 4.5 cm posterior-inferior from the external ear canal, for BC stimuli. The ear with the best FI was tested, while the non-tested ear was masked using narrow-band noise of 70 dB HL. For each frequency (0.5 kHz, 0.75 kHz, 1 kHz, 1.5 kHz, 2 kHz, 3 kHz, 4 kHz, and 6 kHz) the BC stimulus was first applied, followed by the AC stimulus. Subjects could adjust the BC loudness level via the graphical interface by clicking buttons labeled '---','--','-','=','+','++','+++', corresponding to '-4dB','-2dB, '-1dB', '0dB', '1dB', '2dB' and '4dB' respectively (Appendix A). The subjects were blinded to the loudness level of the corresponding steps. Subjects could proceed to the retest condition if an equally loud perception was reached. The outcome measures of this test are used as input for vOCT.

OCT Imaging and Vibrometry

Sample Preparation

One human temporal bone (TB) and one human cadaveric head (CH) from anonymous donors were retrieved from the Vesalius Institute of the University of Leuven, in agreement with the Helsinki declaration and approved by the Medical Ethics Committee of the University Hospitals of Leuven (S65502). Both the TB and CH were preserved in the freezer at -20° and thawed at 4° a day before the experiment. Visualization of the RWM was obtained by a mastoidectomy, posterior tympanotomy, facial nerve removal and widening of the round window niche. In the CH, the RWM was covered by a pseudomembrane, which was carefully removed before OCT imaging. During the experiment, the samples were regularly moisturized with saline solution to optimize the visualization of intracochlear structures on OCT and tissue preservation. The TB was used for two measurements, while each side of the CH was used for one measurement, for a total of 4 different measurements.

To investigate the intracochlear anatomy through the RWM, a Thorlabs SD-OCT system (Telesto TEL320C1) with a central wavelength of 1300 nm, lateral resolution of 20 µm, and axial resolution in air and water of 5.5 µm and 4.2 µm, respectively, and an objective lens (LSM04) with a working distance of 42.3 mm and effective focal length of 54 mm were used. After determining the best point of view using 3D OCT images, the samples were stabilized in a TB holder or plastic container.

OCT vibrometry to investigate motion of intracochlear structures was carried out using the ThorImage software (version 5.1.3RC0) in combination with a MATLAB script driving the RME Fireface sound card, which in turn controls the insert phone in the AC test and the Radioear B71 for the BC test (Appendix B). The stimulus amplitude per frequency was based on the results of the LB test. A sine wave of 4 periods was applied and synchronized with A-scans, a total of 60 frames each at a different phase from the start of the sine wave stimulation were taken. This results in a series of snapshots at a different time during the stimulation, and displacement is visible as a difference in phase of the interference at each pixel. The data were stored as .oct files and analyzed in MATLAB. The velocity (v) was derived from the phase differences (ϕ) between the A-scans, using $v = (\lambda) * \phi / (4 * \pi * (1/f))$, where lambda ($\lambda$) was the central wavelength of the infrared light source (1300 nm) and f the frequency of the A-scans (76 kHz). The signal to noise ratio was improved using a Gaussian smoothing and detrending filter alongside a correlation-filter [27]. If the structure of interest was not perpendicular to the infrared light, a cosine correction for the velocity amplitude was applied (Appendix C).

Statistical Analysis

Statistical analysis was carried out in MATLAB. For all tests an α-level of 0.05 was used. LB data were checked for normality using box-whisker plots, the Shapiro-Wilk, and Levene's test. Following normal distribution, a paired t-test was performed to determine the similarities between the test and retest conditions.

Absolute velocity outcomes of the vOCT measurements were tested for normality using the Kolmogorov-Smirnov test, carried out for every AC and BC measurement per frequency, and Levene's test to check for homoscedasticity between AC and BC measurements in the same sample. Because two measurements on the TB sample happened at different times, a paired t-test was performed to check the intra sample consistency for the same stimulation pathways. To compare absolute amplitude of the velocity in AC and BC, the Welch's t-test was used.

RESULTS

Psychoacoustics

Results for the LB test indicate different force levels are needed to achieve the same perception between AC stimuli of 60 dB HL with BC stimuli (Fig. 2). No significant difference was found for the test-retest conditions (Appendix D).

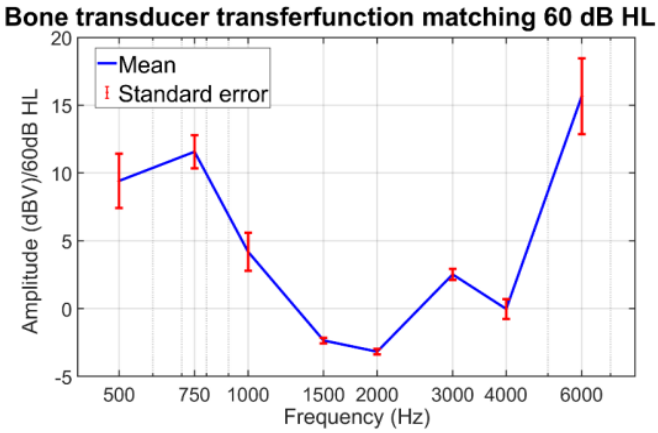


FIGURE 2: Results for loudness balancing tests (n=15), illustrating the bone transducer transferfunction matching 60 dB HL in air conduction. These results illustrate that different force levels for the bone transducer are needed to achieve the same perception.

OCT Imaging and Vibrometry

In the CH the RWM, SV, ST, promontory, posterior crus and head of the stapes, and the OSL were visible. The stapes was imaged upside-down, to image all structures of interest in one view (Appendix E). On top of these, in the TB, the vestibular and tympanic plate of the OSL, BM, parts of the Reissner's membrane, bridge and spiral limbus were also visible (Fig. 3 a, b).

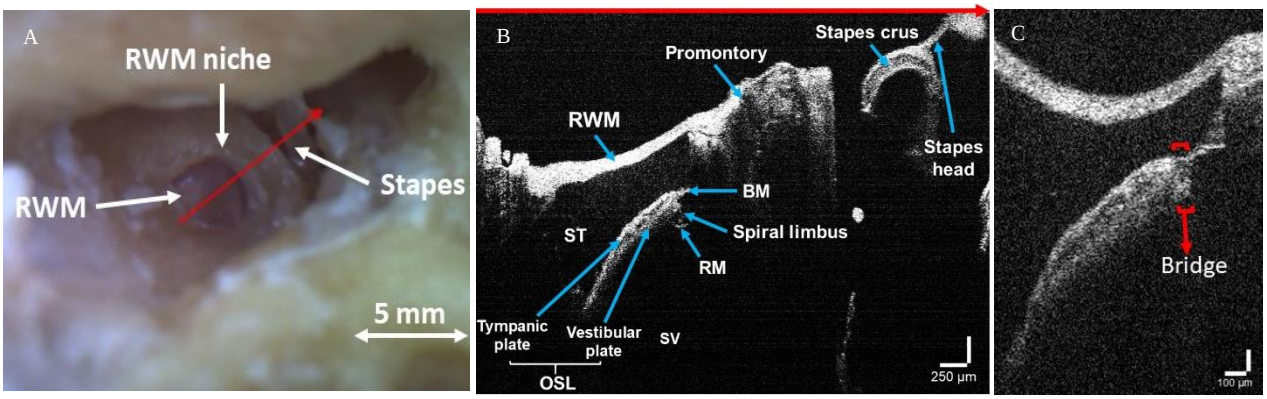


FIGURE 3: Fig. 3a: microscopic view on the round window membrane (RWM) in the temporal bone. Fig. 3b: OCT image of the hook region and stapes, abbreviations: basilar membrane (BM), Reissner's membrane (RM), Scala tympani (ST), scala vestibuli (SV), Osseous spiral lamina (OSL). Fig. 3c: Indication for the location of the cochlear partition bridge

Intracochlear amplitudes of the velocity (mm/s) and phase (°) in the TB for AC and BC stimulation based on the LB are compared using vOCT (Fig. 4 and 5). The phase is in reference to the input signal for all results. The right cochlea of the CH was excluded since the perilymph was found to be partially frozen after measuring.

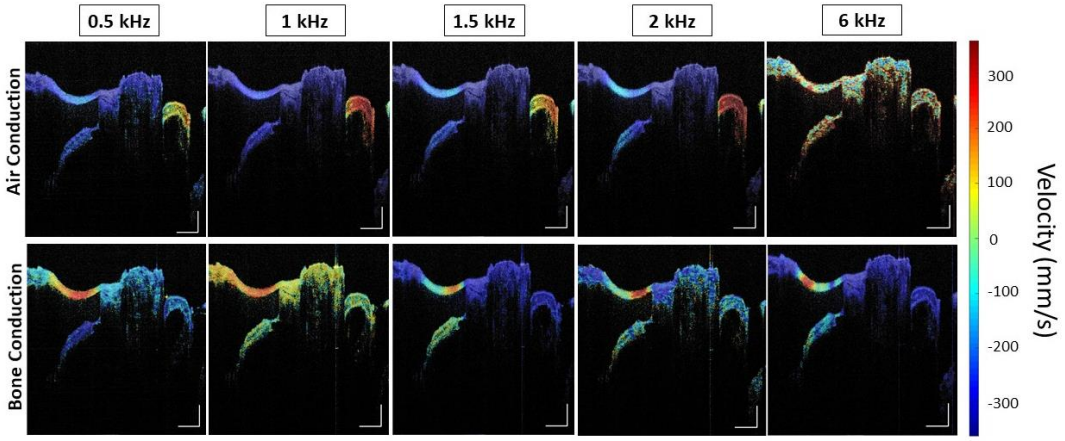


FIGURE 5: Results from OCT vibrometry in the temporal bone at different frequencies for air and bone conduction, with colors indicating the velocity in mm/s.

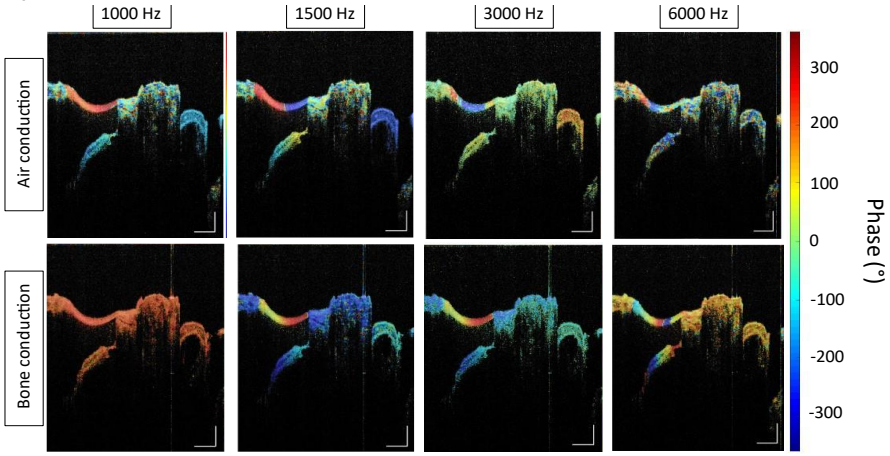


FIGURE 4: Phase results from OCT vibrometry in the temporal bone at different frequencies for air and bone conduction, with colors indicating the phase in degrees (°).

Statistical Analysis of BC and AC Induced Motion

Intra-sample comparison revealed a higher consistency for AC pathways, while BC pathways showed inconsistent results for 0.5, 1.5 and 4 kHz. The Welch's t-test revealed significantly different OSL, RWM and BM motion between AC and BC at various frequencies (Appendix F).

DISCUSSION

Psychoacoustics

The investigation of equal LB indicates validity as the test and retest conditions were not significantly different from each other and robust against outliers. The bone transducer transfer function illustrates that different force levels are needed to equalize BC and AC sound, which can be explained by the frequency-dependent pathways [1,10]. Additionally, the lowest force levels and inter-subject variation between 1.5 and 3 kHz could be due to skull impedances, indicating less force is necessary to transfer mid-frequencies [28]. Higher inter-subject variation and force levels for 0.5, 0.75, 4 and 6 kHz may be explained by harmonic distortion and overtones, occurring at an intensity of 60 dB HL [29,30].

OCT imaging and Vibrometry

In addition to previous findings, the spiral limbus, the Reissner's membrane, and considerable a part of the tectorial membrane are visible [20]. In the CH, the view on the RWM was limited, and fewer structures were visible. The non-bony cochlear partition bridge is assumed present as an area with reduced scattering compared to the bony OSL (Fig. 3c) [31,32].

The vOCT results in fig. 4 illustrates the stapes velocity as an input drive for AC stimuli and the promontory velocity is fixed at approximately -300mm/s over the entire frequency range, indicating no other induced pressures than the from stapes stimulate the BM. For BC the promontory and stapes velocity are similar and varying between 100 and -300 mm/s, indicating other intracochlear pathways and resulting pressures contribute to BC induced motion of intracochlear structures. Based on literature, we might expect a higher stapes velocity between 1.5 and 3 kHz, however it is hypothesized this might be different in a CH compared to a TB due to mass reduction and exclusion of other pathways [7,33–35].

We were able to capture the motion in response to AC and BC on the BM and OSL, a small amplitude difference in the velocity between the BM and OSL in BC, but not for AC, indicating the OSL as a non-stationary contributor in a BC pathway [6,7,36]. At 2 kHz AC the distal part of the OSL has a lower amplitude (~ 100 mm/s) compared to the proximal part and BM (~ 300 mm/s), this hook region, we expect high frequencies to represent a higher velocity amplitude. Unfortunately, the measurements at 6 kHz AC are too noisy to compare the results to.

The stapes and OSL have the same phase up to 1 kHz in AC, and the RWM moves with the opposite phase because of the incompressible fluid (Fig. 5). With increasing frequency, multiple nodes occur on the RWM and OSL in AC and BC, indicating modal vibrations [34]. The most proximal part of the OSL and the BM showed the same phase in BC and AC, which hint at the location of the bridge [7,34,35].

Velocity results using vOCT were compared for AC and BC between and within samples (Appendix F). Since 2D data was collected, comparisons between samples tested at a different time was not straightforward and a Welsh's t-test indicated similarities for AC but not for BC pathways. This might be due to a slightly different location of the bone transducer [37,38]. Results within a sample to compare the amplitude of the velocity and phase in AC and BC stimulation for the TB and CH (paired t-test), indicated differences between the two stimulation modes on the RWM, Stapes, OSL and BM, conform with previous LDV experiments [34,35]. In AC the results confirm the stapes as a major driving force, for BC the importance of other pathways is indicated.

CONCLUSION

These are the first measurements to the author's knowledge using OCT vibrometry to investigate the differences between AC and BC on human intracochlear structures based on psychoacoustics loudness balancing. For LB equalized AC and BC stimuli, we observed intracochlear velocity and phase differences for the BM, OSL, stapes, RWM and promontory are observed. These results illustrate the value of vOCT as an experimental technique for intracochlear investigations and its great potential to investigate the mechanics of hearing without disrupting the intracochlear homeostasis, overcoming limitations and providing in-depth and longitudinal motion information on top of LDV, ICP and LEM.

ACKNOWLEDGMENTS

NV acknowledges funding by Research Foundation Flanders (FWO senior clinical investigator fund 1804816N), TP by Flanders Innovation and Entrepreneurship (IWT155047) and Research Foundation Flanders (FWO 12Y6919N), and AS by Research Foundation Flanders (FWO 1S78519N).

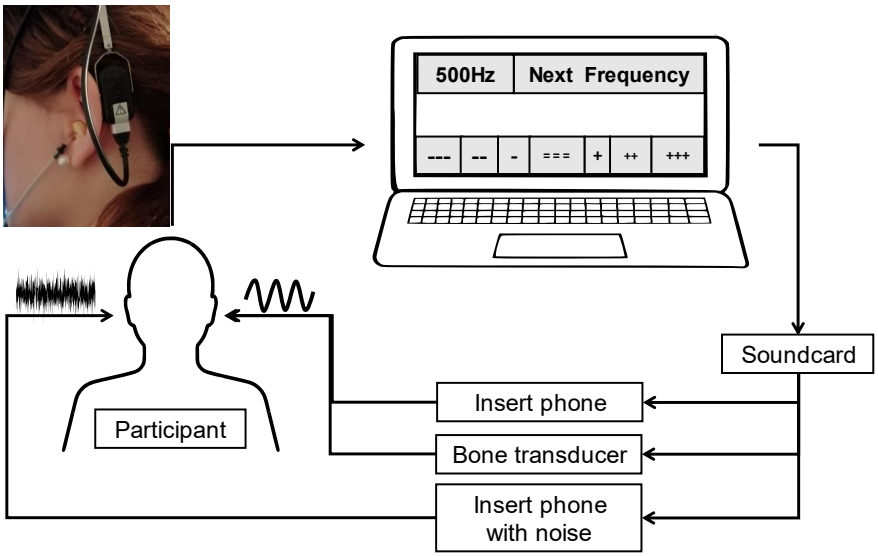
REFERENCES

1. Stenfelt, and Zeitooni, *Hearing Research*, **301**, 85–92 (2013).
2. Stenfelt, *Hearing Research*, **340**, 135–143 (2016) http://dx.doi.org/10.1016.
3. Békésy v G, *Ann Phys*, **5**, (1932).
4. Puria, and Rosowski, *Hearing Research*, **293**, 21–30 (2012) http://dx.doi.org/10.1016.
5. v. Békésy, *Ann Phys*, **405**, 111–136 (1932) https://doi.org/10.1002.
6. Raufer, Idoff, Zosuls, Marino, Blanke, Bigio, et al., *JARO - Journal of the Association for Research in Otolaryngology*, **21**, 171–182 (2020).
7. Stenfelt, Puria, Hato, and Goode, *Hearing Research*, **181**, 131–143 (2003).
8. Dauman, *European Annals of Otorhinolaryngology, Head and Neck Diseases*, **130**, 209–213 (2013) http://dx.doi.org/10.1016.
9. Bergevin, "How exceptional is the ear?", (n.d.).
10. Stenfelt, and Håkansson, "Air versus bone conduction: an equal loudness investigation", (n.d.).
11. Stenfelt, Wild, Hato, Goode, Stenfelt, Wild, et al., *JASA*, **902**, (2005).
12. Tonndorf, TONNDORF, and J., *Acta Otolaryngol*, **213**, Suppl 213:1+ (1966) http://ci.nii.ac.jp.
13. Stenfelt, *J Acoust Soc Am*, **119**, 2848–2858 (2006).
14. Dobrev, and Sim, *Hearing Research*, **364**, 96–103 (2018) https://doi.org/10.1016.
15. Dobrev, Farahmandi, and Röösli, *Hearing Research*, **395**, (2020).
16. Eeg-olofsson, Stenfelt, Taghavi, Reinfeldt, Håkansson, Tengstrand, et al., *Hearing Research*, **306**, 11–20 (2013) http://dx.doi.org/10.1016.
17. Chang, Kim, and Stenfelt, *AIP Conference Proceedings*, **1703**, (2015).
18. Hudde, 3 A Functional View on the Peripheral Human Hearing Organ, in: Communication Acoustics, (2005), pp. 47–73.
19. Roosli, Dobrev, Sim, Gerig, Pfiffner, Stenfelt, et al., *Otology and Neurotology*, **37**, e384–e390 (2016).
20. Starovoyt, Putzeys, Wouters, and Verhaert, *Scientific Reports*, **9**, (2019).
21. van der Heijden, and Vavakou, *Hearing Research*, (2021).
22. Cooper, Vavakou, and van der Heijden, *Nature Communications*, **9**, (2018).
23. Dong, Xia, Raphael, Puria, Applegate, and Oghalai, *J Neuro-Physiol*, **120**, 2847–2857 (2018) www.jn.org.
24. Cho, Wang, and Puria, (n.d.) https://doi.org/10.1101.
25. Matthews, and Adamson, *Curr Opin Otolaryngol Head Neck Surg*, **28**, 296–301 (2020).
26. Cho, Wang, and Puria, *JARO - Journal of the Association for Research in Otolaryngology*, **23**, 195–211 (2022).
27. Hinton, Sayers, Solartron, and Rd, "Impedance Measurement Techniques: Sine Correlation", (1998).
28. Stenfelt, Puria, Hato, and Goode, *Hearing Research*, **181**, 131–143 (2003).
29. Lightfoot, and Hughes, "Bone conduction errors at high frequencies: implications for clinical and medico-legal practice", (1993).
30. Lundgren, "Bone Conduction Transducers and Output Variability Lumped-parameter modelling of state variables", (n.d.).
31. Raufer, Guinan, and Nakajima, *Proc Natl Acad Sci U S A*, **116**, 13977–13982 (2019).
32. Raufer, Idoff, Zosuls, Marino, Blanke, Bigio, et al., *JARO - Journal of the Association for Research in Otolaryngology*, **21**, 171–182 (2020).
33. Stenfelt, *Frontiers in Neurology*, **11**, (2020).
34. Stenfelt, Hato, and Goode, *Hearing Research*, **198**, 10–24 (2004).
35. Stenfelt, Hato, and Goode, *J Acoust Soc Am*, **115**, 797–812 (2004).
36. Raufer, "Human Inner Ear Mechanics Studied with Experimental, Anatomical, and Computational Approaches A dissertation presented", (2019).
37. Dobrev, Stenfelt, Röösli, Bolt, Pfiffner, Gerig, et al., *International Journal of Audiology*, **55**, 439–446 (2016).
38. Fierens, Borgers, Putzeys, Walraevens, van Wieringen, and Verhaert, *BioMed Research International*, **2022**, 1–16 (2022) www.hindawi.com.
39. Stenfelt, *Hearing Research*, **340**, 135–143 (2016).

APPENDICES

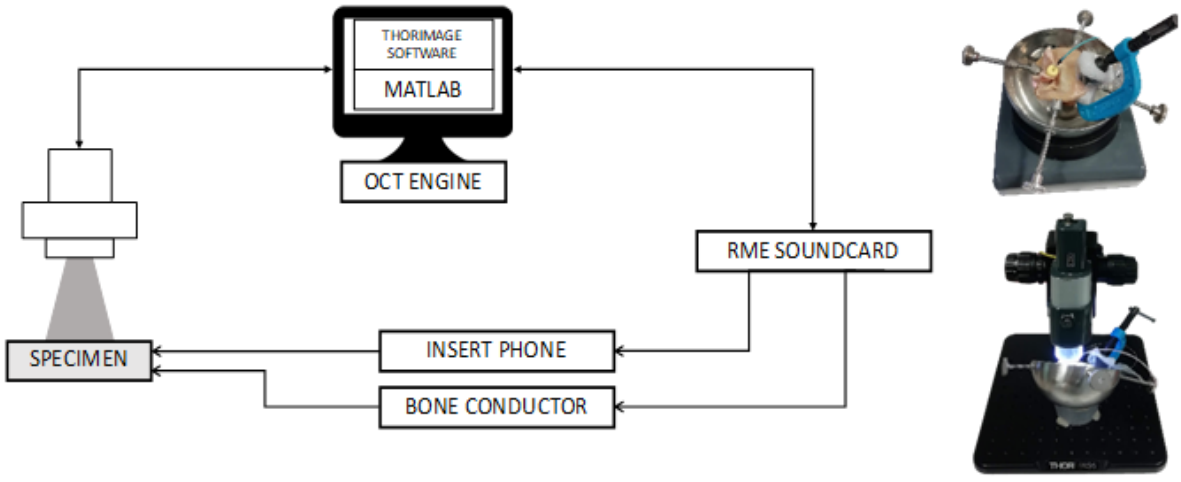
Appendix A – Setup Loudness Balancing Test

Schematic illustration of the loudness balancing test. The image left above shows the position of the bone transducer. In the schematic representation the laptop with the loudness balancing test, inputs and the soundcard are illustrated. The soundcard generates the stimuli through the insert phone and bone transducer on the tested (best) ear. The insert phone with noise was placed in the non-tested ear. The participant could adjust the loudness of the BC stimulus via the laptop, with the scaling illustrated on the laptop representation.



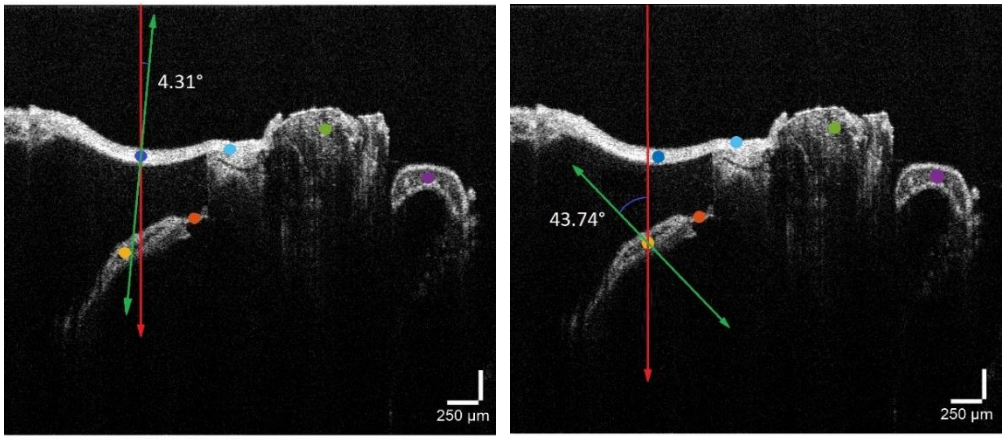
Appendix B – Setup OCT-Vibrometry

A schematic representation of the setup was illustrated on the left an external trigger from MATLAB was generated and send to the RME soundcard. From there the insert phone and bone transducer presenting the stimuli on the specimen. In the middle an illustration of the insert phone and bone transducer mounted on the temporal bone. On the right an illustration of the setup under the OCT scanning system.



Appendix C: Effective Velocity

Correction angle to measure the effective velocity using $Veffective = Vmeasured/cos(\alpha)$, where α is the angle between the infrared light and the effective velocity motion direction. On the left OCT image, the correction angle of 4.31° was measured for the dark blue dot on the RWM. On the right this was done for the OSL (yellow dot) where the correction angle was 43.74°. The red line represents the infrared light from the laser beam, the green line is the direction of the expected effective movement of the structure.



Appendix D: Outcomes Loudness Balancing tests

Statistical outcome measures of the loudness balancing test, to measure the differences between the test and retest condition. These p-values indicate no significant differences.

Freq (Hz)	500	750	1000	1500	2000	3000	4000	6000
p-value	0.1660	0.2973	0.1057	0.1931	0.2618	0.0678	0.8289	0.6041

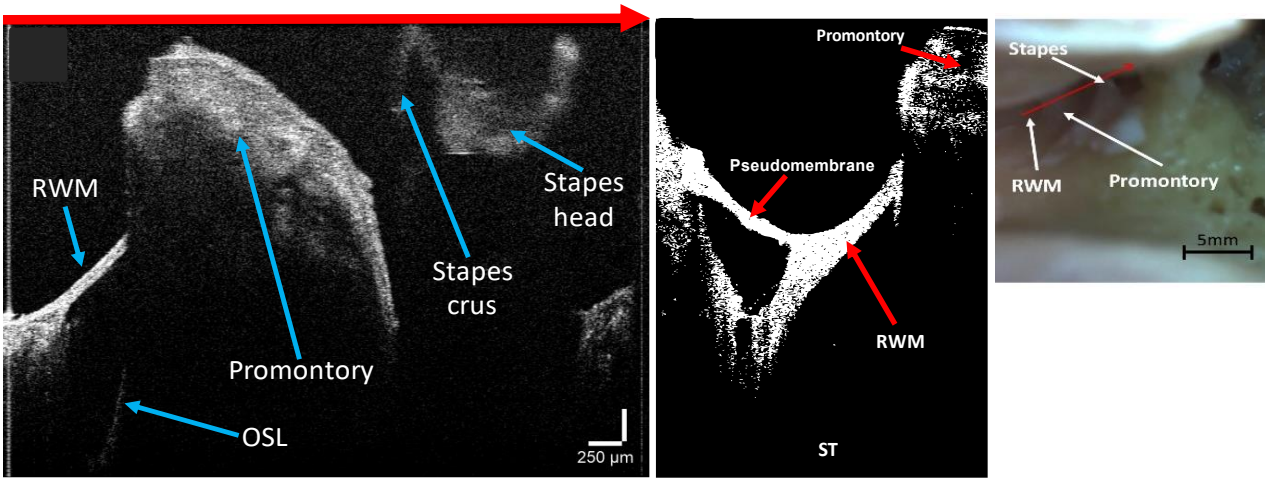
Outcome values of the psychoacoustic test used as stimulation input for the vibrational OCT measurements.

Freq (Hz)	500	750	1000	1500	2000	3000	4000	6000
BC (V)	2.96	3.79	1.62	0.76	0.69	1.33	1.00	6.06
AC (V)	0.70	0.50	0.40	0.50	0.60	0.40	0.35	0.9
Masking (V)	1.20	1.00	0.90	1.20	1.25	1.05	1.10	1.45

Appendix E: Scanning Parameters and OCT in cadaver head

Scanning parameters: The field of view and reflectivity of the structures differed dependent on the sample. With a refractive index of 1.00 in the TB's, the imaging depth was approximately 3 mm and the width was dependent on the orientation of the stapes to the RWM. To visualize both in the same field of view approximately 4.80 mm was needed, similar to the CH. However, for the CH the imaging depth was 3.55 mm.

The image in the middle illustrates the pseudomembrane on the RWM in the cadaver head, before it was removed.



Appendix F: Statistic results of OCT vibrometry

Comparison using the paired t-test of the absolute velocities as a reaction to an AC or BC stimulus, measured on the same structure. A p-value of < 0.05 indicates significant differences in the AC and BC motion.

TBL1: First measurement on the left temporal bone
TBL2: Second measurement on the left temporal bone
CH L: Measurement on the left side of the cadaver head

ABSOLUTE VELOCITIES AC vs BC in same sample (Paired t-test)

RWM - TB L1								
Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
p -value	**0.0006**	**0.00001**	**0.0285**	0.6916	0.2785	**0.0011**	**0.0005**	**0.0145**

RWM - TB L2								
Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
p -value	**0.00003**	**0.0041**	**0.0006**	0.1123	**0.0029**	0.066	**0.0207**	0.175

RWM - CH L								
Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
p -value	0.4661	**0.0074**	0.2991	0.6299	0.9201	0.7193	0.1766	0.1273

Stapes - TB L1								
Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
p -value	0.2868	**0.00018**	0.4899	0.2142	0.1667	**0.0145**	**0.0025**	**0.016**

Stapes - TB L2								
Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
p -value	0.1487	**0.00002**	**0.0009**	**0.00003**	0.1297	0.9352	0.3776	0.1168

Stapes - CH L								
Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
p -value	0.5167	**0.0349**	**0.0446**	**0.0042**	0.0574	0.8029	0.4268	0.0615

OSL - TB L1								
Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
p -value	0.0571	**0.0022**	0.8781	**0.0002**	**0.0006**	0.0971	**0.0401**	**0.0021**

OSL - TB L2								
Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz

p -value	**0.0495**	**0.0006**	**0.0009**	**0.0334**	0.9622	0.6353	0.7493	**0.0124**

OSL - CH L								
Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
p -value	0.0665	0.388	0.3189	0.0569	**0.0262**	0.2325	0.2703	0.2784

BM - TB L1								
Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
p -value	0.5007	**0.0003**	0.0688	**0.0168**	**0.0008**	**0.00005**	**0.0124**	**0.0018**

BM - TB L2								
Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
p -value	**0.0096**	**0.0143**	**0.0026**	0.9878	0.3952	0.3118	**0.0474**	0.0583

RM - TB L1								
Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
p -value	**0.00007**	**0.0074**	**0.0018**	0.9216	0.3099	**0.0388**	**0.0038**	0.0579

RM - TB L2								
Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
p -value	**0.0017**	0.0752	**0.0263**	**0.0136**	0.815	0.155	0.2619	0.3082

Welsh's t-test relative velocities:

Comparison of the relative velocities between samples on the same structure of interest. The AC velocity was normalized to the stapes, while the BC velocity was normalized to the promontory. A red value indicates significant differences for that frequencies on the selected structure of interest between the samples.

RWM

AC	TBL L1 vs. CH L							
Welch's t-test	Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
	p-value	0.2523	0.1626	0.1928	0.1938	**0.021**	0.7445	0.3246	0.1711
AC	**TB L2 vs. CH L**								
Welch's t-test	Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
	p-value	0.2277	0.3847	0.1722	0.0512	0.7907	0.5669	0.1662	0.3146

RWM

BC	TBL L1 vs. CH L								
Welch's t-test	Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
	p-value	0.0961	0.5728	**0.0136**	**0.0099**	**0.0174**	0.1831	0.292	0.5914
BC	**TB L2 vs. CH L**								
Welch's t-test	Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
	p-value	0.3805	**0.0298**	**0.0142**	0.2189	0.4816	0.3149	**0.0122**	0.5261

OSL

AC	TBL L1 vs. CH L								
Welch's t-test	Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
	p-value	0.225	0.0878	**0.0331**	**0.0186**	**0.0089**	0.4758	0.4893	0.2094
AC	**TB L2 vs. CH L**								
Welch's t-test	Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
	p-value	0.456	0.2793	**0.0008**	**0.0052**	0.3442	0.3683	0.8974	0.2141

OSL

BC	TBL L1 vs. CH L								
Welch's t-test	Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
	p-value	0.181	0.5621	0.2381	**0.0223**	0.0737	0.7594	0.9532	0.5448
BC	**TB L2 vs. CH L**								
Welch's t-test	Freq	500Hz	750Hz	1000Hz	1500Hz	2000Hz	3000Hz	4000Hz	6000Hz
	p-value	0.181	0.5621	0.2381	**0.0223**	0.0737	0.7594	0.9532	0.5448

Auditory Evoked Potentials in Comparison to Hearing Thresholds

Fang Wang [a)] and Frank Böhnke [b)]

Department of Otorhinolaryngology, Klinikum rechts der Isar, TU Munich, Germany

a) Corresponding author: harryfangfang520@163.com
b) frank.boehnke@tum.de

Abstract. Auditory Steady State Response (ASSR) is now commonly used to identify the Pure Tone Audiogram (PTA), especially for uncooperative humans. This work examined the correlation between pure tone threshold audiograms and ASSR at four frequencies (500 Hz, 1 kHz, 2 kHz, and 4 kHz) using a modulation frequency 40 Hz. Statistical evaluations were performed by correlation analysis and the Bland-Altman method, where the variance of the differences from measurement pairs was assessed based on the mean of these measurement pairs. The PTA value in decibels (dB) has established as the gold standard. Paired measurements are determined by the gold standard, while ASSR measurements are made by two measuring systems. Results: (1) The difference between ASSR thresholds and PTA thresholds were frequency-dependent, that is, it decreased with increasing frequency. (2) The ASSR is a useful complementary measuring tool to detect hearing thresholds in normal hearing subjects.

INTRODUCTION

Many different auditory evoked responses can be recorded from humans, and each of them can provide useful information about the auditory function. Pure Tone Audiometry (PTA) is the gold standard to evaluate the frequency dependent hearing thresholds. However, we cannot always measure the reliable thresholds with PTA for all patients, such as some children and patients who have neurologic delays. The Auditory Steady State Response (ASSR) is now generally used to identify the PTA, especially in case of uncooperative humans.

The ASSR is divided into two categories according to the modulation frequency: 40 Hz ASSR and 80 Hz ASSR [1]. The higher modulation frequency ASSR is less affected by the alert state [2], so it is suitable for the hearing test of sedated and sleeping children [3]; the low stimulation rate ASSR has the highest response amplitudes when the subject is awake.

ASSR is divided into single-frequency stimulation and multi-frequency stimulation according to the stimulation methods. Single-frequency stimulation means that the stimulation sounds of 500 Hz, 1 kHz, 2 kHz and 4 kHz are applied separately, while the multi-frequency stimulation uses the above four frequencies at the same time. Even eight signals from both ears (four in each ear) can be applied at the same time [4]. The judgment of the response result is automatically completed by the computer according to certain standards, avoiding the judgment of the examiner's subjective bias [5]. In addition, when applying multiple frequencies simultaneously, the test time is greatly shortened and necessary for clinical practice.

Our purpose is to evaluate the correlations between hearing thresholds detected by PTA and the 40 Hz multi-frequency ASSR in normal hearing adults and explore its value for clinical applications. For comparison ASSR was measured by two machines produced by the companies: PATH-Medical, Germany & Interacoustics, Denmark. We test if ASSR is a reliable method to determine hearing thresholds.

MATERIALS AND METHODS

Former studies showed no significant difference between multi-frequency or binaural sound delivery methods and the single-frequency sound delivery method, that is, there is only low mutual interference of cochlear and neural signals, generated by ipsilateral and contralateral ears [3-5]. However, it has also been reported that, except for 500 Hz, there is no significant difference in the results obtained by single-frequency stimulation and multi-frequency stimulation, indicating that multi-frequency stimulation at 1 kHz, 2 kHz and 4 kHz is not affected [6]. The frequency 500 Hz is affected related to the fact that the environmental noise has higher low-frequency sound components, and the response threshold at 500 Hz is susceptible to its interference. When multiple frequencies are given at the same time, the response threshold of high frequencies and the low frequency (500 Hz) is different. A certain inhibitory effect will increase the response threshold [7]. Therefore, it is recommended to use single-frequency stimulation recording at 500 Hz in clinical applications. In this experiment, stimulation is usually started at 50 dB SPL, and eight frequencies are used for simultaneous stimulation in both ears. When it is close to the threshold value, the threshold value is determined by a single ear and single frequency stimulation, so as to exclude possible mutual influence between frequencies as much as possible.

Patient selection

The study group comprised of 12 subjects (24 ears), aged 20–65 years (mean age 29.5 y): 3 females and 9 males. All subjects underwent an otoscopic examination, one PTA, and two ASSR. Only the subjects with normal otoscopic findings were included in the study. Patients with conductive and mixed hearing loss were excluded. All subjects had no history of ear disease or noise exposure.

Instruments and Test Environment

PTA adopts the AT1000 (Auritec GmbH, Germany) pure tone audiometer. ASSR were measured with machines of two companies (PATH-Medical, Germany & Interacoustics, Denmark). Audiological examinations were conducted in double-shielded isolation booths (acoustic and electromagnetic). The background acoustic noise level was less than 20 dB (SPL) at the tested frequencies.

PTA test

The PTA tests were done between 125 Hz and 8 kHz but only four values up to 4 kHz were evaluated in this study. The work examines the correlation between pure tone threshold audiograms of normal hearing subjects due to the World Health Organization (WHO) classification and ASSR at four frequencies (500 Hz, 1 kHz, 2 kHz and 4 kHz). Normal hearing (WHO 0) exists if the average of hearing loss at these four frequencies is less than 25 dB.

ASSR test

The ASSR measurements were conducted using both Interacoustics and PATH-Medical devices. During the examination, subjects were awake and a stimulus was presented via calibrated circumaural headphones. Stimuli were 500 Hz, 1 kHz, 2 kHz and 4 kHz sinusoids modulated with 40 Hz and degree of modulation m = 1 . Electrodes were placed on both mastoids (-Active), at the vertex (+Active), and on the right cheekbone (Ground). The electric impedance for all electrodes was 3 kΩ or less. The initial acoustic pressure level was 50 dB (SPL) for both machines. In case of Interacoustics, the sound level is decreased by steps of 10 dB until no response could be elicited; then the sound level is increased by 5 dB steps, and the minimum sound levels that elicited responsess at four frequencies were determined as ASSR thresholds separately for both ears. In case of PATH the aforementioned steps were implied similarly, but simultaneously for both ears of a subject. The average measuring durations were 35 min in case of Interacoustics and 25 min (PATH) for each subject and both ears.

Statistical methods

Statistical analysis of the test results was carried out using SPSS19.0 statistical software. Pure tone thresholds and ASSR thresholds were expressed as mean values ± standard deviation s. The agreement between the pure tone hearing thresholds and the ASSR thresholds averaged for four frequencies was analyzed by the Bland-Altman method, which constitutes the difference (ordinate) along the sum (abscissa) of values, each scaled in dB (Figure 3).

The study protocol was approved by the local Ethics Committee of the Klinikum rechts der Isar (Munich, Germany). Informed consents were required for all subjects before they could participate.

RESULTS AND DISCUSSION

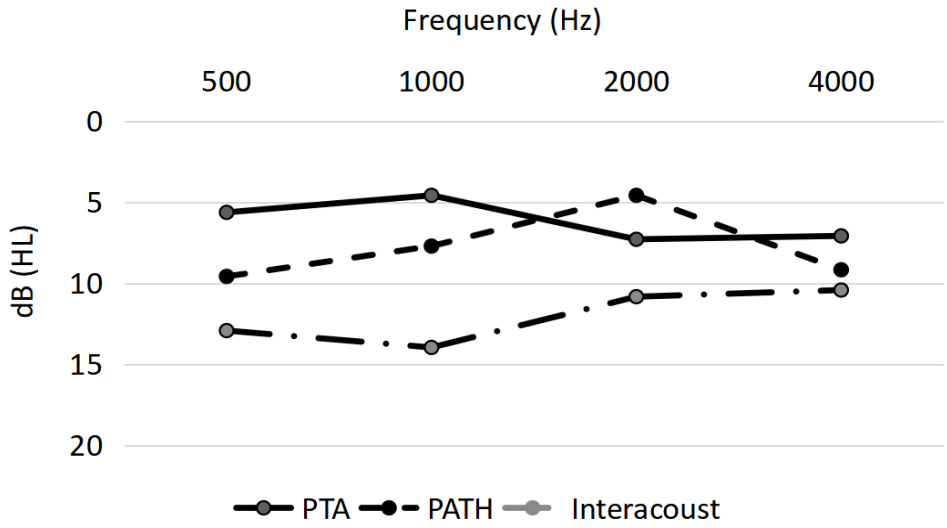


FIGURE 1. Frequency specific average (n = 24) of PTA, ASSR (PATH) & ASSR (Interacoustics)

At 500 Hz the PTA result is 5.63 dB and the ASSR results have large deviations of 3.95 dB and 7.29 dB for the PATH and Interacoustics machines, respectively. At 1 kHz the ASSR PATH result is near to the PTA (3.13 dB) and the Interacoustic has the largest deviation of 9.38 dB. The lowest deviations are found at 4 kHz with 2.09 dB (PATH) and 3.34 dB (Interacoustics). The differences between ASSR thresholds and PTA thresholds is frequency-dependent, i.e., decreases with increasing frequency [6-8], which is confirmed by our study. The ASSR thresholds at each frequency are higher than the PTA thresholds (except PATH result at 2 kHz). It is important to mention, that both manufacturers imply “correction factors” to their measured signals. In case of Interacoustics we used the original values without these factors, whereas in case of PATH a factor decreasing the thresholds by 10 dB was included (personal communication PATH engineers). The manufacturers apply these factors to adapt the ASSR measurements to PTA based on statistical analyses of large groups of normal hearing subjects.

Previous literature reported that response thresholds obtained by ASSR were higher than the PTA thresholds in both normal and hearing-impaired patients, with a difference between 10 and 20 dB [9]. The difference between frequencies is not significant, but the differences in our experimental results show that the ASSR threshold and PTA threshold are smaller than in previous reports.

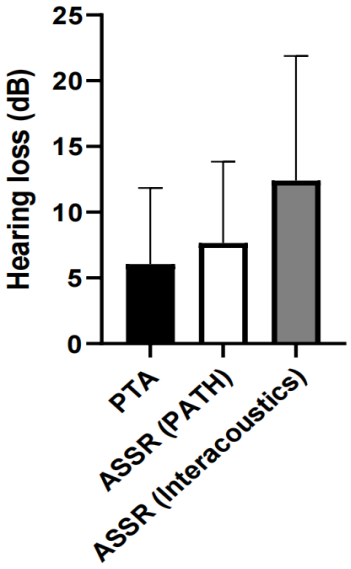
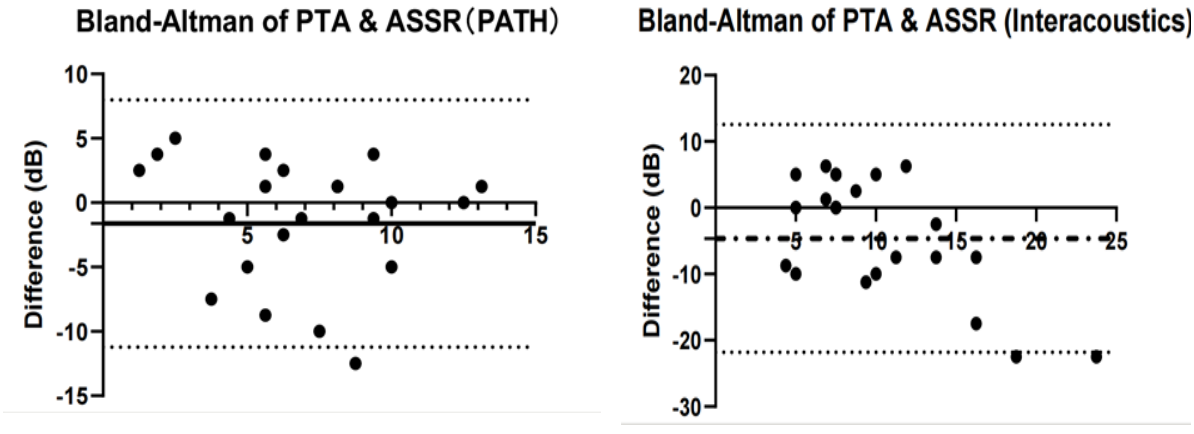


FIGURE 2. Average hearing loss in dB (n = 24) measured by three methods: PTA (Pure Tone Threshold), ASSR (PATH), ASSR (Interacoustics).

The averaged results for all frequencies are shown in Figure 2. The mean value of hearing loss of the PTA was 6.04 dB and a standard deviation s = 5.8 dB. In case of the PATH ASSR result the mean was 7.76 dB (s = 6.15 dB). The ASSR Interacoustics mean was 12.53 dB and had a higher standard deviation of s = 9.49 dB.

The solid line in Figure 3a denotes the bias – 1.61 dB (PATH) and in Figure 3b it is – 4.63 dB (Interacoustics). The dotted lines show the 95 % limits of agreement. In 3a it is from -11.21 dB to 7.98 dB, while in 3b it is from --21.82 dB to 12.54 dB. The statistical analysis by the Bland-Altman method shows the higher consistency to PTA of the PATH ASSR results in comparison to that of Interacoustics.



(a) (b)

FIGURE 3. (a) and (b) comparison of ASSR measurements (PATH & Interacoustics) by the Bland-Altman statistical analysis, x-axis plots the sum of values (dB).

CONCLUSION AND OUTLOOK

This study states a significant difference between the ASSR threshold and the PTA threshold, indicating that the application of ASSR in the evaluation of behavioral auditory thresholds needs to be revised, and the ASSR threshold cannot be directly used as a reference value for behavioral auditory thresholds. Nevertheless the results show that ASSR measurements are a valuable tool to detect the hearing thresholds in normal hearing subjects. ASSR can also be an excellent complementary method, together with other diagnostic methods, e.g. Brainstem Evoked Response Audiometry, represents a valuable tool in more precise determination of hearing thresholds. A next step will be the inclusion of patients with pathological ears.

Technically further ASSR studies are implemented to distinguish cochlear responses from brainstem and cortical responses by unmodulated sinusoidal acoustic stimulation and phase analysis of the auditory evoked responses, to enable a specific diagnosis of the auditory periphery.

REFERENCES

1. Stapells DR, Herdman AH, Smalls SA, et al. (2005) Current status of the auditory steady-state responses for estimating an infant's audiogram. In A Sound Foundation through Early Amplification: Proceedings of an International Conference: 43-59.
2. Cohen LT, Rickards FW, Clark GM (1991) A comparison of steady-state evoked potentials to modulated tones in awake and sleeping humans. J Acoust Soc Am, 90:2467-2479.
3. Aoyagi M, Kiren T, Kim Y, et al. (1993) Optimal modulation frequency for amplitude modulation following response in young children during sleep. Hearing Research, 65:253-261.
4. Dimitrijevic A, John SM, Van Roon P, et al. (2002) Estimating the audiogram using multiple auditory steady-state responses. J Am Acad Audiol, 13:205-224.
5. Picton TW, Skinner CR, Charnpagne SC et al. (1987) Potentials evoked by the sinusoidal modulation of the amplitude or frequency of a tone. Journal of the Acoustical Society of America, 82:165-178.
6. Perez-Abalo MC, Savio G, Torres A, et al. (2001) Steady state responses to multiple amplitude-modulated tones: an optimal method to test frequency-specific thresholds in hearing-impaired children and normal-hearing subjects. Ear Hear, 22(3): 200-211.
7. Herdman AT, Stapells DR (2001) Thresholds determined using the monotic and dichotic multiple auditory steady-state response technique in normal-hearing subjects. Scand Audiol, 30(1): 41-49.
8. Harris DM (1979) Action potential suppression, tuning with single fiber data. Hear Res, 1(2):133- 154.
9. Valdes JL, Perez-Abalo MC, Martin V, et al. (1997) Comparison of statistical indicators for the automatic detection of 80 Hz auditory steady-state responses. Ear Hear, 18(5):420- 429.

Sound-evoked vibrations along the tonotopic axis in the gerbil cochlea

Sebastiaan W.F. Meenderink[1] and Wei Dong[1, 2, a)]

[1)] *VA Loma Linda Healthcare System, Loma Linda, CA, 92374*

[2)] *Department of Otolaryngology – Head and Neck Surgery, Loma Linda University Health, Loma Linda, CA 92350*

[a)] *Electronic mail: wei.dong@va.gov*

Abstract. Our exquisite sense of hearing involves micromechanical inner-ear processes that precede sound transduction by hair cells. These mechanics are however not well understood, especially in the apical regions of the cochlea that are essential for hearing the relatively low frequencies important for the intelligibility of speech. Recent observations in the high-frequency region of the cochlea's sensory organ of Corti (ooC) revealed unanticipated and complex motions, but low-frequency responses are not known. Here, we measured sound-evoked motions from the outer hair cell (OHC) region and the lateral compartment over an extended region in the apex of gerbil cochleae, and compared their phases to the basilar membrane (BM) response. By combining data from multiple tonotopic recording locations, low-frequency traveling waves that propagated along the BM with frequency dependent wavelengths between 1–3 mm were observed. In addition, we found a nonzero OHC−BM phase difference which systematically depended on the angle between the OCT optical beam and the BM along its longitudinal axis. This result establishes that the OHC vibrations were predominantly in the longitudinal direction, orthogonal to the up-down motion of the BM. Such motion is unavoidable in the presence of traveling, surface (BM) waves, but has such a significant magnitude that it masks the detection of OHC vibrations within the cross-sectional plane. With this, interpreting relative ooC motions in terms of the cochlear micromechanisms underlying our ear's remarkable sensitivity and frequency selectivity is tentative because their directions are not *a priori* known and vibrations along the third cochlear dimension must be considered as an important contributor to the recorded response.

INTRODUCTION

Sounds entering near the base of the cochlea create a pressure gradient across the basilar membrane (BM) and the sensory organ of Corti (ooC) it supports. This results in a longitudinally propagating traveling wave (TW) that—due to systematic and gradual changes in several physical properties (e.g., BM stiffness and width)—propagates with an amplitude and speed that depends on frequency, level and intracochlear location. This results in a place-based spectral decomposition of sound, or tonotopic organization, in which high frequencies maximally vibrate the cochlear base and low frequencies the apex. It is believed that active processes associated with outer hair cells (OHC) augment the passive, mechanical intracochlear responses to endow our ears with their capability to accurately discriminate tonal frequencies and intensities while maintaining exquisite sensitivity over a broad spectral and amplitude range [1]. Measuring the TWs, and the effects of OHC responses is challenging: the involved cochlear structures are small and poorly accessible, and their responses are on a (sub-)nanometer scale and vulnerable to physiological insults [2]. Because of this, most direct measurements of sound-evoked intracochlear vibrations were from superficial structures (e.g., basilar membrane; BM) at select locations within each cochlea. Vibrometry that uses spectral-domain optical coherence tomography (SD-OCT) alleviates many of these challenges [3, 4], and its recent application to cochlear mechanics has revealed several unanticipated and complex motions within the ooC, even at a single tonotopic location within the cochlea (e.g., [5, 6, 7, 8]). Here we present sound-evoked vibrations that were recorded using SD-OCT from multiple, adjacent longitudinal locations in the low-frequency, second turn of the gerbil cochlea that allowed us to evaluate the responses of individual low-frequency components across multiple tonotopic locations. This showed TWs along the BM with frequency-dependent wavelengths between 1 and 3 μm that slowed down when approaching the tone's characteristic frequency site along the tonotopic axis. Moreover, the natural curvature of the cochlea varied the (longitudinal) angle between the BM and the OCT optical beam across tonotopic locations. This revealed a drastic and sudden phase reversal of the OHC responses, which we interpret as evidence for substantial longitudinal motions (i.e. parallel to the BM) that "mask" radial vibrations (orthogonal to the BM) in the measured response. Responses in the lateral compartment (Hensen's cells and tectal cells) do not exhibit this phase-flipping behavior, suggesting that the large longitudinal motion is most prominent in the ooC-region occupied by the OHCs and/or Deiter cells.

MATERIALS AND METHODS

ANIMAL PREPARATION. The care and use of animals were in accordance with guidelines of, and approved by, the Institutional Animal Care and Use Committee (IACUC) of the VA Loma Linda Healthcare System. Cochlear vibrations were measured from the left ear in adult, female Mongolian gerbils (*M. unguiculatus*, n=3) that were part of a larger study group. Animals were anesthetized using intraperitoneal injections of a ketamine/xylazine (80 and 10 mg/kg, respectively) cocktail. Supplemental doses were administered to maintain areflexia. Core temperature was kept at $\sim 38^{\circ}$C. Animals were tracheotomized, but not actively ventilated. A ventrolateral surgical approach was used to expose the cochlea. The pinna and the cartilaginous ear canal were resected, and a probe containing a microphone and transducer assembly (ER-10X, Etymotic Research) was placed within a few mm of the tympanic membrane. Animals were not allowed to recover from anesthesia and were euthanized by anesthetic overdose at the end of the experiment.

OPTICAL COHERENCE TOMOGRAPHY AND VIBROMETRY. A spectral domain OCT system (Thorlabs Telesto III TEL321C1 (central wavelength: 1310 nm; bandwidth: 170 nm) with an LSM04 objective) was used to visualize and record the vibratory responses in the middle turn of the intact cochlea. Acquisition of optical spectra was triggered externally (27.9 pulses/ms) and phase-locked to the stimulus generation and microphone acquisition system (RX6: Tucker Davies Technologies system III). Optical spectra were converted into depth-resolved, axial information (A-line) using Fourier analysis. Intensity images (B-scans) were constructed from multiple A-lines by scanning the OCT optical beam. For vibration measurements (M-scans), the OCT beam was at a fixed position and a series of time-stamped A-lines was recorded while an acoustic stimulus was presented to the ear.

ACOUSTIC STIMULATION AND ANALYSIS. Stimuli consisted of 42 equal-intensity frequency components, each one with a random starting phase. Sound pressure levels were calibrated *in situ*. Stimulus frequencies were "zwuis" (e.g., [9]) and chosen such that they all had an integer number of cycles over 334,821 samples ($\sim$ 12 s). Typically, 3–5 of these concatenated periodic stimulus blocks were presented. Response waveforms were averaged across these periodic blocks, without any artifact rejection, and the amplitude and phase at the stimulus frequencies were extracted using Fourier analysis. The response at a frequency component was considered above noise when a Rayleigh's test for uniformity indicated significant ($p \leq 0.001$) phase locking to the stimulus [10]. In each longitudinal OCT image, a tilted ellipse was fitted to the boundary of the lateral compartment. The angle between vertical and the normal of this ellipse served as an estimate for the angle between the OCT beam and the BM in the longitudinal direction (termed "viewing angle" α throughout the manuscript). Frequency-dependent wavelength was calculated by fitting a straight line to the phase-vs-position data at each stimulus frequency using least-square minimization. The estimate was excluded when $r^2 < 0.8$. Custom software, written in MATLAB, was used for stimulus generation, hardware synchronization, and all signal analysis; control of the OCT system was implemented in C$^\sharp$.

RESULTS

With our OCT system we visualized a 0.6–0.8 mm stretch along the longitudinal axis of the second-turn gerbil cochlea (Fig. 1A,B). The cochlea coiled appreciably along this section, resulting in considerable variation in the OCT viewing angle relative to the longitudinal axis of the BM (between -23° and $+19^{\circ}$ across animals; Fig. 1B, inset). We obtained orthogonal images at multiple locations along this axis to visualize the cross-section of the cochlea (Figs. 1C,D). In each of these images, sound-evoked vibrations were recorded from four locations that represent responses of the BM, the OHC, the Hensen's cells, and tectal cells, respectively. Since tonotopic location varied between these cross sections, the tuned frequency response curves (FTCs) of the BM (Fig. 2A) had best frequencies (BFs) that systematically varied: from 2.8 kHz to 2.0 kHz at the most basal and apical locations, respectively (Fig. 2C). By fitting a cochlear place-frequency map obtained from gerbil auditory nerve fiber responses [11] to these BM-BFs, we calculated that the recording locations were between 6.5 and 7.3 mm from the base of the cochlea. For each BM-FTC, the phase (Fig. 2B) accumulated with frequency in a manner characteristic for a longitudinally propagating TW that slowed down when approaching the location's BF. We calculated wavelengths (λ) from BM phase responses across the tonotopic recording locations (Fig. 2D) and found them to vary with frequency between $\lambda \approx 1$ mm near BF and $\lambda \approx 2$–3.5 mm below BF. These values are similar to previous wavelength estimates from more basally obtained cochlear-mechanical [2, 12, 13] or auditory-nerve [14] data. Sound-induced motions of the ooC at one tonotopic location are not uniform in the *in vivo* cochlea. Rather, different parts move at different amplitudes and phases, which results in deformations of the ooC [5, 8, 15]. We systematically compared the response phases of OHC, tectal cells and Hensen's cell to those of the BM (Fig. 3). As previously reported, we found that OHC responses were not in

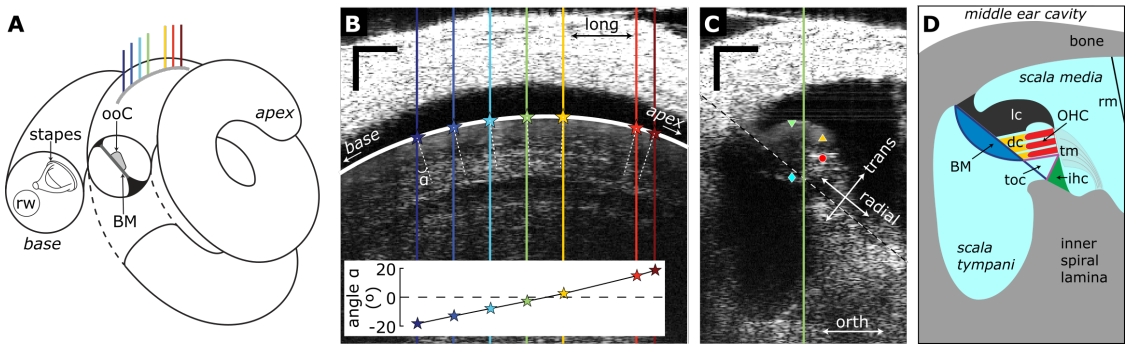


FIGURE 1. Longitudinal recording locations change viewing angle in the 2nd turn of the cochlea. **(A)** Schematic diagram of the spiral gerbil cochlea with a cross-sectional plane exposed to show the BM and ooC. The gray line with colored lines represents the image location shown in (B). Colored lines indicate beam positions used for vibrometry. rw: round window. **(B)** Intensity image along the longitudinal axis with part of the ellipse fitted to the boundary of the lateral compartment (white line). Inset: viewing angles of the OCT beam at the different recording locations. Scale bars, 0.1 mm. **(C)** Cross-sectional intensity image, orthogonal to the longitudinal axis, obtained at the green line in (B) with the recording locations for BM (cyan ◊), OHC (red ◯), Hensen's cell (green ▽) and tectal cell (yellow △) indicated. The green line locates the intersection with the longitudinal plane, the dashed black/white line indicates the BM. View into the paper is towards cochlear apex. Scale bars, 0.1 mm. **(D)** Schematic of the anatomical structures for the cross section shown in (C). Fluid-filled spaces within the cochlea are light blue. dc: Deiters cells, ihc: inner hair cell, lc: lateral compartment, rm: Reissner's membrane, tm: tectorial membrane, toc: tunnel of Corti.

phase with the BM (Fig. 3A). Moreover, the OHC−BM phase difference showed a drastic dependence on the viewing angle between the OCT optical path and the BM, as can be inferred from the color-coding of the lines in Fig. 3A.

The drastic dependence of the OHC−BM phase difference on viewing angle, and its 180° reversal, is further illustrated in Fig. 3D. These data exhibited a robust trend (Fig. 3D): for negative viewing angles, corresponding to the higher-CF tonotopic locations, the OHC responses led the BM by ~ 0.25 cycles, which abruptly transitioned to a ~ 0.25-cycle phase lag when the viewing angles were positive. Only two observations (vertical arrows in Fig. 3A,D), both of which were obtained at two small viewing angles ($-1.8°$ and $-2.8°$, respectively), deviated from the overall trend. These may be explained when accepting an error of a few degrees in the viewing angle between the OCT and the BM, which was calculated from an ellipse fitted to the lateral compartment boundary. This dependence of relative phase on viewing angle, including the 180° change at $\alpha = 0°$, is predicted (Fig. 3E, see also [5]) if the OHC region exhibit significant longitudinal motion in response to sound.

The relative phase of the responses in the lateral compartment, i.e., the tectal cells (Fig. 3B) and Hensen's cells (Fig. 3C), did not show this strong dependence on viewing angle, although there is a tendency for the phase responses

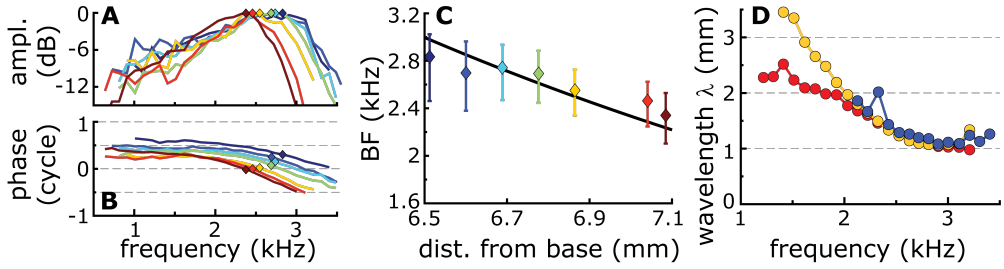


FIGURE 2. Vibratory responses along the tonotopic axis of the cochlea. **(A)** Normalized (re. max) amplitude of BM vibratory frequency response curves (FTCs) measured at different longitudinal locations in the gerbil cochlea with best frequency (BF) indicated (◊). Colors correspond to the color coding in Fig. 1B. **(B)** Corresponding phase (re. middle ear). An overall 0.6-ms delay was removed from these phase curves to emphasize the dependence on frequency. **(C)** Best frequency for BM responses determined from the FTC amplitude curves. Error bars give FTC-bandwidth at 1-dB below the BF response. The BM-BF's were fit (black line) with an gerbil place-frequency map [11] to determine distances from base for the recording locations. These were used to set the abscissa. **(D)** Frequency dependent wavelength (λ) of the TW, calculated across the measured cochlear extent. Colors represent the three animals and are also used in Fig. 3D.

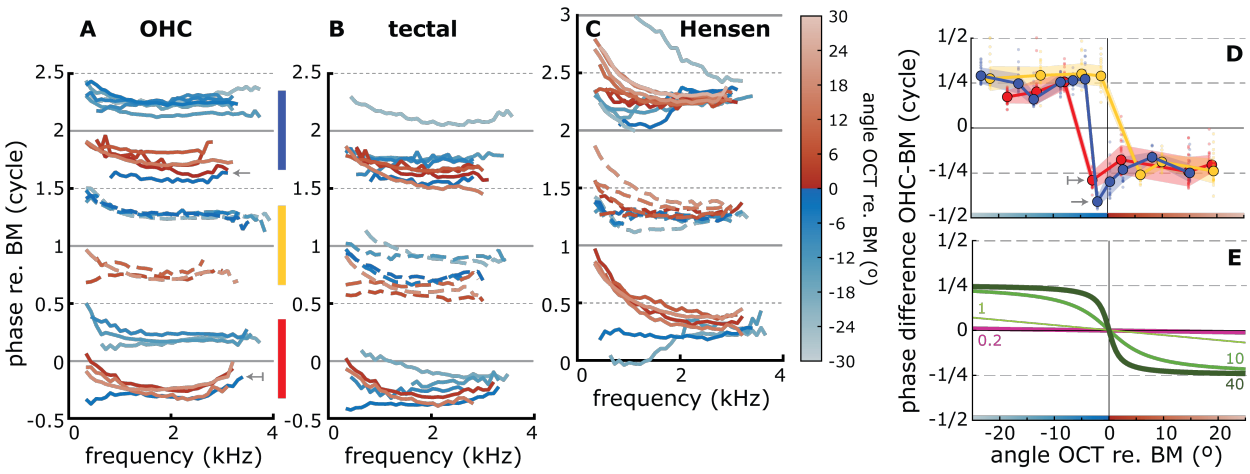


FIGURE 3. Phase of ooC motion relative to the BM response. Response phase for **(A)** OHC, **(B)** tectal cells, and **(C)** Hensen's cells, relative to the phase of the BM. Datasets for all three animals are shown, but they are vertically offset from each other by one cycle. The color of the lines encodes the viewing angle between the OCT beam and the BM in the longitudinal direction with the key provided in (C). **(D)** OHC−BM phase difference, averaged across frequency, as a function of OCT viewing angle. Small colored circles give individual data points for each frequency, shaded areas show ±1 s.d. around the mean. Colors distinguish the three datasets in (A), where they are indicated by the vertical colored bars. The same three colors were used in Fig. 2D. **(E)** Expected phase differences for elliptical motion with different aspect ratio, here defined as the horizontal vibration amplitude divided by the transverse vibration amplitude. This amplitude ratio is given next to each curve. For ratios>1 (i.e., a larger longitudinal component, green lines), the phase difference systematically varies between +0.25 and −0.25 cycles. The 0.5-cycle transition occurs more abruptly for larger ratios. When the transverse motion is larger (ratio<1, magenta line), only a small phase difference exists over this range of viewing angles. Irrespective of amplitude ratio, the phase difference changes sign when the viewing angle = 0°.

observed at the most negative angles (e.g., light blue curves in Fig. 3B) to deviate from the average response. Also, there is a clear trend for the phase difference to change for lower stimulus frequencies, most obvious in the Hensen's cell region, especially for the more apical (red lines) recording locations. This trend may be related to non-perfect cross-sectional imaging in combination with the dispersive nature of the TW. The former may cause a small mismatch between the BF of the BM and the OHC or lateral compartment, while the latter results in a less pronounced phase difference between BM and other ooC structures because sub-BF frequencies propagate fast.

DISCUSSION

In this manuscript we presented data on the relative phase of sound-evoked vibratory responses of different regions within the ooC that were obtained from multiple locations along a 600-800 μm extent of the tonotopic axis in the second, low-frequency turn of the gerbil cochlea. The most striking observation is that the OHC−BM phase difference strongly depends on the the angle between the BM along the longitudinal axis of the cochlea and the incident OCT light beam. Since this viewing angle and the cochlear tonotopic location co-varied, changes in the anatomy or physiology along the longitudinal axes may also potentially cause the observed effect. This, however, it not likely because the phase reversal occurred over an extent of only ∼100 μm, and changes in BM vibrations were small and gradual over this range (e.g., Figs. 2A,B). In addition, the drastic phase reversal would require a sudden change in the local anatomical and/or electro-mechanical ooC properties, which is not observed [16, 17, 18, 19]. Finally, the same phase reversal was observed in all three animals, despite differences in anatomy, experimental conditions and exact BF locations. Rather, the observed phase reversal can readily be explained by considering displacements in the longitudinal direction that lag the transverse vibrations by 0.25 cycle, a scenario encountered in many types of waves [20]. OCT measures the projection of the actual motion onto its optical beam. When the motion contains both longitudinal and non-longitudinal components, their relative contribution to the measurement depends on their relative magnitudes and the OCT viewing angle. A simple geometric relation [5] describes how the OHC−BM phase difference depends on both the viewing angle and the amplitude ratio of the elliptical motion (Fig. 3E). Accepting a few degrees of uncertainty in determining the viewing angles, the data (Fig. 3D) are qualitatively and quantitatively

well described by a situation in which longitudinal motion dominates the recorded OHC vibrations, thus masking any contributions from motion within the cross-sectional (transverse-radial) plane to the measured response. Although a

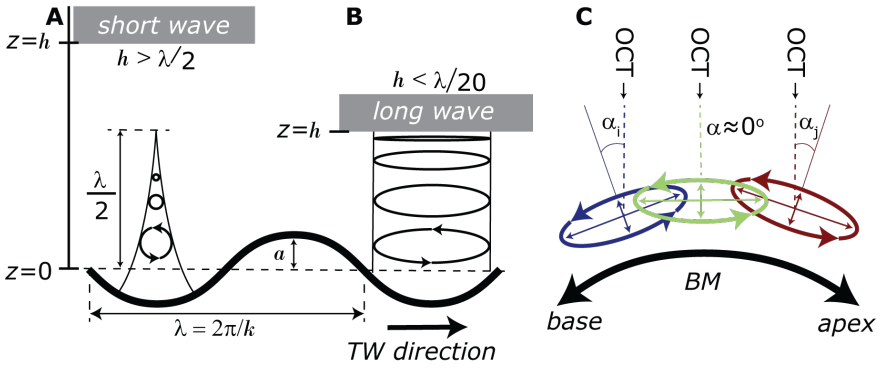


FIGURE 4. . **Elliptical motion associated with a surface wave. (A–B)** A surface wave that travels from left to right (thick black line) is associated with elliptical trajectories of fluid particles. For each cycle of the TW, ellipses are traversed once in the counterclockwise direction. Their shape depends on the depth of the wave medium (h) relative to the TW wavelength (λ) [20]. For short waves (A), the amplitudes of the longitudinal and transverse motion are identical (corresponding to an amplitude ratio = 1 in Fig. 3E), resulting in circular motion. The diameter of each circle decreases exponentially with depth into the wave medium, and at z=$\lambda/2$ is less than 5% of the TW amplitude. For long waves (B), the ellipses are substantially flattened in the transverse direction (i.e., amplitude ratio >1). Moreover, the amplitude of the longitudinal motion decays little with distance away from TW-supporting surface, while the transverse amplitude reduces linearly with depth. This results in an increasing amplitude ratio for distances furhter away from the surface wave. **(C)** Due to the curvature of the cochlea/BM in the longitudinal direction (thick black line), a vertical OCT measurement beam (dashed lines) will view the elliptical motion at an angle ($\alpha\ldots$) that varies with tonotopic location (different colors).

description of the cochlea is undoubtedly more complex due to the presence of the ooC with its internal (anisotropic) structures, it is informative to consider the hydrodynamics of waves on water, i.e., traveling waves on the surface of a homogeneous fluid layer plane ([20, 21, 22], see also Figs. 4A,B). In linear wave theory, the time varying velocity potential (Φ) in the wave medium can be derived from the Laplace equation with the assumptions that the vertical velocity at the "wall" is zero and that particles at the surface of the wave remain there. That is,

$$\Phi = ga\frac{T}{4\pi}\frac{\cosh\{k(h+z)\}}{\cosh kh}\sin\left(kx-\frac{2\pi t}{T}\right) \tag{1}$$

with TW amplitude a, wavelength λ, period T and wavenumber $k=2\pi/\lambda$. The horizontal and vertical displacements at each depth in the wave medium are then given by the partial derivatives:

$$\text{horizontal displacement} = \frac{\partial\Phi}{\partial x} = -\frac{a}{2}\left[\frac{\cosh\{k(z+h)\}}{\sinh kh}\right]\sin\left(kx-\frac{2\pi t}{T}\right) \tag{2}$$

$$\text{vertical displacement} = \frac{\partial\Phi}{\partial z} = \underbrace{\frac{a}{2}\left[\frac{\sinh\{k(z+h)\}}{\sinh kh}\right]}_{\text{amplitude term}}\underbrace{\cos\left(kx-\frac{2\pi t}{T}\right)}_{\text{phase term}} \tag{3}$$

For a full derivation see for example [23]. The difference in the phase terms in Eqs. 2, 3 make that the horizontal displacement (i.e., along the longitudinal cochlear axis) lags the vertical displacement (transverse motion, orthogonal to the BM) by 0.25 cycle, and particles move along ellipses in the counterclockwise direction. The ratio of the amplitude terms, both of which depend on wavelength λ and fluid depth h, determines the shape of these ellipses: they vary from circular when $\lambda < 2h$ (short wave, Fig. 4A) to extremely flattened in the longitudinal direction when $\lambda \gg h$ (long wave, Fig. 4B). The rapid transition in the observed OHC−BM phase difference (Figs. 3D,E) with viewing angle suggests that the OHC motion follows ellipses that are relatively flat, and that the cochlea thus operates in the long-wave regime in which the TW wavelength is at least 20× larger than the effective depth of the wave medium. Wavelengths of 1–3 mm (Fig. 2C) would correspond to an effective cochlear fluid depth of 50–175 μm.

This is on the same order as the ooC dimensions, suggesting that perhaps the tectorial membrane or reticular lamina serve as the fixed "wall" for the wave medium rather than the bony walls of the osseous labyrinth. Irrespective of the origin of the longitudinal motion, the data show that it far exceeded the combined transverse-radial vibrations within the OHC region of the ooC.

In contrast to the OHC vibrations, the relative phase responses in the lateral compartment did not show such a strong dependence on viewing angle (Figs. 3B,C). This is probably related to the its anatomy compared to the OHC/Deiters cell region. That is, the lateral compartment is a cellular structure where cells are tightly connected [24], while OHC are surrounded by cochlear fluid rather than abutting cells. The latter configuration likely results in fewer restrictions on their longitudinal motion. We did observe, especially in the tectal cell region (Fig. 3B), that the relative phase started to change for the most negative viewing angles (light blue lines). This agrees with an amplitude ratio for the elliptical motion that is closer to one (i.e. the amplitude of the longitudinal motion becomes smaller, approaching the amplitude of the orthogonal component), in which the relatively smaller longitudinal motion only starts to appreciable effect the measured response when its projection onto the OCT optical beam is relatively large.

Longitudinal motion has been observed in excised cochleae [25], with indirect evidence from recent in vivo experiments [5]. Without knowing the motion directions, any interpretation of observed complex motion patterns within the ooC becomes tentative. Measured amplitude differences may reflect differences in motion direction rather than in true amplitude along the OCT optical path, and the same holds for the (relative) phase. For example, amplitude and/or phase differences in the responses of two points (e.g., the two ends of an OHC) may be interpreted either in terms of length changes or as a rotation of the structure. Without knowing the directions of the measured vibrations it is impossible to decide between the two. Although this makes interpretation of the complex (relative) OHC motions more difficult and/or speculative, it may explain why, from all the different intracochlear vibrations that have been documented, the classical up-down BM vibrations provide the best description of the tuned auditory nerve fiber (ANF) response [26, 27]. Apparently, the complex ooC vibrations have little effect on the transmission of BM vibrations to the bundle of the inner hair cell, perhaps because they are the result of motions in the "wrong" (longitudinal) direction that do not optimally deflect the bundle. Clearly, an accurate description of the intricate mechano-electrical workings of the auditory sensory system is far from complete, but SD-OCT is (starting to) providing the data necessary to understand the processing and transduction of sounds within the cochlea.

ACKNOWLEDGMENTS

This work was supported by NIH/NIDCD R21 DC019998 to W.D. These contents do not represent the views of the U.S. Department of Veterans Affairs or the United States Government.

REFERENCES

1. W. A. Yost, *Fundamentals of hearing: an introduction*, 2nd ed. (Holt Rinehart and Winston, 1985).
2. L. Robles and M. A. Ruggero, "Mechanics of the mammalian cochlea," Physiol. Rev. **81**, 1305–1352 (2001).
3. N. Choudhury, G. Song, F. Chen, S. Matthews, T. Tschinkel, J. Zheng, S. L. Jacques, and A. L. Nuttall, "Low coherence interferometry of the cochlear partition," Hear. Res. **220**, 1–9 (2006).
4. R. K. Wang and A. L. Nuttall, "Phase-sensitive optical coherence tomography imaging of the tissue motion within the organ of Corti at a subnanometer scale: a preliminary study," J. Biomed. Opt. **15**, 056005 (2010).
5. N. P. Cooper, A. Vavakou, and M. van der Heijden, "Vibration hotspots reveal longitudinal funneling of sound-evoked motion in the mammalian cochlea," Nat. Commun. **9**, 3054 (2018).
6. J. B. Dewey, B. E. Applegate, and J. S. Oghalai, "Amplification and suppression of traveling waves along the mouse organ of Corti: evidence for spatial variation in the longitudinal coupling of outer hair cell-generated forces," J. Neurosci. **39**, 1805–1816 (2019).
7. T. Ren, W. He, and D. Kemp, "Reticular lamina and basilar membrane vibrations in living mouse cochleae," Proc. Natl. Acad. Sci. **113**, 9910–9915 (2016).
8. H. Y. Lee, P. D. Raphael, A. Xia, J. Kim, N. Grillet, B. E. Applegate, A. K. Bowden, and J. S. Oghalai, "Two-dimensional cochlear micromechanics measured in vivo demonstrate radial tuning within the mouse organ of Corti," J. Neurosci. **36**, 8160–8173 (2016).
9. S. W. F. Meenderink and M. van der Heijden, "Reverse cochlear propagation in the intact cochlea of the gerbil: evidence for slow traveling waves," J. Neurophysiol. **103**, 1448–1455 (2010).
10. C. P. C. Versteegh and M. van der Heijden, "Basilar membrane responses to tones and tone complexes: nonlinear effects of stimulus intensity," J. Assoc. Res. Otolaryngol. **13**, 785–798 (2012).
11. M. Müller, "The cochlear place-frequency map of the adult and developing mongolian gerbil," Hear. Res. **94**, 148–156 (1996).
12. O. de La Rochefoucauld and E. S. Olson, "The role of organ of Corti mass in passive cochlear tuning," Biophys. J. **93**, 3434–3450 (2007).
13. W. S. Rhode and A. Recio, "Study of mechanical motions in the basal region of the chinchilla cochlea," J. Acoust. Soc. Am. **107**, 3317–3332 (2000).

14. M. van der Heijden and P. X. Joris, "Panoramic measurements of the apex of the cochlea," J. Neurosci. **26**, 11462–11473 (2006).
15. J. B. Dewey, A. Altoè, C. A. Shera, B. E. Applegate, and J. S. Oghalai, "Cochlear outer hair cell electromotility enhances organ of Corti motion on a cycle-by-cycle basis at high frequencies in vivo," Proc. Natl. Acad. Sci. **118**, e2025206118 (2021).
16. G. Emadi, C.-P. Richter, and P. Dallos, "Stiffness of the gerbil basilar membrane: radial and longitudinal variations," J. Neurophysiol. **91**, 474–488 (2004).
17. S. W. F. Meenderink, C. A. Shera, M. D. Valero, M. C. Liberman, and C. Abdala, "Morphological immaturity of the neonatal organ of Corti and associated structures in humans," J. Assoc. Res. Otolaryngol. **20**, 461–474 (2019).
18. R. Pujol, M. Lenoir, S. Ladrech, F. Tribillac, and G. Rebillard, "Correlation between the length of outer hair cells and the frequency coding of the cochlea," in *Auditory Physiology and Perception* (Elsevier, 1992) pp. 45–52.
19. D. E. Zetes, J. A. Tolomeo, and M. C. Holley, "Structure and mechanics of supporting cells in the guinea pig organ of Corti," PLoS One **7**, e49338 (2012).
20. M. W. Dingemans, *Water wave propagation over uneven bottoms, part 1. Linear wave propagation*, Vol. 13 (World Scientific, 1997) pp. 1–500.
21. J. Lighthill, *Waves in Fluids* (Cambridge University Press, 1978).
22. J. Lighthill, "Energy flow in the cochlea," Journal of Fluid Mechanics **106**, 149–213 (1981).
23. R. G. Dean and R. A. Dalrymple, *Water Wave Mechanics for Engineers and Scientists*, Vol. 2 (World Scientific, 1991).
24. J. J. Zwisilocki, N. B. Slepecky, L. K. Cefaratti, and R. L. Smith, "Ionic coupling among cells in the organ of Corti," Hear. Res. **57**, 175–194 (1992).
25. K. Karavitaki and D. C. Mountain, "Imaging electrically evoked micromechanical motion within the organ of Corti of the excised gerbil cochlea," Biophys. J. **92**, 3294–3316 (2007).
26. S. S. Narayan, A. N. Temchin, A. Recio, and M. A. Ruggero, "Frequency tuning of basilar membrane and auditory nerve fibers in the same cochleae," Science **282**, 1882–1884 (1998).
27. M. A. Ruggero, S. S. Narayan, A. N. Temchin, and A. Recio, "Mechanical bases of frequency tuning and neural excitation at the base of the cochlea: comparison of basilar-membrane vibrations and auditory-nerve-fiber responses in chinchilla," Proc. Natl. Acad. Sci. **97**, 11744–11750 (2000).

Two operating modes for outer hair cells and implications for cochlear tuning.

Jonathan Ashmore,
UCL Ear Institute and Dept of Neuroscience, Physiology and Pharmacology London WC1E 6BT, UK
j.ashmore@ucl.ac.uk

Abstract. The role of cochlear outer hair cells (OHCs) in mammal hearing is compromised by the 'RC time constant problem'. The issue arises because if the cells are to operate at high acoustic frequencies conventional voltage driven 'electromotility' is low pass filtered by the cell membrane. By synthesising a description of the OHC as a piezoelectric actuator and its position between resonant basilar and tectorial membranes it is shown that sharp tuning can arise at high frequency if prestin/SLC26A5 is dynamically tensioned. The model predicts an imaginary (dissipative) component to the OHC capacitance as well as two distinct frequency tuning curve (ftc) shapes: a near symmetrical ftc at low CF, and a low frequency 'tail' on the more sharply tuned high CF ftcs.

Introduction

The outer hair cells (OHCs) of the mammalian cochlea play an essential role amplifying incoming sound and making normal hearing possible. The evidence stems from a variety of different experiments but most significantly from the observation that OHCs elongate and shorten when their membrane potential is changed, from the cells' position in the organ of Corti and from evidence that the gene for prestin/SLC26A5 determines cochlear sensitivity (reviewed for example in [1]). The role of outer hair cells and the requirement for an actuator molecule prestin/SLC26A5 thus seems sufficient for cochlear performance over the entire acoustic range. Nevertheless direct measurement from OHC movement both *in vitro* [2] and *in vivo* from basilar membrane measurements of gerbil OHCs [3] highlights a long term concern that the OHC membrane acts like a lowpass filter of voltage, a problem which has come to be known as the 'RC-time constant problem' [4]. One escape from the paradox is to propose that the OHC RC time constant, τ, continues to decrease towards the basal end of the cochlea and so mitigate the effect, at least for time scales imposed by patch clamp recording [5].

A line of modelling studies of single OHCs, based on the properties of prestin's non-linear capacitance, suggest that OHCs are capable of injecting power at the higher end of the acoustic range [6] [7] [8]. The ideas owe much to the suggestion that mechanical loading extends the OHC bandwidth [9].The present work combines prestin-dependent properties of the OHCs with a simple 1D model of cochlear micromechanics to investigate how such a scheme might enhance cochlear tuning. In the present case, the varying geometry of the organ of Corti along the cochlea and the precise configuration of hair cell excitation will, to first approximation, be neglected. The inclusion of such geometric factors has been considered elsewhere [10] in a model on which some the present considerations are based**.**

Formulation of the problem

The governing set of equations for the 1-D amplitude motion, ξ, of the basilar membrane (BM) is given by a set of coupled oscillators embedded in fluid,

$$\sum_j \left(G_{ij} + m\delta_{ij} \right) \ddot{\xi}_i + \; h_i \dot{\xi}_i + \; k_i \xi_i = f_i(\xi, t) \qquad (1)$$

represented here by the discretised version (where the dots signify time derivative). The BM amplitude ξ is adopted to distinguish it from longitudinal cochlear position, x, indexed as i. The term f_i (ξ,t), is the local forcing term at position i. The parameter k = k(x) is the stiffness of the partition at each point x along the cochlea, h=h(x) is a viscosity parameter, accounting for both the local fluid viscosity in scala media as well as that arising in the subtectorial space.

Although m=m(x) in Eqn (1) is described conventionally as a partition mass, the additional term G(x,x') is included as non-local fluid mass term and performs the function of fluid coupling. It is the discretised hydrodynamic Green's function G(x,x') for the problem and can be calculated from a knowledge of the dimensions of the cochlear fluid volumes [11], [12]. For simplicity the total point mass will be termed m in what follows. The problem term in Eqn (1) is the damping h_i dξ/dt and leads to dissipative energy loss, reducing sharp tuning of the BM. It includes any viscosity contributions from cells of the organ of Corti as well as the viscous forces at the interface between the TM and the reticular lamina. OHCs enter the description by the addition of a second forcing term to the right hand side, an 'undamping' term $f_{i,OHC}$ [13].

$$f_i(\xi, t) \rightarrow f_i(\xi, t) + f_{i,OHC}(\xi, t)$$

Cochlear micromechanics with OHC feedback

To calculate f_{OHC} at each point when the OHCs are part of a feedback loop it is necessary to include the effects of the forces arising from the OHCs and any response of prestin/SLC26A5. A simple feedback scheme of the OHCs sandwiched between a resonant tectorial membrane tectorial (TM) and a resonant basilar membrane (BM) (Figure 1) will be used and closely follows the treatment elsewhere [10]

Fig 1A shows a point hair cell model with two resonant structures. Both resonant structures, the TM and the BM are referenced to the modiolus. The equations for the movement, ξ, of the BM, and y for the TM are thus (dropping the place indices) are

$$H_{BM}\xi = f_s - f_{OHC} \quad ; \quad H_{TM}y + f_{OHC} = 0 \qquad (2a,b)$$

where f_s is the stapes forcing term applied as a result of sound entering at the basal end, and f_{OHC} is the combined force due to OHCs and the organ of Corti. In Fourier space, with $\omega = 2\pi f$ and $j = \sqrt{-1}$, the operators are: $H_\alpha(\omega) = k_\alpha\left(1 + j\,\bar{\omega}_\alpha/Q_\alpha - \bar{\omega_\alpha}^2\right)$ where $\bar{\omega}_\alpha = \omega\sqrt{m_\alpha/k_\alpha}$; and $Q_\alpha = \sqrt{\frac{m_\alpha}{h_\alpha k_\alpha}}$ are best frequency and quality factors respectively for α = BM or TM.

If the OHC electromotility is absent the system is purely passive; the organ of Corti itself couples the TM and the BM with an effective force $Z_{OHC}(y - \xi)$ where Z_{OHC} is the impedance of the OHC including the organ of Corti. In this case Eqn 2 shows that the BM will experience a force of

$$f_{OHC} = -H_{TM}\,y = -\frac{H_{TM}\,Z_{OHC}}{(H_{TM} + Z_{OHC})}\xi \qquad (3)$$

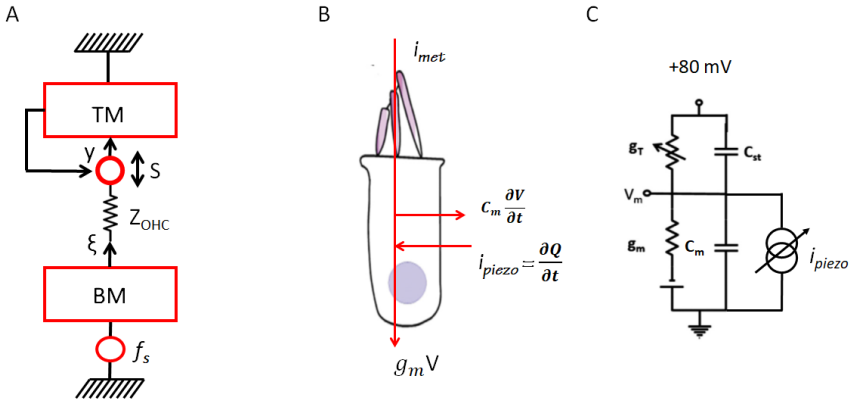


Figure 1: A, a 1-D model with the OHC driven by feedback from the TM. Note the direction in which the BM and TM directions are defined relative to the modiolar reference points. B, model of the OHC: transducer current $i_{met,}$, C_m the membrane capacitance. i_{piezo} the displacement current produced by OHC tension, arising from anion and intrinsic charge movement associated with prestin/SLC26A5. C, the equivalent circuit with i_{piezo} appearing as a tension dependent current in addition to the current through the m.e.t. conductance g_T.

Both the OHCs and the TM will behave like mechanical impedances in series. In general the force f_{OHC} will be a function of frequency and the BM mechanics will reflect loading by the TM even in this passive case.

OHCs as piezoelectric actuators

The OHC can be described as 1D piezoelectric device expanding its length by two (linearised) equations [7]:

$$S = c\delta T + d_{31}\,\delta V\,; \quad \delta Q = d_{13}\delta T + \varepsilon\delta V \tag{4a,b}$$

where S (=δL/L) is the incremental length (or strain), and δQ is the charge displaced when the device is subject to increments in the tension T and voltage V around the equilibrium. The voltage dependent parameters d_{31} and d_{13} characterise the piezoelectricity (the subscripts indicating the tensor nature of the parameters). The parameter c is the compliance (inverse stiffness) of the cell. The equations can be seen as modifications of Hooke's Law and charge storage by a capacitance ε , both terms of which can be absorbed into the passive mechanics before any OHC feedback is included.

Hair cell transduction

Deflection of the hair bundle activates a transduction current i_{met} which drives the membrane voltage (Fig 1B,C)

$$i_{met} = g_m V + C_m \frac{dV}{dt} + i_{piezo} \tag{5}$$

The first two terms $g_m V$ and C_m dV/dt are the conventional currents flowing through the membrane conductance and the total membrane capacitance (Fig 1B,C). The conventional RC time constant is $\tau = C_m/g_m$ An additional term is included in Eqn 5: $i_{piezo} = \frac{\partial Q}{\partial t}$ representing the current arising in OHCs due to charge movement when prestin/SLC26A5 is deformed. This displacement current can be observed experimentally [7], [14]; [15]. The transducer current, i_{met}, is a nonlinear function of displacement but can be linearised for small signals as $i_{met} = -\beta y$ in the convention of Fig 1A

As the OHC changes length, some of the displacement change is taken up by the strain change S of the cell. The force produced by the cells between the BM and TM will therefore be Z_{OHC} (y-ξ – S). Other (passive) mechanical components of the organ of Corti, in parallel with the OHCs, can be absorbed into the OHC spring impedance Z_{OHC}. As a result, algebraic combination of Eqns 4 and 5 in the Fourier domain allows the force f_{ohc} to be written as the sum of two terms

$$f_{OHC} = -H_{TM}K\left(\frac{1+ j\omega\tau(1-K\gamma H_{TM})}{(\alpha+ j\omega\tau(1-K\gamma H\_TM)}\,\xi + \gamma H_{TM}K\,\frac{j\omega\tau(H_{BM}-KH_{TM})}{(\alpha+ j\omega\tau(1-K\gamma H_{TM})\ )}\xi\right. \tag{6}$$

where $\boldsymbol{\alpha} = 1-K\, d_{31}\beta/g_m$, $\boldsymbol{\gamma} = \frac{d_{31}d_{13}}{C_m}$ and $K = \frac{Z_{OHC}}{H_{TM}+Z_{OHC}}$.

The parameters d_{31}, d_{13} have been measured in guinea pig OHCs [7]; the m.e.t. slope β and the conductance g_m have been measured in gerbil [5] so that we estimate $d_{31}\,\beta/g_m \sim 5.3$ for a 25 μm OHC and thus $\alpha<0$. This implies that the force f_{OHC} is real and positive in the low frequency limit ($\omega\tau<<1$). The force is therefore acts as negative feedback and increases the stiffness of the BM at low frequencies.

The full expression for the force in Eqn 6 depends on γ, a coupling parameter that is the product of both OHC piezoelectric parameters, d_{31} and d_{13} Since both parameters are voltage dependent, the value of the coupling parameter γ can also depend on the membrane potential V of the OHC, opening the possibility of dynamic tuning of cochlear force feedback and sensitivity. It should be noted that the expression, Eqn 6, reduces to the TM loading of the BM (Eqn 3) when $d_{31}=0$ (no electromotility); when $d_{13}=0$ (i.e. no tension sensitivity of prestin) Eqn (6) reduces to the expression given by Geisler and Shan [16] [10].

Effective capacitance of OHCs

The first term in Eqn 6 exhibits an effective time constant $\tau' = \tau(1 - K\gamma\, H_{TM}) = R_m C_{eff}$. It depends on the piezo electric parameter γ and the OHC and TM impedances. If H_{TM} is real (i.e. no viscosity) then the time constant τ'can become arbitrarily small for suitable parameter choice or even negative [6]. If the TM has viscosity, then τ', and hence C_{eff} , can take on complex values as well. Such results have been reported for OHC macropatches [17].The results are shown in Figure 2.

Since $\gamma=d_{31}d_{13}/C_m$ depends on the square of the piezo parameters, C_{eff} measured when the OHC is loaded may have a stronger voltage dependence than that of the isolated cell's nonlinear capacitance. This conclusion can be experimentally investigated.

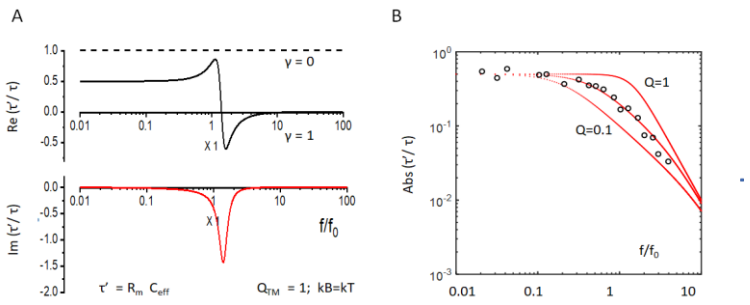


Figure2 Effective OHC time constant τ'. A, real and imaginary parts of the function $\tau' = \tau(1 - K\gamma\, H_{TM})$ *with* $k_{ohc} = k_{TM}$ *γ as shown .B, Absolute magnitude of effective capacitance* $C_{eff} = \tau'/R_m$. *circles, data from single OHC patches (Gale and Ashmore 1997) with* f_0 *= 12 kHz. The curve fits a resonant system* H_{TM} *with high damping (Q=0.5) compatible with data from OHC membrane patches..*

The second term in Eqn 9 vanishes when γ =0 (i.e. when i_{piezo} is not included) but depends explicitly on the BM impedance H_{BM}. Notably the term also depends on the product of the BM and TM stiffnesses and so can compensate for the increasing stiffness of BM towards the basal end of the cochlear end

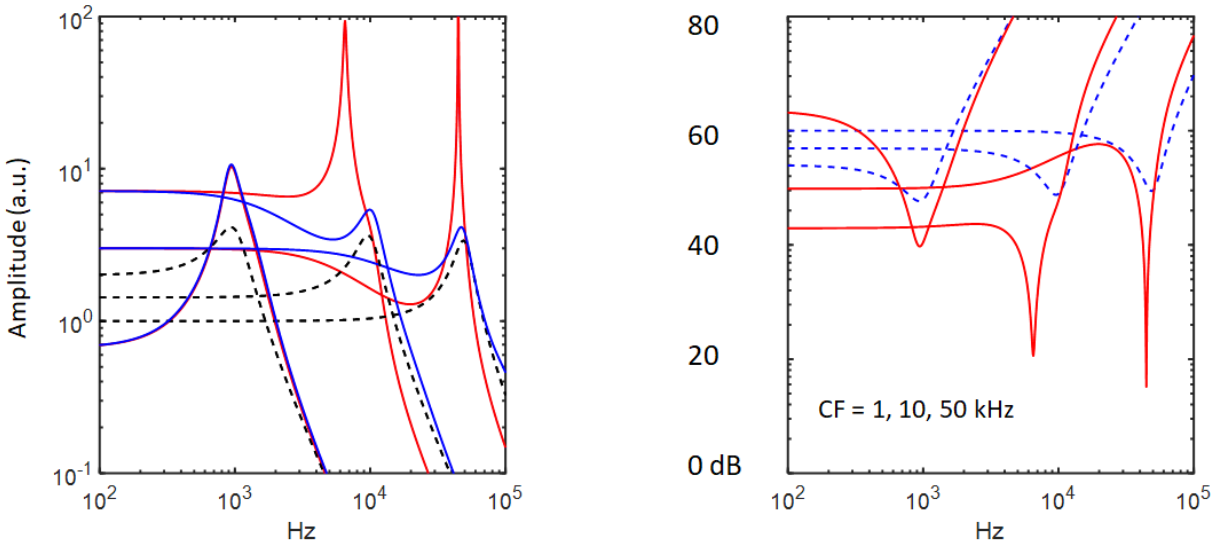


Figure 3: Indicative tuning curves for three cochlear positions inferred from Eqn 6. Left, BM tuning; Right inferred frequency tuning curves, with a scale given in dB. Black dashed line, the transfer function of the BM alone; red, transfer function of the BM with effects of f_{ohc} *(i.e. 1/(*$H_{BM} - f_{ohc}$*); blue, transfer function of BM with* f_{ohc}*, but with γ=0. The positions correspond to apical 1 kHz, mid-10 kHz and basal 50 kHz locations. The underlying tuning is established by the passive mechanics of the BM with only small changes in the mechanical* Q_{3dB} *of the BM. The range of parameters chosen were (kHz,* $d_{31}\beta/g_m$ *, γ): (1, 2.5, 0.8); (10, 4..5, 3,2); (50, 5, 3,3)*

Although the parameter space is large, to see how this term contributes to tuning, consider the simple case where both the BM and the TM are resonant structures with the same corner frequencies. The real part of the second in Eqn 6 term will vanish when Re $(H_{BM} - KH_{TM}) = 0$ or equivalently when the BM impedance is matched by the impedances of the TM and OHC in series. Hence we choose for simplicity $Z_{OHC} = k_{ohc}=k_{TM} =2\ k_{BM}$. Thus the impedance of the TM and OHC together match that of the BM at low frequencies. In this case f_{ohc} is purely imaginary and proportional to the negative viscous damping of the BM, with a correction for any resonant effects from the TM. The gain is determined by the piezo parameter γ. The cancellation may be imperfect for it will be modified by the presence of the non-local Green's function coupling term G(x,x') in Eqn 1.

With suitable choice of parameters, the effect of OHC enhancement over the passive BM tuning can reach 40 dB or more . Figure 3 shows the that characteristically sharp tuning curve can be obtained for a resonant TM. It can also be obtained for a non-resonant TM, the main difference being that the resonant TM generates a notch on the low frequency side of the high CF tuning curve. The mechanical power dissipation by this system of OHCs will be $\dot{\xi} f_{OHC}(\xi)$. A critical role is played by $\boldsymbol{\gamma} = d_{31}d_{13}\ /C_m$ which determines not only the mechanical undamping in Eqn 6, but also determines the effective imaginary capacitance of the OHC (i.e. the dissipative effects) though the expression for $\tau' = R_m C_{eff}$..

Discussion

Below the cut-off frequency determined by the cell membrane time constant, OHC 'electromotility' provides positive feedback; the above treatment shows that there is also a regime above the cut-off frequency where the charge flowing through the prestin, not just the m.e.t. channel, contributes to feedback.

Before the identification of prestin/SLC26A5 the models of OHC function depended upon a phenomenological fit to the data rather than a traceable molecular mechanism. OHC feedback was required to be specified up to a gain factor, adjusted to obtain agreement enhancement of the BM tuning [18]. Piezoelectric descriptions reduce that ambiguity by basing the feedback on experimentally determined cellular data. The piezoelectric OHC properties are energy dependent and the mechanical power dissipation at each point is $\dot{\xi}\ f_{OHC}(\xi)$. Physically this can be thought of as either viscous losses of the protein-lipid environment, or viscous loading of the whole OHC. It has also been argued that such mechanisms might arise from energetically favoured binding of the anionic charge in prestin [19]. The source of the energy, however, is the potential across the cell membrane, maintained by the endocochlear potential, through which anions or intrinsic charges move.

It is noticeable that when measured in isolated OHCs, the piezo coefficients d_{13}, d_{31} are not completely equal but that charge displacement vs force shows that $d_{13} < d_{31}$ particularly at the most sensitive range (see Fig 5B in [7]). This may indicate some pre-existing degree of dissipation built into the OHC system.

The system described here is linear since all parameters have been linearised around the operating point. The incorporation of any non-linear component that generates distortion products depends upon the inclusion of the mechano-electric transduction I-X curve. The voltage dependence of the prestin is not sufficient to contribute a significant nonlinearity for small signals. One time domain model that does produce distortion products is that of Nobili and Mammano, [18]. The approach there was to expand the first term of Eqn 6 to first order in β and then identify β as the I-X function under assumptions that simplify Z_{OHC} and H_{TM}.. The full set of cochlear equations can then be numerically computed in the time domain as the m.e.t. nonlinearity enters in a tractable form. The price of the expansion is the introduction of a variable gain parameter which substitutes for the positive feedback afforded by experimental parameters in Eqn 6.

Two cochlear tuning regimes

The development of the model and Figure 3 shows that at low frequencies the effect of i_{piezo} can largely be ignored and the system exhibits positive feedback driven by the transduction current. This is a conventional electromotility

model. In this case the tuning is quite broad. At high frequencies, beyond the membrane cut-off, the system dynamics changes and the tuning becomes sharper. It has been argued that at least up to 20 kHz, the membrane time constant τ continues to fall and so that the low frequency cut off might never be approached and sufficient membrane potential change exists to drive a voltage sensitive prestin [5]. Although attractive, there are technical and bandwidth limitations to testing this proposition further using conventional patch clamp techniques.

A feature of auditory tuning curves often remarked upon is that there are discontinuities in their sharpness of tuning (the Q values) as well as the high frequency slopes [20]. The discontinuities distinguish tuning curves with CFs above or below 5 kHz. In addition, tuning curves with high CFs have a noticeable low frequency 'tail' or plateau in a wide variety of mammalian species (see for example [21] [22]). The above treatment suggests that the origin of such features are the two modes of OHC contributions to BM mechanics Eqn 6). These characteristics are shown in Figure 3B where the inverse of the transfer curves ($1/(H_{\mathbf{BM}}\text{-}f_{OHC})$ have been plotted as an indication of the threshold frequency tuning curves.

Conclusion

Although the effect of fluid coupling has been ignored, the results indicate semi-quantitively the curves that would be obtained in the intact cochlea. The low CF tuning curve is broader than for high CF; the high CF curve shows a low frequency plateau which is due mainly to the contribution of the m.e.t. driven OHC motility. The contribution to low frequency responses comes from RC-time constant limited motility of the OHC; the contribution to high frequency components of the curve arising from the charge movement inherent in the movement of anions at the cytoplasmic vestibule of prestin/SLC26A5 as well as charge reorganization of the protein. Recent cryo-e.m. structures of SLC26A5 help to resolve the precise molecular mechanisms underlying such dynamics [23] [24].

Time domain models, including the effects of wave propagation in the cochlear duct, are computationally harder to implement but the solutions defy easy analytical expression. Eqn 6 does admit a time domain formulation but necessarily involves adding the effect of the Green's function G(x,x'). It is unfortunate that the resulting computational complexity often obscures the underlying biophysics even though such models can automatically include the non-linear consequences of hair cell transduction including distortion products and compressive behaviour of the mechanics away from a linear threshold. The present highly simplified treatment can provide a framework in which the full time-dependent solution can be understood.

Acknowledgements

I thank Robert Fettiplace for comments on an earlier version of this work.

References

[1] J. Ashmore, 'Outer Hair Cells and Electromotility', *Cold Spring Harb Perspect Med*, vol. 9, no. 7, Jul. 2019, doi: 10.1101/cshperspect.a033522.

[2] J. Santos-Sacchi, K. H. Iwasa, and W. Tan, 'Outer hair cell electromotility is low-pass filtered relative to the molecular conformational changes that produce nonlinear capacitance', *J. Gen. Physiol.*, vol. 151, no. 12, pp. 1369–1385, Dec. 2019, doi: 10.1085/jgp.201812280.

[3] A. Vavakou, N. P. Cooper, and M. van der Heijden, 'The frequency limit of outer hair cell motility measured in vivo', *Elife*, vol. 8, 24 2019, doi: 10.7554/eLife.47667.

[4] P. Dallos and B. N. Evans, 'High-frequency motility of outer hair cells and the cochlear amplifier', *Science*, vol. 267, no. 5206, pp. 2006–9, Mar. 1995.

[5] S. L. Johnson, M. Beurg, W. Marcotti, and R. Fettiplace, 'Prestin-driven cochlear amplification is not limited by the outer hair cell membrane time constant', *Neuron*, vol. 70, no. 6, pp. 1143–54, 06-23 2011, doi: 10.1016/j.neuron.2011.04.024.

[6] K. H. Iwasa, 'Negative membrane capacitance of outer hair cells: electromechanical coupling near resonance', *Sci Rep*, vol. 7, no. 1, p. 12118, 21 2017, doi: 10.1038/s41598-017-12411-6.

[7] X. X. Dong, M. Ospeck, and K. H. Iwasa, 'Piezoelectric reciprocal relationship of the membrane motor in the cochlear outer hair cell', *Biophysical journal*, vol. 82, no. 3, pp. 1254–9, Mar. 2002, doi: 10.1016/S0006-3495(02)75481-7.
[8] R. D. Rabbitt, 'The cochlear outer hair cell speed paradox', *PNAS*, vol. 117, no. 36, pp. 21880–21888, Sep. 2020, doi: 10.1073/pnas.2003838117.
[9] D. C. Mountain and A. E. Hubbard, 'A piezoelectric model of outer hair cell function', *J Acoust Soc Am*, vol. 95, no. 1, pp. 350–354, Jan. 1994, doi: 10.1121/1.408273.
[10] C. D. Geisler, 'A realizable cochlear model using feedback from motile outer hair cells', *Hear Res*, vol. 68, no. 2, pp. 253–262, Aug. 1993, doi: 10.1016/0378-5955(93)90129-o.
[11] Allen and Sondhi, 'Cochlear macromechanics: tiem domain solutions', *J. Acoust. Soc. Am.*, vol. 66, no. 1, pp. 123–132, 1977.
[12] F. Mammano and R. Nobili, 'Biophysics of the cochlea: linear approximation', *J Acoust Soc Am*, vol. 93, no. 6, pp. 3320–3332, Jun. 1993, doi: 10.1121/1.405716.
[13] S. T. Neely and D. O. Kim, 'A model for active elements in cochlear biomechanics', *J Acoust Soc Am*, vol. 79, no. 5, pp. 1472–1480, May 1986, doi: 10.1121/1.393674.
[14] J. E. Gale and J. F. Ashmore, 'Charge displacement induced by rapid stretch in the basolateral membrane of the guinea-pig outer hair cell', *Proc Biol Sci*, vol. 255, no. 1344, pp. 243–249, Mar. 1994, doi: 10.1098/rspb.1994.0035.
[15] S. X. Sun, B. Farrell, M. S. Chana, G. Oster, W. E. Brownell, and A. A. Spector, 'Voltage and frequency dependence of prestin-associated charge transfer', *J Theor Biol*, vol. 260, no. 1, pp. 137–44, Sep. 2009, doi: 10.1016/j.jtbi.2009.05.019.
[16] C. D. Geisler and X. Shan, 'A Model for Cochlear Vibrations Based on Feedback from Motile Outer Hair Cells', in *The Mechanics and Biophysics of Hearing*, New York, NY, 1990, pp. 86–95. doi: 10.1007/978-1-4757-4341-8_11.
[17] J. Santos-Sacchi and W. Tan, 'Complex nonlinear capacitance in outer hair cell macro-patches: effects of membrane tension', *Sci Rep*, vol. 10, no. 1, p. 6222, Dec. 2020, doi: 10.1038/s41598-020-63201-6.
[18] R. Nobili and F. Mammano, 'Biophysics of the cochlea. II: Stationary nonlinear phenomenology', *J Acoust Soc Am*, vol. 99, no. 4 Pt 1, pp. 2244–2255, Apr. 1996, doi: 10.1121/1.415412.
[19] J. Santos-Sacchi and L. Song, 'Chloride-driven electromechanical phase lags at acoustic frequencies are generated by SLC26a5, the outer hair cell motor protein', *Biophys J*, vol. 107, no. 1, pp. 126–133, Jul. 2014, doi: 10.1016/j.bpj.2014.05.018.
[20] A. N. Temchin, N. C. Rich, and M. A. Ruggero, 'Threshold tuning curves of chinchilla auditory-nerve fibers. I. Dependence on characteristic frequency and relation to the magnitudes of cochlear vibrations', *J. Neurophysiol.*, vol. 100, no. 5, pp. 2889–2898, Nov. 2008, doi: 10.1152/jn.90637.2008.
[21] A. M. Taberner and M. C. Liberman, 'Response properties of single auditory nerve fibers in the mouse', *J Neurophysiol*, vol. 93, no. 1, pp. 557–569, Jan. 2005, doi: 10.1152/jn.00574.2004.
[22] M. C. Liberman, 'Auditory-nerve response from cats raised in a low-noise chamber', *The Journal of the Acoustical Society of America*, vol. 63, no. 2, pp. 442–455, Feb. 1978, doi: 10.1121/1.381736.
[23] J. Ge *et al.*, 'Molecular mechanism of prestin electromotive signal amplification', *Cell*, vol. 184, no. 18, pp. 4669-4679.e13, Sep. 2021, doi: 10.1016/j.cell.2021.07.034.
[24] N. Bavi *et al.*, 'The conformational cycle of prestin underlies outer-hair cell electromotility', *Nature*, vol. 600, no. 7889, Art. no. 7889, Dec. 2021, doi: 10.1038/s41586-021-04152-4.

The Functional Contributions of Links in Mammalian Cochlear Hair Bundles

Zenghao Zhu,[1, a)] Wisam Reid,[1, 2, b)] and Dáibhid Ó Maoiléidigh[1, c)]

[1)] *Department of Otolaryngology-Head and Neck Surgery, Stanford University, Stanford, CA, USA*
[2)] *Current affiliation: Harvard Medical School, Boston, MA, USA*

[a)] *Electronic mail: perterzhu@stanford.edu*
[b)] *Electronic mail: wisam@g.harvard.edu*
[c)] *Corresponding author: dmelody@stanford.edu*

Abstract. In the mammalian cochlea, hair bundles of the sensory outer and inner hair cells detect mechanical signals. A hair bundle comprises a set of rod-like stereocilia that pivot around their insertion points in the hair-cell's apex. Stereocilia are linked by gating springs and connectors, also known as top or shaft connectors, side, lateral, or ankle links. Gating springs link neighboring stereocilia of differing height, while connectors link all neighboring stereocilia. Sound-induced gating-spring oscillations open and close mechanoelectrical transduction channels attached to the gating springs, causing oscillations in the hair cell's receptor current. In contrast to gating springs, connectors are not attached to channels and their functional role is unclear. To determine how the specific properties of gating springs and connectors contribute to outer-hair-cell bundle function, we use a computational model of an outer-hair-cell bundle, which accounts for nonlinear hair-bundle splaying at rest, nonlinear fluid forces on stereocilia, and nonlinear channel gating. The model is validated by reproducing many experimental observations, including stereocilium splaying at rest and hair-bundle stiffness decreases caused by breaking gating springs or connectors. We discuss how varying the gating-spring and connector stiffnesses affects the receptor current in response to stimulation at the characteristic frequency of the hair cell.

INTRODUCTION

In the auditory, vestibular, and lateral-line organs of vertebrates, hair bundles convert forces induced by sound, head motion or position, and fluid into receptor currents [1]. A hair bundle comprises rows and columns of rod-like stereocilia protruding from the hair-cell's apex and increasing in height toward one edge of the bundle. In a column, stereocilia of increasing height are linked by gating springs, which transmit forces that gate (open and close) mechanoelectrical transduction (MET) channels. Rows are formed by stereocilia of similar height, which are not linked by gating springs. Deflection of a hair bundle toward its tallest row (row 1) causes its stereocilia to pivot, gating-spring forces to increase, and MET channels to open. Gating springs comprise tip-links, made of the proteins protocadherin 15 (PCDH15) and cadherin 23 (CDH23), and other elastic elements in series with the MET channels [1, 2]. Stereocilia are also linked within rows and columns by connectors, which do not gate channels. It has been proposed that connectors increase the coherence of stereociliary motions and decrease hair-bundle damping, but recent work in outer-hair-cell (OHC) bundles show that connectors can decrease the uniformity of stereociliary motions and increase hair-bundle damping, calling the role of OHC connectors into question [3, 4, 5]. Connectors comprise different elements in different types of bundles and include top and shaft connectors, side, lateral, and ankle links [6, 7, 8, 9, 10].

The structure and function of hair bundles differ between species, between organs, and between different locations within an organ, but we have a limited understanding of the relationship between a hair bundle's structure and its function [1]. Gating-spring and connector stiffnesses differ between different types of bundles and between different locations within an organ, but we do not know how specific gating-spring or connector stiffnesses contribute to bundle function in the mammalian cochlea [3, 11, 12, 13]. In addition, connector stiffness may differ between rows and columns within the same bundle and we have no information about gating-spring or connector damping [12]. Here we focus on the contribution of gating-spring and connector stiffness to OHC-bundle function in the mammalian cochlea.

The function of an OHC bundle is to convert sound-induced forces into receptor currents. OHC structure changes from the cochlear base to apex: bundle stiffness decreases, stereocilium number decreases, stereocilium height increases, pivot and gating-spring stiffness decrease, and gating-spring resting tension decreases [11, 14, 15, 16, 17]. OHC connectors comprise stereocilin, immunolabeling of which decreases in intensity from base to apex, suggesting a decrease in connector stiffness and damping from base to apex [9]. Because OHC mechanics changes with characteristic place along the cochlea, we study OHC bundles corresponding to a specific characteristic-frequency (CF) region (4 kHz). Using a computational model, we determine the OHC bundle receptor current in response to

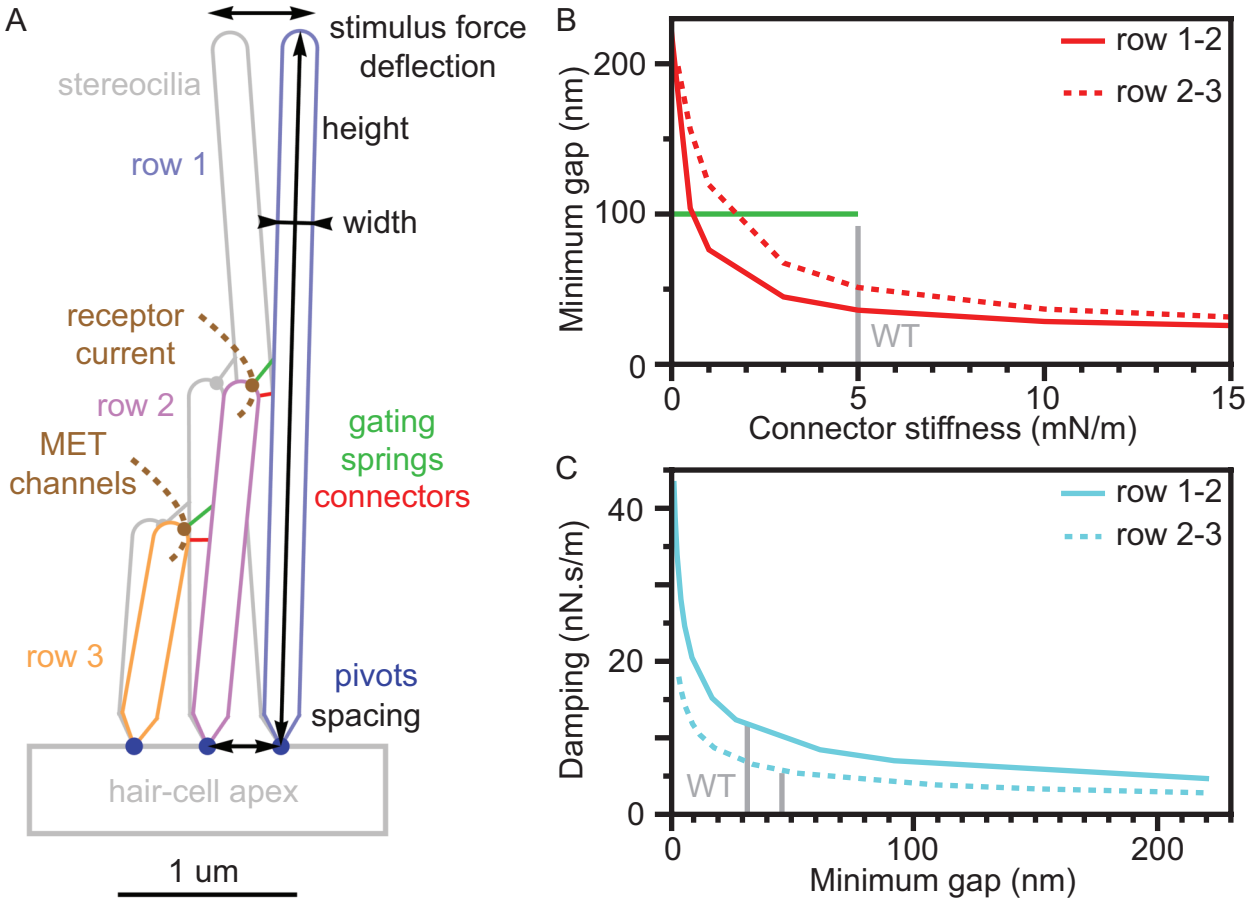


FIGURE 1. (A) An OHC bundle column. Stereocilia of different rows have different heights, widths, and pivot spacing (black). In response to a sound-induced stimulus force, stereocilia deflect around their pivots (blue) in the hair-cell apex. Stereocilia are linked by connectors (red) and gating springs (green) that gate MET channels (brown circles) through which the receptor current (brown dashed lines) flows. To help visualize the deflection, a deflection much larger than the physiological maximum (less than a stereocilium's radius) is shown. (B) In a bundle without gating-springs, the minimum gaps between neighboring stereocilia are shown versus the connector stiffness. The minimum gap is along the undeflected connectors in panel A. The wild-type (WT) connector stiffness is indicated by the vertical gray line. The length of a reforming gating spring composed of PDH15-PDH15 is indicated by the horizontal green line. (C) Fluid damping coefficients associated with changing stereocilium separations along the connectors are shown versus the minimum gap between stereocilia. The WT values of the minimum gaps are indicated by vertical gray lines.

oscillatory tectorial-membrane forces versus the stimulus frequency. To determine the contributions of gating-springs and connectors, we vary the properties of these components and calculate the changes in the receptor current.

OUTER HAIR CELL BUNDLE MECHANICS

To demonstrate some of the consequences arising from varying gating-spring and connector stiffness, we use an OHC bundle model consisting of a single column of stereocilia (Fig. 1A). Because the stereocilia pivot rather than bend, a single connector between neighboring stereocilia is sufficient to capture the connector's mechanical effects [18, 19, 20, 21]. Pivots, gating-springs, and connectors are viscoelastic elements. Fluid between neighboring stereocilia causes damping and couples the stereocilia [3, 22, 23]. We describe each MET channel using two-state kinetics [24]. Myosin motors set the gating-spring resting tensions and the resting open probabilities [1, 11, 25, 26, 27]. We do not include adaptation for three reasons [1, 26]. First, adaptation's mechanism in OHCs remains a matter of debate [27, 28, 29]. Second, adaptation has small effect on OHC-bundle mechanics [27, 29]. Third, by omitting

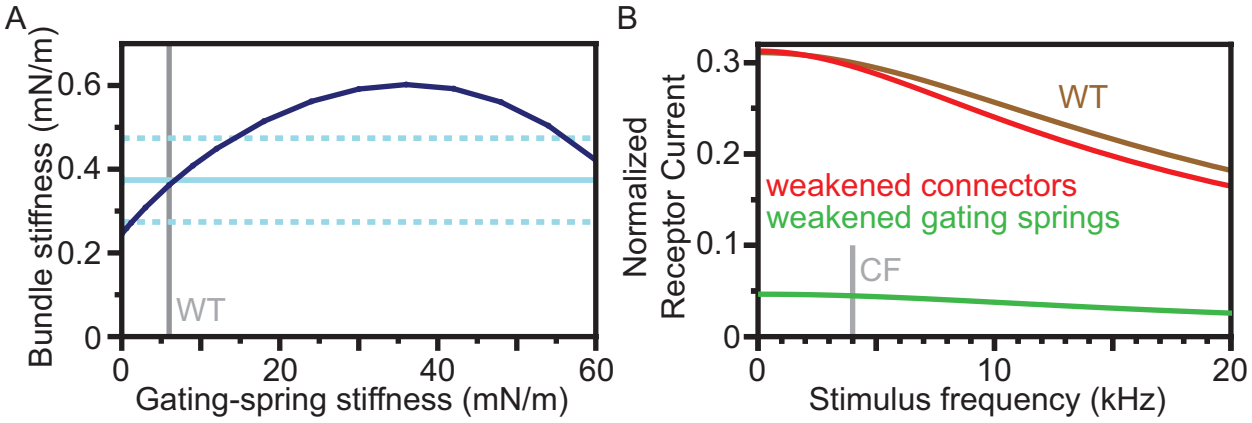


FIGURE 2. (A) The bundle's stiffness is shown versus the gating-spring stiffness (dark blue). For comparison, the experimentally-measured hair-bundle stiffness is shown as horizontal light-blue lines (mean (solid) ± standard deviation (dashed)) [11]. For a single column, this stiffness is the measured bundle stiffness divided by the number of columns in the bundle. The WT gating-spring stiffness (gray line) is chosen such that the model bundle's stiffness matches that of experiments. (B) The normalized receptor current is shown versus the stimulus frequency for the WT bundle (brown). Receptor currents for weakened (10% of the WT values) gating-springs (green) or connectors (red) are also shown. A vertical gray line indicates the bundle's CF.

adaptation, we distinguish its effect on the receptor current from the viscoelastic effects we discuss here. Based on prior measurements, we expect adaptation to high-pass filter the receptor current with a corner frequency of about 600 Hz for a 4 kHz CF OHC [30]. The tectorial membrane displaces the OHC bundle and the fluid around the bundle almost simultaneously, thus there is little fluid damping caused by the surrounding fluid on an OHC *in situ* [31, 32]. We stimulate the row 1 tip of the OHC bundle using sinusoidal tectorial-membrane forces. The receptor current is the sum of the MET currents through each channel and is normalized to the maximum current for fully open channels.

Wild-type (WT) OHC bundle parameters correspond to the 4 kHz CF region in the rat and mouse. We determine the stereocilium heights, widths, pivot spacings and resting gating-spring lengths from published electron and light microscopy data [11, 33, 34]. We validate the model by reproducing many different experimental observations. For example, by matching the bundle's stiffness and damping to that of experimental measurements, we determine the pivot, gating-spring, and connector stiffness and damping [4, 11]. We describe channel dynamics using a recent estimate of the channel time constant and a gating swing such that the width of the receptor-current activation curve matches experimental measurements [24, 29, 35]. By integrating analytical expressions for the damping between separating and sliding cylinders, we calculate the fluid damping between stereocilia and validate these calculations by comparing them with prior computational fluid dynamics results [3, 23, 36, 37].

RESULTS

The connectors and gating springs cause the stereocilia to lean toward each other at rest (gray stereocilia, Fig. 1A). The leaning angles and the gaps between neighboring stereocilia depend nonlinearly on the link stiffnesses (e.g. Fig. 1B). Noise breaks OHC tip links, the extracellular parts of gating springs, increasing the gaps between neighbors [2, 38, 39, 40]. In other words, noise causes stereocilia to splay in their resting state. Tip links can reform, first as 100-nm PDH15-PDH15 filaments before they are replaced by PDH15-CDH23 filaments [2, 41]. Without connectors, however, the minimum gaps between neighboring stereocilia would be too large (> 200 nm) to allow the tip links to reform. In our model, increasing the connector stiffness decreases the minimum gaps between neighbors, which would allow the tip links to reform (Fig. 1B). As suggested previously, the WT connector stiffness is sufficiently large to constrain the minimum gaps and limit stereocilium splaying at rest [9, 42].

Fluid damping between neighboring stereocilia depends nonlinearly on the gaps between the stereocilia [3, 22, 23]. In our model, the fluid damping coefficient associated with stereocilia separating decreases as the minimum gap between stereocilia increases (Fig. 1C). Thus, decreasing the gating-spring or connector stiffness decreases the fluid damping between neighbors. Because the row 2-3 gap exceeds that of row 1-2, the row 2-3 damping coefficient is less than that of row 1-2.

Owing to MET channel gating, hair-bundle stiffness depends nonlinearly on gating-spring stiffness (Fig. 2A). This nonlinearity arises because hair-bundle stiffness depends on the gating-spring stiffness and the MET channel's open probability, which depends nonlinearly on the gating-spring stiffness [13, 24, 26, 43]. At all gating-spring stiffness values, MET channel opening decreases bundle stiffness, a phenomenon known as gating compliance [13, 43]. Although bundle stiffness initially increases with gating-spring stiffness, gating compliance also increases with gating-spring stiffness, causing bundle stiffness to eventually decrease with gating-spring stiffness. For the WT bundle, decreasing gating-spring stiffness decreases bundle stiffness, in agreement with experimental observations [11].

Although the bundle displaces less in response to stimulus forces (bundle stiffness increases) as the gating-spring stiffness increases, the CF receptor current in response to an oscillatory stimulus force increases markedly with gating-spring stiffness (Fig. 2B). In contrast, the CF receptor current increases very little with increasing connector stiffness. Because of bundle damping and MET channel kinetics, the receptor current decreases with increasing stimulus frequency, but the decrease is small at the bundle's CF.

DISCUSSION AND CONCLUSION

Although we model only a single column of an OHC bundle, we gain insight into three nonlinear contributions of gating springs and connectors to OHC bundle function. First, connectors constrain the gaps between neighboring stereocilia, allowing broken gating springs to reform. Second, gating-springs and connectors regulate the damping between neighbors by constraining the gaps between them. Third, gating springs can increase or decrease bundle stiffness, because they also regulate MET-channel gating. Additionally, we find that the receptor current is much more sensitive to gating-spring than connector stiffness.

The increase in gaps between neighboring stereocilia with decreasing connector stiffness agrees with the increase in splaying at rest seen in OHC mutants lacking connectors [4, 9, 10]. Connectors may also protect gating springs from breaking in response to high stimulus intensities, but we expect this protection to come at a cost [42]. If connectors act in parallel with gating springs and reduce gating-spring forces at high stimulus intensities, then they will also reduce gating-spring forces at low stimulus intensities, decreasing bundle sensitivity. Whether connectors increase OHC bundle coherence remains to be determined [3]. We also show that connectors regulate bundle damping and have little effect on CF receptor currents.

Bundle stiffness varies nonlinearly with gating-spring stiffness. The size of this effect increases with the gating swing size, the change in the gating-spring's length when the MET channel opens [1, 26, 43]. Here we use a gating-swing of 0.5 nm to match the measured width of the current activation curves [29]. It is likely, however, that the activation curve width is overestimated and that we have consequently underestimated the gating-swing size [12, 28, 29]. Thus, the gating-spring stiffness at which the bundle stiffness peaks may be smaller than we show here. Because of the experimentally-observed increase in gating-spring stiffness toward the cochlear base, gating compliance may contribute more to the stiffness of basal bundles than apical bundles [11, 14].

The contributions of gating-spring and connector damping remain to be determined. We find that links regulate fluid damping and prior work shows that MET channel gating dynamics causes damping [24]. It is difficult to disentangle these contributions to bundle damping from intrinsic link damping, even though we have data showing that bundle damping changes when links are broken [4]. We expect link damping to decrease bundle displacements and receptor currents and because mammalian hair bundles must respond to high stimulus frequencies, we expect link damping to be small. Links may, however, have unavoidable damping owing to physical constraints on their biomolecular structure. Link damping may contribute to calcium-independent adaptation or the coherence of stereociliary motions at high stimulus frequencies [3, 44].

Because of the minimal number of assumptions and because the single-column model is validated by matching experimental observations, we expect our single-column 2D results to apply to 3D OHC bundles with adaptation. However, we also expect the results to be modified in 3D bundles with adaptation and for these bundles to exhibit additional effects associated with gating-spring and connector mechanics. Our single-column model does not account for the fact that OHC columns do not move in parallel *in vitro*, there are connectors between columns, and adaptation affects the receptor current in response to low-frequency stimuli [5, 7, 8, 9, 10, 30, 45]. Comparing 2D bundles, 3D bundles, and bundles with and without adaptation will further clarify the contributions of gating-springs and connectors to OHC function.

ACKNOWLEDGMENTS

This work was funded by Stanford Otolaryngology-Head and Neck Surgery startup funding for DÓ, the Stanford Initiative to Cure Hearing Loss, and National Institute on Deafness and Other Communication Disorders grant 2R01 DC003896-21 (PI Anthony Ricci) for DÓ.

REFERENCES

1. D. Ó Maoiléidigh and A. J. Ricci, "A bundle of mechanisms: Inner-ear hair-cell mechanotransduction." Trends Neurosci **42**, 221–236 (2019).
2. P. Kazmierczak, H. Sakaguchi, J. Tokita, E. M. Wilson-Kubalek, R. A. Milligan, U. Müller, and B. Kachar, "Cadherin 23 and protocadherin 15 interact to form tip-link filaments in sensory hair cells." Nature **449**, 87–91 (2007).
3. A. S. Kozlov, J. Baumgart, T. Risler, C. P. C. Versteegh, and A. J. Hudspeth, "Forces between clustered stereocilia minimize friction in the ear on a subnanometre scale," Nature **474**, 376–379 (2011).
4. A. X. Cartagena-Rivera, S. Le Gal, K. Richards, E. Verpy, and R. S. Chadwick, "Cochlear outer hair cell horizontal top connectors mediate mature stereocilia bundle mechanics." Sci Adv **5**, eaat9934 (2019).
5. A. W. Peng, A. L. Scharr, G. A. Caprara, D. Nettles, C. R. Steele, and A. J. Ricci, "Fluid jet stimulation of auditory hair bundles reveal spatial non-uniformities and two viscoelastic-like mechanisms." Front Cell Dev Biol **9**, 725101 (2021).
6. G. P. Richardson and C. Petit, "Hair-bundle links: Genetics as the gateway to function." Cold Spring Harb Perspect Med **9** (2019), 10.1101/csh-perspect.a033142.
7. R. J. Goodyear, W. Marcotti, C. Kros, and G. P. Richardson, "Development and properties of stereociliary link types in hair cells of the mouse cochlea." J Comp Neurol **485**, 75–85 (2005).
8. V. Tsuprun, P. A. Schachern, S. Cureoglu, and M. Paparella, "Structure of the stereocilia side links and morphology of auditory hair bundle in relation to noise exposure in the chinchilla." J Neurocytol **32**, 1117–1128 (2003).
9. E. Verpy, M. Leibovici, N. Michalski, R. J. Goodyear, C. Houdon, D. Weil, G. P. Richardson, and C. Petit, "Stereocilin connects outer hair cell stereocilia to one another and to the tectorial membrane." J Comp Neurol **519**, 194–210 (2011).
10. P. Avan, S. Le Gal, V. Michel, T. Dupont, J.-P. Hardelin, C. Petit, and E. Verpy, "Otogelin, otogelin-like, and stereocilin form links connecting outer hair cell stereocilia to each other and the tectorial membrane." Proc Natl Acad Sci U S A **116**, 25948–25957 (2019).
11. M. Tobin, A. Chaiyasitdhi, V. Michel, N. Michalski, and P. Martin, "Stiffness and tension gradients of the hair cell's tip-link complex in the mammalian cochlea." Elife **8** (2019), 10.7554/eLife.43473.
12. J.-H. Nam, A. W. Peng, and A. J. Ricci, "Underestimated sensitivity of mammalian cochlear hair cells due to splay between stereociliary columns," Biophys. J. **108**, 2633–2647 (2015).
13. P. Martin, A. D. Mehta, and A. J. Hudspeth, "Negative hair bundle stiffness betrays a mechanism for mechanical amplification by the hair cell," Proc. Natl. Acad. Sci. USA **97**, 12026–12031 (2000).
14. D. Strelioff and A. Flock, "Stiffness of sensory-cell hair bundles in the isolated guinea pig cochlea." Hear Res **15**, 19–28 (1984).
15. D. J. Lim, "Functional structure of the organ of corti: a review." Hear Res **22**, 117–146 (1986).
16. B. Roth and V. Bruns, "Postnatal development of the rat organ of corti. ii. hair cell receptors and their supporting elements." Anat Embryol (Berl) **185**, 571–581 (1992).
17. Y. M. Yarin, A. N. Lukashkin, A. A. Poznyakovskiy, H. Meissner, M. Fleischer, J. Baumgart, C. Richter, E. Kuhlisch, and T. Zahnert, "Tonotopic morphometry of the lamina reticularis of the guinea pig cochlea with associated microstructures and related mechanical implications." J Assoc Res Otolaryngol **15**, 1–11 (2014).
18. J. Howard and J. F. Ashmore, "Stiffness of sensory hair bundles in the sacculus of the frog." Hear Res **23**, 93–104 (1986).
19. A. C. Crawford, M. G. Evans, and R. Fettiplace, "Activation and adaptation of transducer currents in turtle hair cells," J Physiol **419**, 405–434 (1989).
20. K. D. Karavitaki and D. P. Corey, "Sliding adhesion confers coherent motion to hair cell stereocilia and parallel gating to transduction channels." J Neurosci **30**, 9051–9063 (2010).
21. Y. Wang, C. R. Steele, S. Puria, and A. J. Ricci, "In situ motions of individual inner-hair-cell stereocilia from stapes stimulation in adult mice." Commun Biol **4**, 958 (2021).
22. D. E. Zetes and C. R. Steele, "Fluid-structure interaction of the stereocilia bundle in relation to mechanotransduction." J Acoust Soc Am **101**, 3593–3601 (1997).
23. K. K. Miller, P. Atkinson, K. R. Mendoza, D. Ó Maoiléidigh, and N. Grillet, "Dimensions of a living cochlear hair bundle." Front Cell Dev Biol **9**, 742529 (2021).
24. V. Bormuth, J. Barral, J.-F. Joanny, F. Jülicher, and P. Martin, "Transduction channels' gating can control friction on vibrating hair-cell bundles in the ear," Proc. Natl. Acad. Sci. USA **111**, 7185–7190 (2014).
25. J. A. Assad, G. M. G. Shepherd, and D. P. Corey, "Tip-link integrity and mechanical transduction in vertebrate hair cells," Neuron **7**, 985–994 (1991).
26. D. P. Corey, D. Ó Maoiléidigh, and J. F. Ashmore, "Mechanical transduction processes in the hair cell," in *Understanding the Cochlea*, edited by G. A. Manley, A. W. Gummer, A. N. Popper, and R. R. Fay (Springer International Publishing, Cham, 2017) pp. 75–111.
27. G. A. Caprara, A. A. Mecca, and A. W. Peng, "Decades-old model of slow adaptation in sensory hair cells is not supported in mammals." Sci Adv **6**, eabb4922 (2020).
28. D. Ó Maoiléidigh and A. J. Hudspeth, "Effects of cochlear loading on the motility of active outer hair cells," Proc. Natl. Acad. Sci. USA **110**, 5474–5479 (2013).
29. G. A. Caprara and A. W. Peng, "Mechanotransduction in mammalian sensory hair cells." Mol Cell Neurosci **120**, 103706 (2022).
30. A. J. Ricci, H. J. Kennedy, A. C. Crawford, and R. Fettiplace, "The transduction channel filter in auditory hair cells." J Neurosci **25**, 7831–7839 (2005).

31. J.-H. Nam and R. Fettiplace, "Theoretical conditions for high-frequency hair bundle oscillations in auditory hair cells." Biophys J **95**, 4948–4962 (2008).
32. D. M. Freeman and T. F. Weiss, "The role of fluid inertia in mechanical stimulation of hair cells." Hear Res **35**, 201–207 (1988).
33. D. N. Furness, S. Mahendrasingam, M. Ohashi, R. Fettiplace, and C. M. Hackney, "The dimensions and composition of stereociliary rootlets in mammalian cochlear hair cells: comparison between high- and low-frequency cells and evidence for a connection to the lateral membrane." J Neurosci **28**, 6342–6353 (2008).
34. D. N. Furness, Y. Katori, B. Nirmal Kumar, and C. M. Hackney, "The dimensions and structural attachments of tip links in mammalian cochlear hair cells and the effects of exposure to different levels of extracellular calcium." Neuroscience **154**, 10–21 (2008).
35. M. Beurg, J.-H. Nam, and R. Fettiplace, "The speed of the hair cell mechanotransducer channel revealed by fluctuation analysis." J Gen Physiol **153** (2021), 10.1085/jgp.202112959.
36. A. J. Hunt, F. Gittes, and J. Howard, "The force exerted by a single kinesin molecule against a viscous load." Biophys J **67**, 766–781 (1994).
37. J. Baumgart, *The Hair Bundle: Fluid-Structure Interaction in the Inner Ear*, Ph.D. thesis, Technische Universität Dresden (2011).
38. J. O. Pickles, M. P. Osborne, and S. D. Comis, "Vulnerability of tip links between stereocilia to acoustic trauma in the guinea pig." Hear Res **25**, 173–183 (1987).
39. J. A. Clark and J. O. Pickles, "The effects of moderate and low levels of acoustic overstimulation on stereocilia and their tip links in the guinea pig." Hear Res **99**, 119–128 (1996).
40. E. L. Wagner and J.-B. Shin, "Mechanisms of hair cell damage and repair." Trends Neurosci **42**, 414–424 (2019).
41. A. A. Indzhykulian, R. Stepanyan, A. Nelina, K. J. Spinelli, Z. M. Ahmed, I. A. Belyantseva, T. B. Friedman, P. G. Barr-Gillespie, and G. I. Frolenkov, "Molecular remodeling of tip links underlies mechanosensory regeneration in auditory hair cells," PLoS Biol. **11**, e1001583 (2013).
42. W. Han, J.-O. Shin, J.-H. Ma, H. Min, J. Jung, J. Lee, U.-K. Kim, J. Y. Choi, S. J. Moon, D. W. Moon, J. Bok, and C. H. Kim, "Distinct roles of stereociliary links in the nonlinear sound processing and noise resistance of cochlear outer hair cells." Proc Natl Acad Sci U S A **117**, 11109–11117 (2020).
43. J. Howard and A. J. Hudspeth, "Compliance of the hair bundle associated with gating of mechanoelectrical transduction channels in the bullfrog's saccular hair cell," Neuron **1**, 189–199 (1988).
44. J. Howard and A. J. Hudspeth, "Mechanical relaxation of the hair bundle mediates adaptation in mechanoelectrical transduction by the bullfrog's saccular hair cell," Proc. Natl. Acad. Sci. USA **84**, 3064–3068 (1987).
45. G. A. Caprara, A. A. Mecca, Y. Wang, A. J. Ricci, and A. W. Peng, "Hair bundle stimulation mode modifies manifestations of mechanotransduction adaptation." J Neurosci **39**, 9098–9106 (2019).

Not so presto? Can prestin be sluggish?

Kuni H. Iwasa[a)]

NIDCD - National Institutes of Health Bldg. 35A Rm 1F121A,
Bethesda Maryland 20892 USA

(Dated: 14 July 2022)

Abstract. Prestin (SLC26A5), a protein essential for the sensitivity of the mammalian ear, was so named from *presto*. The assumption was that this membrane protein supports fast movement of outer hair cells (OHCs) that matches the mammalian hearing range, up to 20 kHz in general and beyond, depending on the species. *In vitro* data from isolated OHCs appeared to be consistent with such frequencies. However, some recent reports cast doubts on this assumption, suggesting that the intrinsic transition rates of this protein are much lower, about 3 kHz for guinea pigs, not covering the auditory frequency range of the animal. Recent *in vivo* data also show that the amplitude of OHC motion rolls off well below the best frequency of the location. The present report examines whether or not these recent observations are compatible with the physiological function of OHCs by using simple piezoelectric models.

INTRODUCTION

The significance of prestin for the sensitivity and frequency selectivity of the mammalian ear has been well established [1]. However, the detailed mechanism, with which this piezoelectric membrane protein plays its physiological role, remains not as clear. The frequency range that this protein is capable of responding is a critical issue.

Earlier *in vitro* studies on isolated OHCs confirmed that their fast motile response is based on piezoelectricity [2, 3, 4]. Force generation under quasi-isometric condition was shown to have flat frequency dependence up to 60 kHz [5]. The characteristic frequency of the power spectrum of membrane current due to prestin was about 40 kHz [6]. Those frequencies were considered to be lower bounds imposed by experimental conditions.

However, recent reports on OHCs in *in vitro* [7, 8] appear to contradict these earlier observations. In addition, *in vivo* data obtained with optical coherence tomography (OCT) [9] have been interpreted as evidence against fast motile response of OHCs.

The present report examines the implications of these two kinds of reports by asking two questions. First, can OHCs counteract local viscous drag incurred by the movement of the organ of Corti in the cochlea if gating rates of prestin are as low as these recent reports suggest? At the basal end, where the traveling wave initiates, viscous drag must be counteracted by OHCs. This condition determines the frequency limit of the ear.

Second, should the movement of OHCs large at their best frequencies *in vivo*? The first issue is addressed by evaluating power generation by OHCs driven by prestin with finite transition rates. The second issue is addressed using a simple model system, which consists of two harmonic oscillators, one of which has a driver and the other with a damper.

1. TRANSITION RATES

First, let us derive the equation of motion for a cell, which is driven by a motile molecule with finite intrinsic transition rates. An earlier version has been published as Appendix to Santos-Sacchi et al. [8].

[a)]Corresponding author: kuni.iwasa@gmail.com

Rate equation

Consider a membrane molecule with two discrete conformational states C_0 and C_1 and let the transition rates k_+ and k_- between them, with k_+ from C_0 and C_1, and k_- the opposite, i.e.

$$C_0 \underset{k_-}{\overset{k_+}{\rightleftharpoons}} C_1 \tag{1}$$

Let P_1 be the probability that the molecule in state C_1, which elongate the cell. Suppose charge transfer $q(>0)$ is associated with a change $a(>0)$ in the length of the cell. Then P_1 satisfies

$$\frac{P_1}{1-P_1} = \frac{k_+}{k_-} = \exp[\beta[q(V-V_0)+aF]], \tag{2}$$

where F is the axial force applied on the cell. The transition rates k_+ and k_- can be expressed as

$$k_+ = \exp\left[\alpha\beta[q(V-V_0)+aF]\right], \tag{3a}$$
$$k_- = \exp\left[(-1+\alpha)\beta[q(V-V_0)+aF]\right], \tag{3b}$$

by introducing a parameter α. For the rest of the present paper, the dependence on the value of the parameter α does not appear except for $\omega_g(=k_++k_-)$.

The time dependence of P_1 can be expressed by the rate equation

$$\frac{d}{dt}P_1 = k_+ - (k_++k_-)P_1. \tag{4}$$

Now assume that the voltage V consists of a constant term $\overline{V}$ and a small sinusoidal component with amplitude v , i.e. $V = \overline{V} + v\exp[i\omega t]$, where ω is the angular frequency and $i=\sqrt{-1}$. Then the transition rates are time dependent due to the voltage dependence, satisfying

$$\frac{k_+}{k_-} = \frac{\overline{k}_+}{\overline{k}_-}(1-\beta qv\exp[i\omega t]), \tag{5}$$

where $\overline{k}_+$ and $\overline{k}_-$ are time independent, and small amplitude v implies $\beta qv \ll 1$. A set of k_+ and k_- that satisfies Eq. 5 can be expressed

$$k_+ = \overline{k}_+(1-\alpha\beta qv\exp[i\omega t]), \tag{6a}$$
$$k_- = \overline{k}_-\{1-(1-\alpha)\beta qv\exp[i\omega t]\}. \tag{6b}$$

If we express $P_1 = \overline{P}_1 + p_1\exp[i\omega t]$, for the 0th and 1st order terms we have, respectively [10]

$$\overline{P}_1 = \frac{\overline{k}_+}{\overline{k}_+ + \overline{k}_-}, \tag{7a}$$
$$p_1 = -\frac{\overline{k}_+\overline{k}_-}{\overline{k}_+ + \overline{k}_-}\cdot\frac{\beta qv}{i\omega + \overline{k}_+ + \overline{k}_-}. \tag{7b}$$

This simple result can be obtained in the case $\alpha = \overline{k}_-/(\overline{k}_+ + \overline{k}_-)$.

By using the relationship $\overline{P}_1 = \overline{k}_+/(\overline{k}_+ + \overline{k}_-)$, Eq. 7b turns into

$$p_1 = \frac{\beta q\overline{P}_\pm}{1+i\omega/\omega_g}\cdot v, \tag{8}$$

with $\overline{P}_\pm = \overline{P}_1(1-\overline{P})$ and $\omega_g = \overline{k}_+ + \overline{k}_-$. Here, the average values and the amplitudes of sinusoidal components are represented by notations similar to the voltage. Notice that Eq. 8 satisfies the Boltzmann distribution. That means the frequency dependence is determined only by the transition rates and unaffected by mechanical load or by mechanical relaxation.

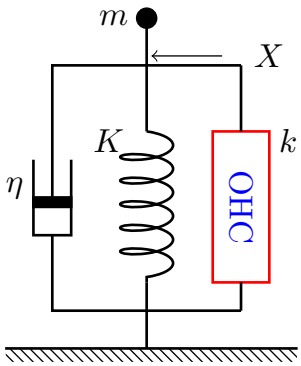


FIGURE 1. A schematic representation of a cell with mechanical load. The stiffness of the cell is k, the stiffness of the external elastic load is K, the inertial load is m, and drag coefficient is η. The state of the outer hair cell (OHC) is described by P. The mechanical displacement of the cell is represented by X. The properties of OHC cannot easily be rendered as a simple combination of elastic- and displacement elements, even though such attempts have been made earlier [11, 12, 13].

Equation of motion

Consider a system, where an OHC is associated with mass m, drag coefficient η, and external elastic load K (Fig. 1). Assume that the material stiffness of the OHC is k [14]. The sum of inertial force and drag is balanced with elastic force $F(t)$. The displacement $X(t)$ satisfies the equation

$$\left(m\frac{d^2}{dt^2}+\eta\frac{d}{dt}\right)X(t)=F(t). \tag{9}$$

If we are interested in the response to the stimulus with an angular frequency ω, the equation of motion Eq. 9 can be expressed as

$$(-\omega^2 m+i\eta\omega)x=f(\omega), \tag{10}$$

where x and f are, respectively, the amplitudes of displacement X and force F with the frequency ω.

If the force F in Eq. 3 is due to the external load alone, we have $F=-KX$ and $f=-Kx$, which are the familiar equations of motion. In such a case, the equilibrium condition is $X=0$ and $x=0$. In our system, however, movement is driven by a deviation from the Boltzmann distribution. If the displacement can respond instantaneously to voltage changes, the system goes from one equilibrium to another.

Note that the quantity p_∞ is similar to p_1 in Eq. 8 in that it satisfies the Boltzmann distribution. However, it depends on both v and p because the energy term has both electrical and mechanical terms as expressed by Eqs. 3. If the difference between p_∞ and p is small, the driving force can be proportional to the difference $p_\infty-p$. Thus, the driving force can be expressed

$$f=k\cdot aN(p_\infty-p). \tag{11}$$

The presence of external elastic load K makes the displacement x expressed by $x=aNp\cdot k/(k+K)$, where n is the number of motile elements. By choosing p as the variable of the equation, we have

$$\left[-\omega^2 m+i\eta\omega\right]p=(k+K)(p_\infty-p), \tag{12}$$

which has a familiar form for the equation of motion.

For a given set of values for v and p, the quantity p tends to move to the value that satisfies equilibrium condition, which is given by

$$p_\infty=\beta\overline{P}_\pm(qv+a^2N\tilde{K}p). \tag{13}$$

with $\tilde{K}=kK/(k+K)$ because voltage changes affects length changes as well as charge transfer [11].

With finite transition rates of prestin, the response is to an earlier state of p_∞. The time delay is characterized by ω_g as expressed by Eq. 7b. By introducing the explicit form of p_1, this equation turns into

$$[-(\omega/\omega_r)^2+i\omega/\omega_\eta+1+\delta^2]p=\frac{\beta\overline{P}_\pm}{1+i\omega/\omega_g}\cdot qv, \tag{14}$$

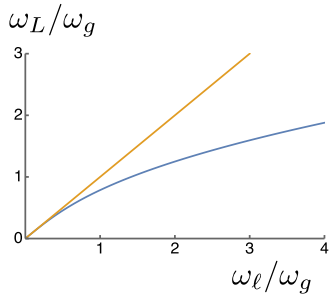


FIGURE 2. Limiting frequency ω_L for prestin with gating frequency ω_g against that of ω_ℓ, for infinitely fast gating. Here, ω_L/ω_g, the limiting frequency normalized with respect to the gating frequency is plotted against ω_ℓ/ω_g (blue). Only in the limit of large ω_g, i.e. for small ω_ℓ/ω_g andω_L/ω_g , does ω_L approach ω_ℓ, the straight orange line.

where $\omega_r^2 = m/(k+K)$, $\omega_\eta = \eta/(k+K)$, and $\delta^2 = \beta\overline{P}_\pm na^2\tilde{K}/(1+i\omega/\omega_g)$. The quantity δ^2 has only a minor effect on p. Notice if we let $\omega_g \to \infty$, factor $1/(1+i\omega/\omega_g)$ turns into unity, we obtain the equation [12] for infinitely fast gating.

In a special case of $m = K = 0$, this equation turns into

$$(1+i\omega/\omega_\eta)(1+i\omega/\omega_g)p = \beta\overline{P}_\pm \cdot qv, \tag{15}$$

showing low-pass behavior with two time constants $1/\omega_g$ and $1/\omega_\eta$.

Power output

Power output $W(\omega)$ by OHCs can be evaluated by calculating the work done against viscous drag because the work against elastic load is recovered after a cycle. The mean power output $\langle W(\omega)\rangle$ is given by $\langle W(\omega)\rangle = (\eta/2)|\omega aNp/(k+K)|^2$.

Eqs. 14, with the observation that δ^2 has a relatively minor effect on p, indicates that the mean power generation $\langle W_\infty(\omega)\rangle$ for the special case of infinitely fast gating differs from $\langle W(\omega)\rangle$ only in the absence of the attenuation factor $1/[1+(\omega/\omega_g)^2]$ if v remains the same. However, the factor $1/(1+i\omega/\omega_g)$ reduces v because it reduces movement of the motile element. For this reason, we may put

$$\langle W(\omega)\rangle < \langle W_\infty(\omega)\rangle/[1+(\omega/\omega_g)^2]. \tag{16}$$

Recall how an optimal limiting frequency was determined for the case, where gating is infinitely fast. In the presence of inertia, power generation has a peak value $P_\infty^{\max}$ due to piezoelectric resonance. A limiting frequency ω_ℓ for that case is obtained from $P_\infty^{\max} = \mu\omega_\ell^2$, equating the maximal power production with viscous loss. Here μ is proportional to viscous coefficient η.

If we can assume that the resonance peak is sharp, the corresponding limiting frequency ω_L for a finite gating frequency ω_g can be approximated by $P_\infty^{\max}/[1+(\omega_L/\omega_g)^2] < \mu\omega_L^2$.

The combination of these two equations leads to

$$(\omega_L/\omega_g)^2 < \frac{1}{2}\left(\sqrt{1+4(\omega_\ell/\omega_g)^2}-1\right). \tag{17}$$

For $\omega_\ell = 2\pi\times 10$ kHz [12] and $\omega_g = 2\pi\times 3$ kHz [8], $\omega_L/\omega_g < 1.1$ (See Fig. 2). The limiting frequency is 3.3 kHz, not much higher than the gating frequency.

The assumption that led to the inequality (17) is not always satisfied because the resonance peak disappears as the gating frequency decreases. In such cases, however, the inequality still holds even though the real frequency limit is lower than indicated because power production is lower without resonance peak.

2. COUPLED OSCILLATORS

The evaluation of maximum power output as mentioned above is based on the assumption that OHCs operate at near resonance frequencies. That assumption in turn requires multiple modes of motion in the organ of Corti because of

the mismatch in stiffness between OHCs and the basilar membrane [12]. In the following, a coupled oscillator is used as a model illustrate this issue.

Here we assume an OHC is associated with an oscillator, which is weakly coupled with another oscillator. It would be useful to examine a system of interacting coupled harmonic oscillators explicitly due to its simplicity, even though the cochlea is not a linear system. The following analysis can be justified only for low input level.

Coupled harmonic oscillators

A textbook example of coupled harmonic oscillators are connected by an elastic element. Typically described by

$$\left(M_1\frac{d^2}{dt^2}+\eta_1\frac{d}{dt}+K_1\right)X_1 = K_c(X_2-X_1) \tag{18a}$$

$$\left(M_2\frac{d^2}{dt^2}+\eta_2\frac{d}{dt}+K_2\right)X_2 = K_c(X_1-X_2)+F_2(t), \tag{18b}$$

where M_j is the mass, η_j drag coefficient, and K_j the spring of each oscillator (for j=1,2). K_c is the stiffness of the spring that connects two oscillators. $F_2(t)$ is the force that stimulates the system. The classical analyses of such coupled oscillators are available [15, 16]. Despite the simplicity of the idea, the frequency dependence of the energy transfer between the two oscillators is rather complex [16].

Amplification with negative drag

The simplest introduction of active or amplifying mechanism to the system of interacting harmonic oscillator would be to flip the sign of the drag term in one of the oscillators.

Let us assume that the first one with displacement x is an amplifying oscillator (AO), i.e. $\eta_1 < 0$, and the second one with displacement y is a dissipating oscillator (DO) and stimulus is applied to the DO. Let the stimulus have a sinusoidal waveform with at an angular frequency ω. Then $F_2 = f_2\exp[i\omega t]$ and the amplitudes of AO and DO are, respectively, x and y.

The equation of motion can be expressed in the form

$$[-(\overline{\omega}/\overline{\omega}_1)^2 - i\overline{\omega}/\overline{\omega}_a + 1 + cs]x - csy = 0, \tag{19a}$$

$$-cx + [-\overline{\omega}^2 + i\overline{\omega}/\overline{\omega}_\eta + 1 + c]y = f, \tag{19b}$$

where frequencies are normalized with respect to the resonance frequency ω_2 (where $\omega_2^2 = M_2/K_2$) of the DO, i.e. $\overline{\omega} = \omega/\omega_2$, $\overline{\omega}_\eta = \omega_\eta/\omega_2$, and $\overline{\omega}_a = \omega_a/\omega_2 (> 0)$. The negative sign on the term involving $\overline{\omega}_a$ indicates negative drag that amplifies the oscillation of the system.

The parameters introduced are $f = f_2/K_2$, $c(=K_c/K_2) < 1$ and $s(=K_2/K_1) < 1$, because it is reasonable to assume the stiffness of the basilar membrane is larger than the elastic load of OHCs. Notice that the ratio x/y is larger than X_1/X_2 because $s\,X_1/X_2 = x/y$.

OHC as amplifier

Eq. 12 for an OHC with respect to the variable p can be rewritten again by using the displacement x, which is defined by $aNpk/(k+K)$, as

$$(-\omega^2 m + i\eta\omega + 1)x = (k+K)aNkp_\infty, \tag{20}$$

assuming the gating of prestin depends on the mechanical factor alone. Here, the component of the receptor potential v of the stimulation frequency ω is determined by $-i_0\hat{r} = (\sigma + i\omega C_0)v - i\omega Nqp$. Here $i_0 = (e_{ec} - e_K)/(\overline{R}_a + R_m)$, where e_{ec} is the endocochlear potential, e_K the resting potential of the OHC, $\overline{R}_a$ the resting level of hair bundle resistance, R_m the basolateral resistance, and C_0 the regular capacitance.

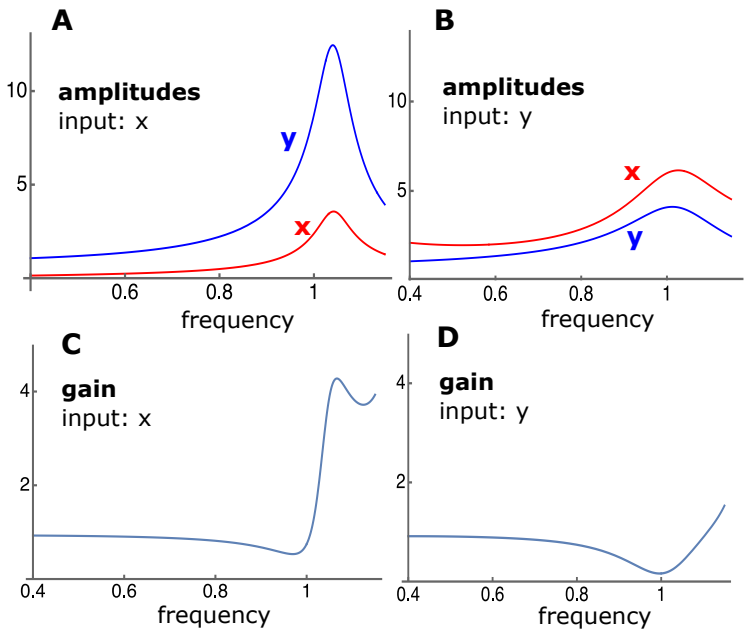


FIGURE 3. Amplitudes and gain plotted against ω/ω_r, the frequency normalized to the resonance frequency. Here both oscillators are assumed to have the same resonance frequency. The amplitudes are normalized to external force f. Gain is the square of ratio of the amplitude y while OHC is engaged to the amplitude y while OHC is not engaged. Left (A and C): input to the hair bundle is x. Right (B and D): input to the hair bundle is y. Gain is the ratio of the squared amplitude of y relative to the condition, where OHC is turned off. The reduction of the gain near reduced frequency unity is associated to a positive peak shift due to OHC, i.e. the effect of the parameter B in the equations. Parameter values are, $\omega_1 = 1$, $c = 0.1$, $k = 5$, $\overline{\omega}_\eta = 10$, $s = 0.01$, $B = 0.57$, $A = 1.5$.

The quantity $\hat{r}$ is the relative change of the hair bundle conductance. If $\hat{r}$ depends on x, the equations of motion can be expressed as

$$[-\overline{\omega}^2 - iA/\overline{\omega} + 1 + cs - B]x - csy = 0, \tag{21a}$$

$$-cx + [-\overline{\omega}^2 - i\overline{\omega}/\overline{\omega}_\eta + 1 + c]y = f, \tag{21b}$$

where A and B are determined by the operating point. This set of equations is similar to the one with anti-drag amplifier, Eq. 19a even though that the frequency dependence of the amplifier term is different.

If the relative change $\hat{r}$ of hair bundle resistance depends on y, the other oscillator, the set of equations take another form:

$$[-\overline{\omega}^2 + 1 + ck - B]x - [iA/\overline{\omega} + cs]y = 0, \tag{22a}$$

$$-cx + [-\overline{\omega}^2 + i\overline{\omega}/\overline{\omega}_\eta + 1 + c]y = f. \tag{22b}$$

Some of numerical examinations of these equations are shown in Fig. 3.

Numerical examination

Having identified variable x as associated with OHC, and y with the BM to derive Eqs. 21 and 22, it could be more convenient to call the two oscillators as OHC-associated and BM-associated rather than amplifying and dissipating oscillators.

Fig. 3 shows that the sensitivity of the hair bundle to OHC displacement is critical for the amplifier gain (Fig. 3C) near the resonance frequency and that the amplitude of OHC displacement could be much smaller than that of the BM. This result appears consistent with experimental observations [9]. However, the small amplitude of OHC movement does not increase at lower frequencies and it is not consistent with the reported experimental observation [9].

If the hair bundle is driven by the BM, the amplitude of OHC displacement is larger. At the same time the gain is smaller. This condition could be regarded as similar to the experimental observation at low frequencies. Thus, the experimental observations can be explained if we can assume that the hair bundle is stimulated by the BM at low frequencies and by OHC displacement near the characteristic frequency.

This idea is incompatible with the assumption that hair bundle displacement is a weighted sum of BM displacement and OHC displacement throughout the frequency range. That is because the much smaller amplitude of the OHC needs to dominate hair bundle stimulation to make the amplitude of the BM larger.

CONCLUSIONS

The present analysis is based on a set of simple coupled oscillators, which likely over-simplifies possible modes of motion in the cochlea. Indeed, the organ of Corti is an anisotropic, heterogeneous-three dimensional object. In addition, this analysis ignores nonlinearity, characteristic to the cochlea. These conditions therefore limit the validity to low stimulation intensity. With these reservations, the following conclusions can be drawn.

The intrinsic transitions of prestin must be fast enough to exceed the upper bound of the auditory range for electromotility to counteract local drag. Even though energy can flow laterally along the cochlea, a large gap between the intrinsic gating frequency and the best frequency of the location makes it unlikely for OHCs function as the cochlear amplifier.

A simple local coupled oscillator model shows that OHCs can function as cochlear amplifier, at frequencies lower than the gating frequency of prestin, even if the amplitude of OHC length changes is much less than the amplitude of basilar membrane movement, consistent with OCT data [9].

For these reasons, low frequency gating of prestin in recently reports [7, 8], which contradicts older reports [5, 6], is not compatible the OCT data [9] because low frequency gating is incapable with amplifying effect of OHCs at best frequencies.

The model presented by itself cannot explain the amplitude roll-off of OHC movement as observed with the OCT [9]. However, the present analysis suggests that such a roll-off can be explained as the consequence of mode changes. Given the observed dependence of the mode of motion of the organ of Corti [17, 18], one can argue nonetheless that the modes of motion at low frequencies are indeed different from that of best frequencies. Verifying this expectation requires a detailed analysis of modes of motion *in vivo* and their dependence on the frequency.

REFERENCES

1. P. Dallos, X. Wu, M. A. Cheatham, J. Gao, J. Zheng, C. T. Anderson, S. Jia, X. Wang, W. H. Y. Cheng, S. Sengupta, D. Z. Z. He, and J. Zuo, "Prestin-based outer hair cell motility is necessary for mammalian cochlear amplification." Neuron **58**, 333–339 (2008).
2. K. H. Iwasa, "Effect of stress on the membrane capacitance of the auditory outer hair cell," Biophys. J. **65**, 492–498 (1993).
3. J. E. Gale and J. F. Ashmore, "Charge displacement induced by rapid stretch in the basolateral membrane of the guinea-pig outer hair cell," Proc. Roy. Soc. (Lond.) B Biol. Sci. **255**, 233–249 (1994).
4. X. X. Dong, M. Ospeck, and K. H. Iwasa, "Piezoelectric reciprocal relationship of the membrane motor in the cochlear outer hair cell," Biophys. J. **82**, 1254–1259 (2002).
5. G. Frank, W. Hemmert, and A. W. Gummer, "Limiting dynamics of high-frequency electromechanical transduction of outer hair cells," Proc. Natl. Acad. Sci. USA **96**, 4420–4425 (1999).
6. X. Dong, D. Ehrenstein, and K. H. Iwasa, "Fluctuation of motor charge in the lateral membrane of the cochlear outer hair cell." Biophys J **79**, 1876–1882 (2000).
7. J. Santos-Sacchi and W. Tan, "The frequency response of outer hair cell voltage-dependent motility is limited by kinetics of prestin." J. Neurosci. **38**, 5495–5506 (2018).
8. J. Santos-Sacchi, K. H. Iwasa, and W. Tan, "Outer hair cell electromotility is low-pass filtered relative to the molecular conformational changes that produce nonlinear capacitance." J. Gen. Physiol. **151**, 1369–1385 (2019).
9. A. Vavakou, N. P. Cooper, and M. van der Heijden, "The frequency limit of outer hair cell motility measured in vivo." eLife **8**, e47667 (2019).
10. K. H. Iwasa, "Current noise spectrum and capacitance due to the membrane motor of the outer hair cell: theory," Biophys. J. **73**, 2965–2971 (1997).
11. K. H. Iwasa, "Energy output from a single outer hair cell." Biophys. J. **111**, 2500–2511 (2016).
12. K. H. Iwasa, "Negative membrane capacitance of outer hair cells: electromechanical coupling near resonance." Sci. Rep. **7**, 12118 (2017).
13. K. H. Iwasa, "Kinetic membrane model of outer hair cells," Biophys. J. **120**, 122–132 (2021).
14. The actual axial stiffness is reduced by conformational transitions of the motile elements [19]. This effect is analogous to "gating compliance" [20].
15. P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton University Press, Princeton, New Jersey, 1968) p. 927.
16. C. A. Mercer, P. L. Rees, and F. J. Fahy, "Energy flow between two weakly coupled oscillators subject to transient excitation," J. Sound Vibr. **15**, 373–379 (1971).
17. S. S. Gao, R. Wang, P. D. Raphael, Y. Moayedi, A. K. Groves, J. Zuo, B. E. Applegate, and J. S. Oghalai, "Vibration of the organ of corti within the cochlear apex in mice." J. Neurophysiol. **112**, 1192–1204 (2014).
18. N. P. Cooper, A. Vavakou, and M. van der Heijden, "Vibration hotspots reveal longitudinal funneling of sound-evoked motion in the mammalian cochlea." Nature Commun **9**, 3054 (2018).
19. K. H. Iwasa, "Effect of membrane motor on the axial stiffness of the cochlear outer hair cell," J. Acoust. Soc. Am. **107**, 2764–2766 (2000).
20. J. Howard and A. J. Hudspeth, "Compliance of the hair bundle associated with gating of mechanoelectrical transduction channels in the bullfrog's saccular hair cell," Neuron **1**, 189–199 (1988).

Digitization and Reanalysis of Fiber-optic Displacement Data: An Elephant in the Chamber?

Eric L. LePage

PO Box 2564, Mount Claremont, Western Australia, 6010, Australia
Corresponding author: ericlepage@innerearmechanisms.org

Abstract. Over the last half century, the intense activity upon the mechanics of mammalian hearing falls into two broad categories: 1) Studies by sensory physiologists (e.g. Brownell, Flock, Russell) interested in the holistic active mechanical properties of the outer hair cells (OHC) and 2) studies by theoreticians and practitioners who have been attracted into the field seeking to explain the central question of audition, i.e. What accounts for the extreme sensitivity and vulnerability of mammalian hearing? How can the critical mechanism which goes missing *not* be best described as amplification, if, when sensitivity is lost, it can be compensated by use of an external amplifier? The historic starting point has been the Bekesy two-chamber model of the cochlea. The general, the Fourier transform of basilar membrane (BM) vibrations would normally have a zero-frequency (or dc-) component. However, because of the historic basis of the standard model, any such component has been systematically disqualified from consideration. This presenter was the first to publish that a dc-component of BM motion not only does exist, but it covaries with the vibrational component in ways which mimicked the polarity variation of the summating potential (LePage JASA,1987). The later published data (full method details, LePage, Hear.Res., 1989) did not have the full support of all collaborators beyond the development of the fiber-optic sensor. However, using the finished device, it was shown, to the real-time disbelief of B.M.Johnstone, that the BM displacement displayed properties very similar to the contractile components of the OHC both in temporal dependence and amplitudes equalling fractions of the length of the OHC. These fiber-optic data have been extracted from many paper chart records, digitised, then reprocessed with MATLAB(TM) to generate a video to bring to new life what was originally only rapid large excursions of the pen on the single channel chart recorder. These BM dc-shifts are indeed highly tuned. Their tuning varies as expected with the manipulations originally described. Moreover, the velocity of these excursions is of the same order with the estimates of velocity of BM vibrations from dozens later investigations which have uniformly assumed the unseen motion was exclusively a vibration at the stimulus frequency. This presentation therefore raises deep questions about whether the two-chamber model has remained a dominant influence on the progressively-refined design of BM measurements – as well as the widespread rejection of data from many studies whose results made no sense in conventional terms of a fragile and "mysterious" cochlear amplifier (CA). Could it be that the expected super fragile tuning mechanism was the oft-employed explanation, when an equally plausible explanation could be a robust mechanism based on the OHC, modulated by an unknown but critical tiny BM baseline bias?

BACKGROUND

There have been two streams of experimental investigation into basilar membrane motion in mammals – those preferring to use velocity sensors and those using displacement sensors. The velocity sensors delivered much cleaner data particularly coupled with liberal use of standard noise-rejection techniques. By contrast the displacement measures of basilar membrane motion produced such confusing data they have been widely dismissed, particularly because they tended to need invasion of the cochlea to get access to the basilar membrane. Such a disturbance was unacceptable essentially because it tends to destroy the high sensitivity and sharp tuning which is the target objective: 1) vibrations due to incoming sound at threshold levels are submicroscopic demanding very high sensor sensitivity, and 2) the tuning mechanism appears very delicate – it can be disturbed by a host of influences, each of which must be carefully controlled. In a half-century of achievement, the source of OHC motility has been tied to contractile behavior of isolated OHC soma easily observed under a regular microscope (Brownell et al 1985). The unresolved issue has been how that is tied to hearing sensitivity. On the

one hand negative damping needs a motor to give a fast cycle-by-cycle kick to the motion, but on the other hand, nature seems instead to have supplied a motor which can deliver large but relatively slow displacements. This is due to the OHC membrane time constant being so long that even if the OHC stereocilia were faithfully following high frequency sounds, in bats and cetaceans, the OHC motor could not respond quickly enough.

There appears to exist a pair of non-intersecting cycles keeping two sets of activity apart: 1a) The vibrations are submicroscopic requiring high sensitivity 1b) the sensitivity can be increased arbitrarily if low-frequency components of the motion are excluded, for which velocimeters are ideally suited and 1c) the approach typically needs copious artifact rejection methodologies, 1d) results in very clean data which modelers need. Alternately, 2a) the allowance of low-frequency components of the motion using displacement techniques result in labile behavior, 2b) it cannot easily be categorized as vibrations, 2c) it does not fit with the standard model, 2d) it has been taken almost universally as a non-specific measure of pathology of the preparation, 2e) many years of promising physiological data have been summarily dismissed because they did not seem to be attacking the core of the problem: the mechanism of undamping.

The first cycle has led to e.g. the need for a Hopf bifurcation. For all of its theoretical relevance this realization has not provided a ready avenue for general progress. By contrast, the second set seems rich with evolved physiological behaviors like muscle contractions and essential homeostatic control mechanisms, such behaviors also observed extensively by Brundin et al (1991, 1992), LePage et al (1993) and Zenner et al (1994).

This investigator's history is relevant to this report. In the early days the favored approach was the Doppler-shift Mössbauer technique. It proved unsuitable for a project requiring to control basilar membrane waveform for neural experiments. Instead, a displacement-sensitive capacitive probe was built and waveform control easily achieved. That project also confirmed the existence of the expected compressive nonlinear behaviour (LePage and Johnstone, 1980). A capacitive probe requires draining scala tympani for the measurement. However, an advantage of this technique is that using a low frequency complex periodic stimulus, the electrical output of the probe can be also monitored with headphones. It provides a much more intimate real-time connection of the investigator with the experiment. When the sound level was varied, within a minute the basilar membrane started moving toward the probe tip – observed in hundreds of experiments. Contact with the film of perilymph shorted-out the probe often requiring restarting the run. The Ockham's Razor view was that the perilymph was refilling. There was really only one way to dispute that assessment – a new displacement technique was needed which could operate without draining. Nonetheless, a second set of capacitive probe experiments took place at Washington University in which the summating potential (SP) was recorded simultaneously on a second channel (LePage 1987b). This series provided a good control that these preparations were not unduly compromised because the nonlinear behaviour, indeed the peaked tuning was pronounced. Also, it was clearly demonstrated that the SP correlated with mechanical responses. Moreover, another important finding was of baseline displacements, not only occurred during the tone bursts, (recently observed by Dewey etal, 2021), but also in the longer term.

By courtesy of the late Dr. B.M. Johnstone, a fiber optic lever (FOL, Cook and Hamm, 1979) was constructed in Perth, and improved. It was calibrated and tested to measure basilar membrane displacements in guinea pigs. The progression to a comprehensive project did not eventuate and so the tryout was preliminary. But in just five weeks it was demonstrated that the guinea pig basilar membrane responses were not merely "vibrations". Despite the rich set of physiological and pharmacological manipulations reported, the extensive review by Robles and Ruggero (2001) neglected to comment on any of the data. Perhaps, again this was likely because of the widespread assessment that the displacement techniques were unsuitable for studying threshold phenomena.

After thirty-three years of attempts to understand the cochlear amplifier (Ashmore et al, 2010) it seems timely to re-examine these old displacement data for any clues that nature managed to get around the speed problem with less mystery. A fresh hypothesis driving this reanalysis is to assume that the speed problem simply does not exist; it may be a manifestation of the heavy filtering needed to find cycle-by-cycle kicks amongst the Brownian noise.

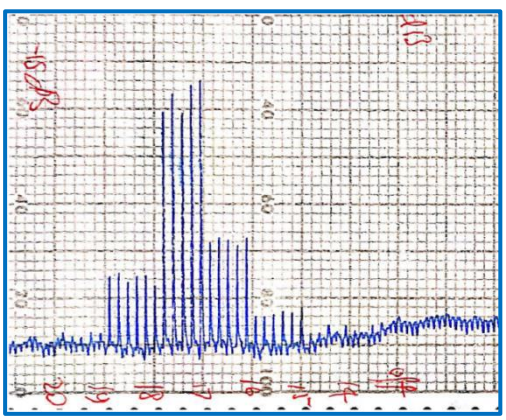


Figure 1 Early time-series chart record. The chart speed is 6 seconds per cm) from the basal turn of guinea pig GP87-13 (1987) from fiber-optic lever (FOL), showing pulsations in the transverse position of the gold foil mirror resting flush upon the basilar membrane, under perilymph. Full scale is 1.5µm displacement determined by moving the probe mounted on a Narishige micromanipulator towards the BM. Positive-going displacement on the chart record indicates the gap between the probe tip and the small gold mirror resting on the basilar membrane is decreasing, i.e. the displacement is towards scala tympani. Stimuli are a series of 5 individual tone bursts having 1ms rise and fall times, in 1kHz steps from 20kHz down to 14kHz. Tone level is 105dB SPL less 15dB, 90dB SPL. Despite mimicking the protocol and the IHC data of Russell & Sellick (1978) these pulsations were deemed orders too large. The tiny ripples are due to respiration.

METHODS

This report is a revisitation of the many interesting data generated (LePage, 1989), to draw out one relevant result and to encourage repetition using the *same long-range displacement* methodology, on the basis that the abundant criticism displacement methods have received is here invalidated. All relevant details and calibration are contained in the original publication. The motion of the reflecting surface varies the light returned which is quantified with an infrared detector. The device was mounted on a Narishige micromanipulator to allow positioning of the sensor tip at precisely controlled distances from the vibrating surface. A mirror of gold foil approximately 60µm square was placed central onto the BM in the basal turn, within the perilymph, placement using a guinea pig hair. With a binocular microscope the mirror was checked to lie flush with the BM before the experiment and rechecked afterward. The probe was aligned normal to the gold mirror. For this arrangement the sensor had an input noise level equivalent to 1nm vibration. Calibrated pure tones were delivered from to the right ear bar and sound levels were monitored with a probe microphone which ranged from 60dB to 110dB SPL.

The basilar membrane displacement was recorded in real time on a chart recorder without signal averaging nor other artifact rejection. The selected chart speed was 6cm per minute, The pure tone generator could be programmed to deliver cosine-shaped tone bursts with frequency increasing or decreasing 1kHz steps, automatically or each time the investigator hit an advance button.

Historic chart records like that shown in Figure 1 were digitized then reprocessed with MATLAB®. It was possible to expand the records on a large monitor screen for precise selection of data points.

RESULTS

In the figures to follow are charted two records of dc-displacements from guinea pig (GP87-17 and GP87-13). Placing the mirror on the bony spiral lamina produced no comparable results. A protocol was used, similar to that used by Russell and Sellick (1978) for IHC potential recordings for tone bursts.

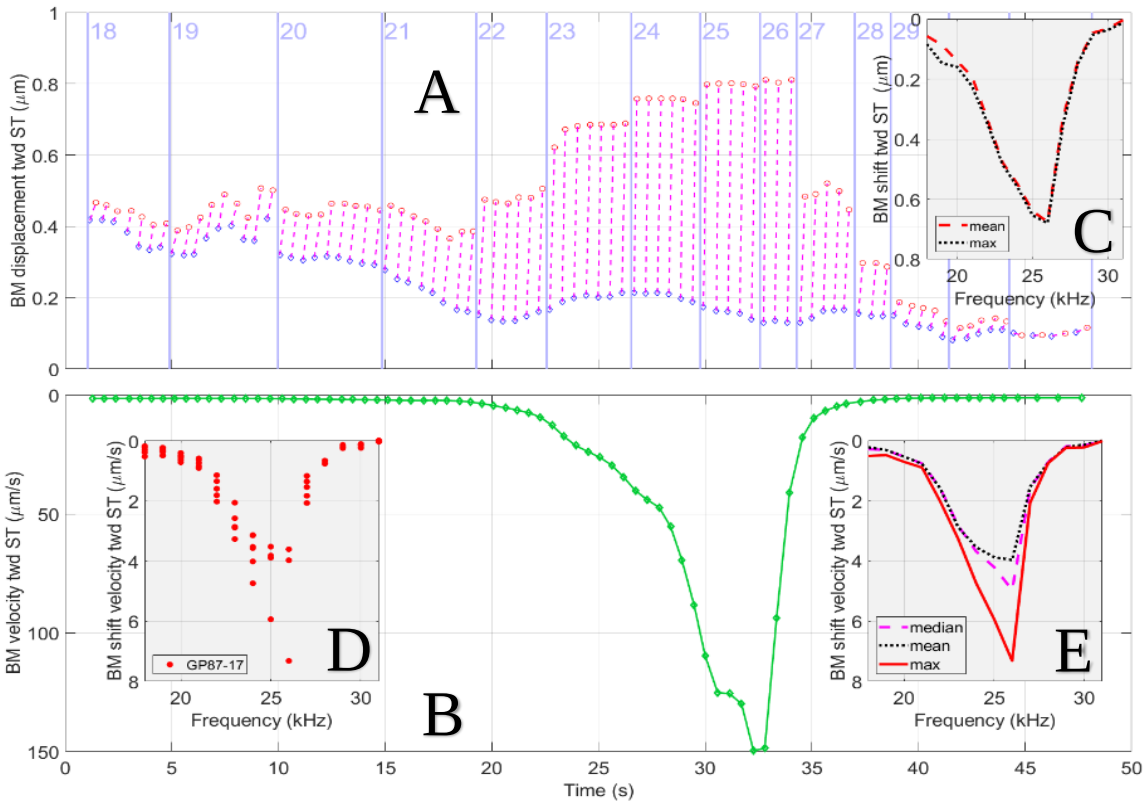


Figure 2 The wide top and bottom panels (Panels A and B) are time series records. Shown are the steep baseline shifts for frequencies ranging from 18 to 31 kHz. The position of the basilar membrane at tone onset is represented by the blue diamonds, while the red circles represent the position when the basilar membrane abruptly reaches its maximum displacement. The dashed red lines represent a linearized track between the end points of the rapid displacement at tone on. The return to baseline displacement during tone off is omitted for clarity. The flanks of the green tuning curve representing transition velocity (axis inverted) display an asymmetry which is characteristic of neural tuning curves (LePage, 1987c) despite this not being obtained as an iso-response curve. The insets Figs. 2C 2D and 2E show measures of central tendency as indicated.

Figure 3 also shows that the maximum displacement is frequency-dependent. As discussed below the response is "tuned", but not in the sense of a vibratory resonance. It is tuned in the sense of a place-localized peak in the basilar membrane baseline pulsations with frequency (Panel C). The slope of each rapid transition (giving the velocity) is also changing (diamonds connected by the green trace in Panel B, smoothed with a moving average of 11 points). These changes of velocity are also displayed versus frequency in the lower insets which are shown with an inverted scale. The left inset (D) gives the computed velocity for individual bursts, while the lower right inset (E) shows the median, mean and maximum velocity (dashed, dotted and continuous line) for each set.

Figure 4 presents the same data presentation for another guinea pig (GP87-13) as Fig 3 drawing out the change in baseline with change in frequency. Panel 4A shows the raw time series record plotted as tone ON and tone OFF responses as well. In the original publication LePage, 1989 (Fig.4) the exponential rises and falls was shown.

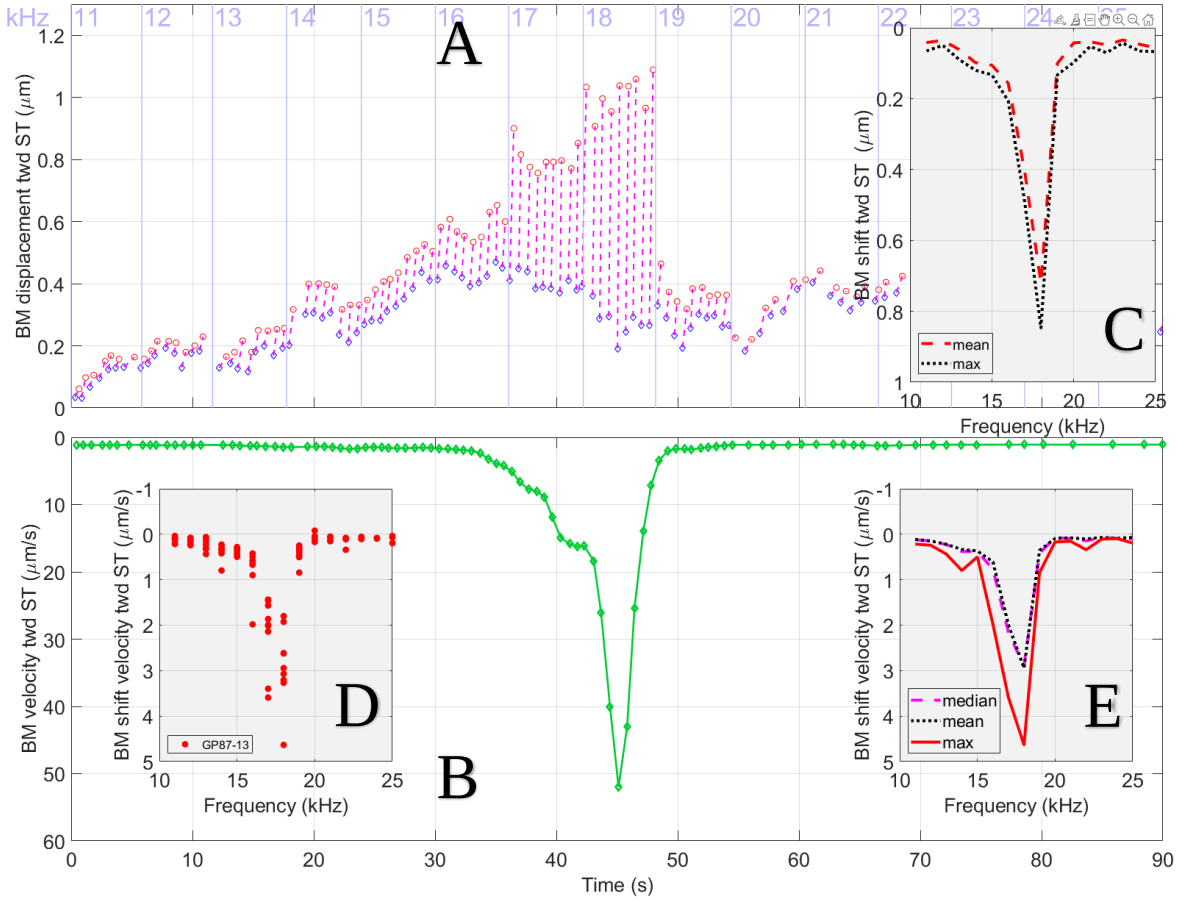


Figure 3 For guinea pig GP87-13 obtained similarly to the record in Fig. 2 for a similar set of raw (unaveraged) baseline excursions. In this case there are exactly 10 bursts each frequency from 11 to 25 kHz. The best frequency for dc-shifts is 18 kHz. The baseline begins with a low bias value at low frequencies. Again, as the BF is approached the baseline rises. There are marked changes in the upper limiting value as well as the slopes of the pulsations (curves inverted). A feature of neural and mechanical tuning curves is the so-called cutoff which phenomenon is prominent in these data worthy of replication.

Some of these bursts shows a clearer exponential decay than others. Panel 4B shows the different frequencies color coded with the value of the baseline highlighted as a dashed line. In Panel 4C the baseline is subtracted giving a response much more consistent with the results of velocimeter experiments as well as the expectations of the two-chamber model. In Panel 4D these tone-burst responses from Panel 4B are averaged *synchronous with the start of each tone burst* and the result expanded to fill each box. In Panel 4E these are differentiated to give the velocity of the averaged response which at the BF is >5 µm/s and clearly peaking at the end of the rise time.

Figure 5 shows a 3-D color-coded scatterplot, same data, vs starting baseline value, one point per tone-burst.

DISCUSSION

1) No signal averaging is used because these signals were far in excess of the noise level of the FOL. 2) The small pulsations on the baseline are respiratory movements. 3) For each tone burst of the same frequency the excursions rise from whatever was controlling the pre-burst baseline position to a common saturated value almost as if it had hit a hard stop different for each frequency. 4) Although basilar membrane displacement is by no means similar

to the iso-response parameters of velocimeter tuning curves, the similarity of the form is worthy of further investigation. 5) The baseline value is changing during the series due to factors which will be critical to determine for a complete explanation (LePage 1987). 6) Possibly the most prominent feature of this set of data is the cutoff region mimicking that of neural and velocimeter measurements. 7) At each frequency these *long-range* displacement responses are reminiscent of the twitch responses of the frog gastrocnemius muscle given a transient stimulus. The strongest twitch occurs at the best frequency (BF) of the place (18kHz). 8) These responses, may now fairly safely be assumed to be due to the somatic contractility of prestin (consistent with the macro-responses in the original videos of Brownell et al, 1985). 9) They are clearly tuned, but not in the sense of a vibratory resonance, but as a *place-specific OHC length change*. This does not imply there is not a resonance.

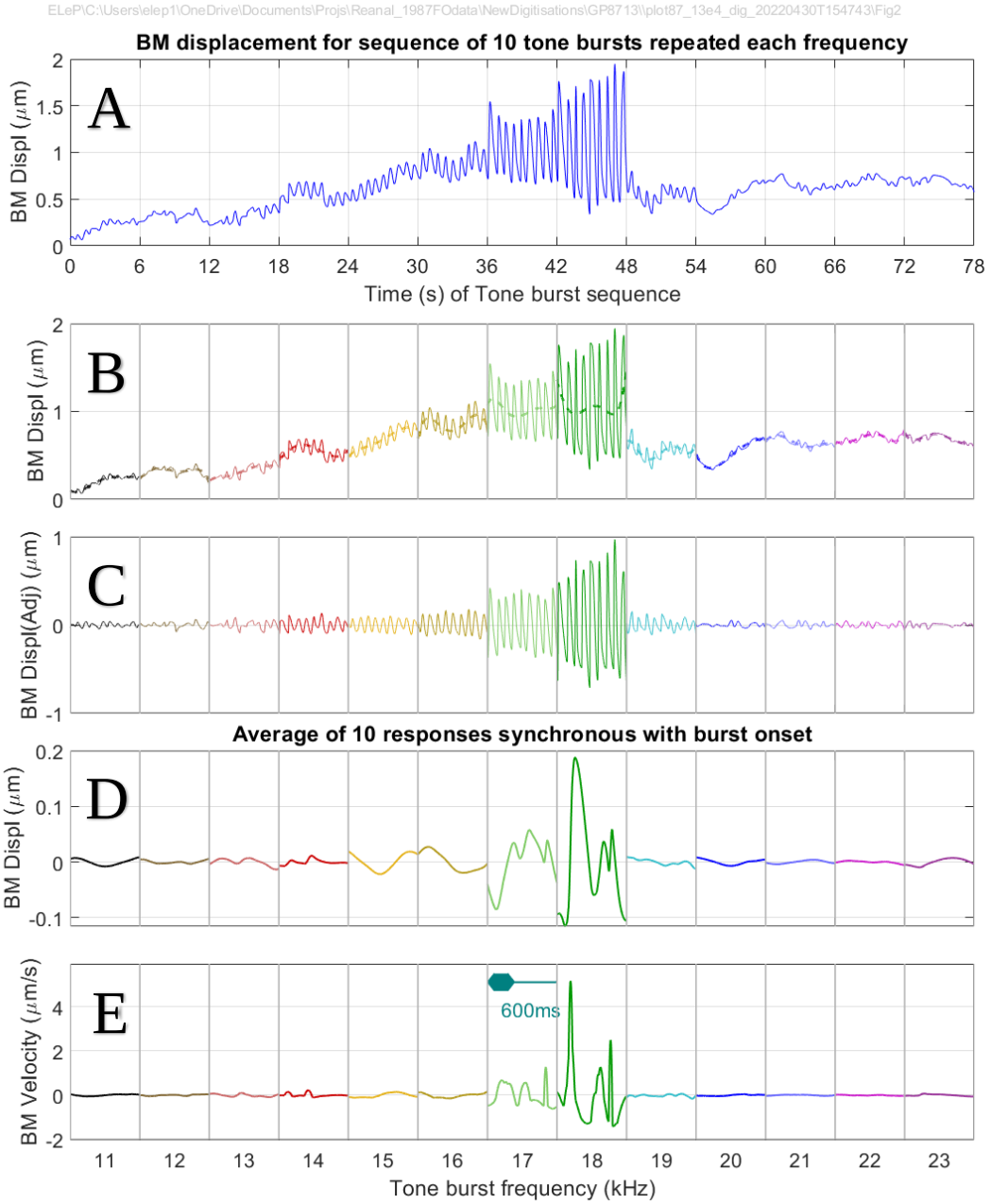


Figure 4. The same data as Fig. 3 showing it as time series in top panel, then showing against frequency axis. In the top three panels are individual tone bursts showing both rises and falls; the bottom two panels show the set of averages of the ten bursts and their time derivatives. In this series ten tone bursts are presented sequentially at each frequency (timing of bursts shown in panel A). Panels B through E are color coded by frequency and the same coding carried through to the scatterplots in Fig 5. The superimposed dashed line represents the moving average baseline (11-point) during each series of tone bursts. In Panel C this moving average is subtracted leaving an apparently high-pass filtered version of the mechanical response -- much more in keeping with the expectations of velocimeter measurements. In Panel D the time series displacement responses of each tone burst with the same frequency (from Panel C) are averaged together showing only the deviations from the starting baseline value. Panel E shows the time-derivatives of the averaged displacement responses in Panel D against the envelope of the tone burst. Clearly the peak velocities correlate with the tone-burst onset risetimes.

However, it does suggest that any resonance does not involve prestin. 10) Instead, it challenges the usual idea that the fragility of tuning is due lack of robustness of the structure or the OHC when the key reason may merely be a shift in the bias setting the operating points and affecting sensitivity. 11) In keeping with explants, this manifestation of robust activity has not been severely compromised by draining of the cochlea either – it has simply modified the bias values needed to view pronounced compression. 12) These data do not support the notion of a vibratory cochlear amplifier. Instead, Fig. 4E suggests, that the adequate stimulus for OHC motility is the *onset* of a sound at the appropriate place which is much more in keeping with conventional muscle physiology. 13) The important outcome of both sets of displacement studies is the suggestion that the pre-occupation with cochlear amplification of a miniscule vibration has influenced the design of experiments which have continued to assume that high sensitivity is necessary to see basilar membrane vibrations. 14) The polarity of these motor responses was shown to switch with frequency (LePage, 1987b). 15) This fits with assertions of the bipolarity of flexoelectric force generation within the soma of the OHC (Brownell and Manis, 2014).

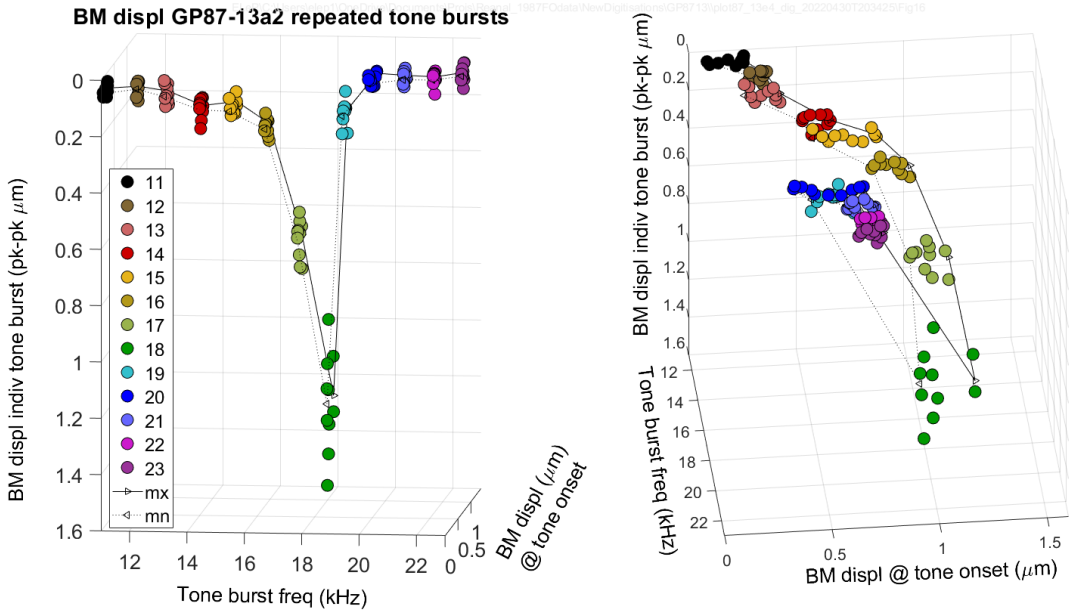


Figure 5 3-D Scatterplots of the same data set shown in Figs 3 and 4. The left and right panels present the same data. The right is rotated about the vertical axis so as to reveal the mechanical bias needed for sharpest tuning. The displacement (Z)axis is shown inverted, consistent with Figs 2 and 3 in the manner of iso-response neural tuning curves. The size of the BM shift at the tuning peak at 18kHz appears most strongly dependent upon the baseline displacement at the onset of the tone burst. The magnitude of the transverse BM displacement appears enhanced with a baseline offset of 0.5µm at tone onset.

CONCLUSIONS

Cochlear tuning need not limited to the way it has always been recorded i.e., at low sound levels, plotted in acoustic terms with an associated Q10dB value. Based upon these fiber-optic data a new idea is tendered. It is that *the "cochlear amplifier" is not necessarily a vibration (ac)-amplifier*. Here sharp tuning is manifest as a "macro-contractile" response, localized to central frequencies defining that peak. When these results and those of Brundin were considered, they were rejected because they were not vibrations. These few fiber-optic data are sensitive to OHC contractions as they affect the position of the basilar membrane. They suggest it may be profitable to reinvestigate basilar membrane displacement with a method which can span up to 10µm at speed.

When Gold provided a solution to this speed problem, he had no choice but to accept the severe constraints of the two-chamber model of von Békésy. His creative solution set a generation of investigators hunting for a submicroscopic mechanism buried in noise. Alas, the extraction of expected waveforms left no choice but to assume the response had the form of a vibration at the stimulus frequency. Without special attention any data produced by phase-locking of the data collection with the sinusoidal stimulus will not include baseline shifts, and therefore could be used to support the cochlear amplifier concept and lead to overconfidence that the description is complete.

NEW POSSIBILITIES

The hair cell responses obtained by Russell and Sellick (1978) are not inconsistent with the idea that the OHC sensory response produces OHC length change which is the adequate stimulus for the IHC stereocilia. These data lead suggest that cochlear homeostasis may incorporate a relatively slow signaling system which may also modulate the radial tension in the basilar membrane produced by the fibrocytes of the spiral ligament (Furness, 2019). These data not only suggest a resolution of the speed problem, they do so with an ironic twist: fluid damping may be an advantage, serving to isolate analog and "flagging" modes in top and bottom layers, from each other.

SIGNIFICANCE

The human ear has a set of sensory hair cells called the outer hair cells (OHC), which are progressively depleted throughout life with age or noise exposure leading to hearing loss. Dr. Hallowell Davis, the "father of auditory science" wore hearing aids. He wrote a paper in 1983 which made use of a new finding back then that the OHC are not just sensory, they had in-built motors as well. The paper made a lot of sense: the motors must have evolved to act as an internal amplifier to boost the vibration. That idea was supported by the best data describing vibrations of the basilar membrane. The problem is that three decades later, modelers have not been able to resolve long-term puzzles in cochlear mechanics as well as in auditory psychophysics. This paper revisits mechanical data directly measured in live animals in 1987 for new clues. It was an unconventional approach because it did not cut out slow baseline changes which could be happening in the basal turn (high frequency region). The approach revealed that the action of the OHC motors may not be limited to boosting vibrations. Instead, the OHC length changes may be 1) involved with fine regulation of the resting position of the BM and 2) they may directly flag that their stereocilia have detected sound, by driving the tectorial membrane and the IHC a bit like "fingers playing on a keyboard" – with a speed not so fast as to be significantly damped by the fluid.

ACKNOWLEDGEMENTS

I am indebted to very many people, particularly the late Dr. Brian Johnstone for his many insights and two periods of sponsorship. The most significant encouragement came from the late Dr. Charles Molnar and his team at Washington University. I enjoyed many lunch discussions with the late Dr. Hallowell Davis who kindly offered me co-authorship of his "cochlear amplifier" paper but graciously accepted my alternate view. I thank the late Dr. David Mountain at Boston University, Prof. Dr. Hans Peter Zenner at the University of Tübingen and Dr. Reginald Cook for assisting my reconstruction of this fiber-optic lever.

REFERENCES

1. Ashmore et al (2010). The remarkable cochlear amplifier. Hear.Res. 266, 1-17.
2. Brownell etal, (1985). Evoked mechanical responses of isolated cochlear OHC. Science, 227: 194-196.
3. Brownell and Manis (2014) Structures, mechanisms and energetics. Perspectives on Aud Research.Ch2.
4. Brundin, L. et al (1991). Frequency-specific position shift in the GP organ of Corti. Hear.Res.128, 77-80.
5. Brundin, L. et al (1992). The tuned displacement response of the hearing organ. Neurosci. 49, 607-616.
6. Cook and Hamm (1979). Fiber-optic displacement transducer. App.Optics, 18, 3230-9.
7. Davis, H. (1983). An active process in cochlear mechanics. Hear.Res. 9, 79-90.
8. Dewey, J.B. etal (2021). Cochlear OHC electromotility enhances a cycle-by-cycle motion. PNAS 118,43.
9. Furness, D.N. (2019). Forgotten fibrocytes: a neglected supporting cell type. Frontiers.CellNeurosci,13,532.
10. Gold, T. (1948). Hearing. II. The physical basis of action of the cochlea. Proc.Roy.Soc.(Lond)SerB,492-8.
11. Guinan et al (2005). Medial-olivocochlear-efferent inhibition. J.Acoust.Soc.Am. 118,2421-2433.
12. LePage and Johnstone, (1980). Nonlinear mechanical behaviour of the basilar membrane.Hear.Res.2,183-9.
13. LePage (1981) Ph.D. thesis, The role of nonlinear processes. The University of Western Australia. In: The mechanics of cochlear homeostasis, 2006, OAEricle Australia ISBN 0-9775547-0-8.
14. LePage (1987a). The application of a capacitive probe technique. J.Acoust.Soc.Am. 82, 126-138.
15. LePage (1987b). Frequency-dependent self-induced bias of the basilar membrane and its potential for controlling sensitivity and tuning in the mammalian cochlea. J.Acoust.Soc.Am. 82, 139-154.
16. LePage (1989). Functional role of the olivo-cochlear bundle: A motor unit system. Hear.Res.38, 177-198.
17. LePage etal (1993). Summating baseline shifts with two optical displacement-sensing techniques. In Proc. Biophysics of hair cell sensory systems (In: MoH series Paterswolde, The Netherlands, 240-248).
18. LePage (2018). A role for the otoliths in the mechanics of cochlear homeostasis? AIP Conf Proc.
19. Robles and Ruggero, (2001). Mechanics of the mammalian cochlea. Physiol Rev. 81, 1305-1352.
20. Russell and Sellick (1978). Intracellular studies of hair cells in mammalian cochlea. J.Physiol,284,261-290.

High Speed Imaging of Active Motility in Hair Cells

Martín A. Toderi,[1, a)] Dzmitry Vaido,[1, b)] and Dolores Bozovic[1, 2, c)]

[1)] *Department of Physics & Astronomy, UCLA, Los Angeles, California, 90095, USA.*
[2)] *California NanoSystems Institute, UCLA, Los Angeles, California, 90095, USA.*

[a)] *Electronic mail: mtoderi@physics.ucla.edu*
[b)] *Electronic mail: dvaido1@physics.ucla.edu*
[c)] *Corresponding author: bozovic@physics.ucla.edu*

Abstract. Hair cells of the auditory and vestibular systems are capable of detecting sub-nanometer deflections due to air-borne or ground-borne vibrations. Further, hair bundles of some end organs are known to spontaneously oscillate without external stimulation, indicating the presence of internal active processes. Ionic influx upon the deflection of hair bundles elicits changes in the somatic potential, leading to a conversion of mechanical input to electrical signals. While the transduction process has been well-studied in hair cells, the possible presence of mechanical distortions in the hair cell soma has not been explored in non-mammalian species. In this work, we applied optical techniques based on transmitted and scattered light to detect active motility in the hair cell soma under different experimental conditions. We developed image processing tools tailored for semi-intact preparations of the sensory epithelium, observed with wide-field microscopy. Accordingly, we recorded light intensity distributions of hair cells of the American bullfrog (Rana catesbeiana) at high speed during spontaneous oscillations and mechanical deflection of the hair bundles. We imaged the hair cells at various optical sections along the z axis, with a specific focus on the innervation area, and observed mechanical fluctuations at localized positions within the hair cell. We aim to further develop optical techniques that would enable the study of hair cells somatic activity with a non-invasive and dye-free approach, so as to minimize mechanical or physiological disturbance of natural conditions.

INTRODUCTION

The auditory system plays a vital role in living organisms, allowing them to navigate in space and communicate and providing important cues for the avoidance of danger. These features require not only sensitivity that is sufficient to detect signals as low as 0 dB [1] but also the ability to extract that information from environments containing multiple competing streams of information. Specialized cells located in internal sensory organs within the ear are responsible of the hearing and balance capabilities of vertebrates [2]. The key functional units performing the first step of detection are known as hair cells (HCs) due to their characteristic stereocilia that emerge from a cuticular plate (Fig. 1(A)). The hair bundle is composed of 30 to 50 stereocilia that protrude from the apical surface of the cell body, which is embedded in the surrounding tissue. HCs of the auditory and vestibular systems are finely tuned sensory detectors, capable of perceiving sub nanometer mechanical perturbations of the hair bundle. These displacements can be induced by air- or ground-borne vibrations that are transduced into electrical signals [3]. Therefore, mechanical inputs trigger a mechanism that elicits spike trains in the neurons that innervate the cell body.

The process by which HCs detect mechanical input from the surrounding environment and transmit that information to the brain involves a cascade of cellular events. Upon deflection of the hair bundle, which pivots the stereocilia at the base, tension increases in the tip links that connect adjacent pairs of stereocilia (Fig. 1(B)). The increased tension opens mechano-sensitive transduction channels that are coupled to the tip links, enabling the influx of ionic currents, leading to depolarization of the HC. As a consequence, vesicular release of neurotransmitters is triggered at the afferent synapse, which triggers spike trains in the neurons and ultimately conveys a signal to the brain [4]. The opposite pathway, transmission of signals from higher order processing centers of the brain back to the sensory epithelium, is managed by the efferent neurons, which provide a protective feedback mechanism [5, 6]. Both conduits interact with the HC at the lower area of the cell body through the synaptic boutons of innervating fibers, as shown in Fig. 1(A).

The auditory system further contains an active amplifier, which expends energy to enhance the response to extremely small signals. At the level of the hair bundle, internal adaptation process involving myosin motor activity have been shown to lead to spontaneous oscillations. One of the manifestations of the active process, these oscillations can be modulated by mechanical, electrical, and chemical manipulations and provide a useful probe of internal mechanics of the bundle.

In this study, we explore the use of widefield microscopy combined with analysis of light intensity fluctuations, to study mechanical perturbations of the hair bundle as well structures with the hair cell soma. We aim to develop

a methodology that would enable the study of hair cells somatic activity with a non-invasive and dye-free approach, so as to minimize mechanical or physiological disturbance of natural conditions. Our results indicate the presence of mechanical fluctuations, localized to specific layers within the hair cell soma and associated with hair bundle motility. We propose that this technique provides a useful approach for a sensitive and noninvasive recording of mechanical movement, for structures deep within the biological tissue.

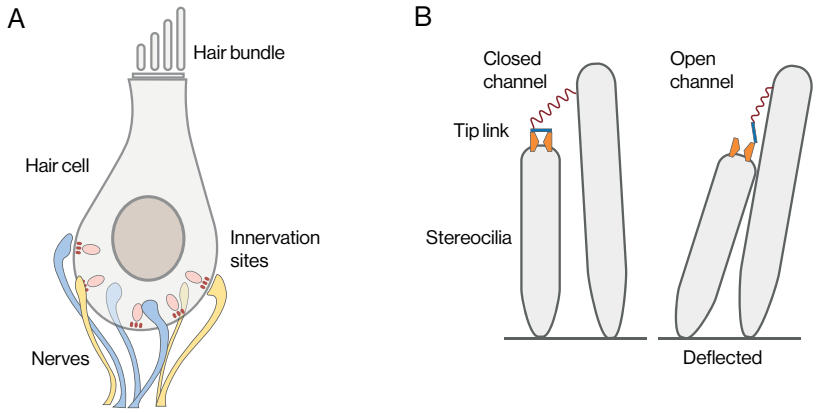


FIGURE 1. (A) Schematic diagram of a HC and the innervating neurons. The cluster of stereocilia at the apical surface comprise the mechanosensitive organelle of a hair cell, the hair bundle. Synaptic contacts are present at the basal pole of the HC corresponding to fiber innervations. (B) Schematic of the transduction process. Stereocilia are inter-connected by tip links, which are mechanically coupled to transduction channels. Upon increase in tension of the tip link due to deflection, the transduction channel opens, allowing the influx of ions.

MATERIALS AND METHODS

The biological model used in these experiments was the North American bullfrog (Rana catesbeiana). HCs from dissected sacculi were imaged *in vitro* while maintaining their physiological integrity. Observations were carried out during spontaneous oscillations of the hair bundle and mechanical stimulation of the stereocilia. These perturbations served to trigger or modulate mechanical motions of the hair bundle or within the cell. Optical imaging was then used to detect fluctuations in light intensity associated with mechanical perturbations of the tissue.

Biological Preparation

Frogs of either gender were anesthetized (pentobarbital: 150 mg/kg), pithed, and decapitated following protocols approved by the University of California, Los Angeles Chancellor's Animals Research Committee. Sacculi were excised from the inner ears of the animals and placed in oxygenated artificial perilymph solution (in mM as follows: 110 Na^+ , 2 K^+ , 1.5 Ca^{2+} , 113 Cl^- , 3 D-(+)-glucose, 1 Na^+ pyruvate, 1 creatine, 5 HEPES). The epithelium was mounted in a two-compartment chamber, emulating the fluid partitioning of the *in vivo* physiological conditions. In this arrangement, apical surfaces were bathed in artificial endolymph (in mM as follows: 2 Na^+, 118 K^+, 0.25 Ca^{2+}, 118 Cl^-, 3 D-(+)-glucose, 5 HEPES) and basolateral membranes in perilymph, as depicted in Fig. 2. In order to allow direct mechanical access to the hair bundles, the otolithic membrane was carefully removed from the epithelium after an 8 min enzymatic dissociation with 15 g/ml Collagenase IV (Sigma-Aldrich).

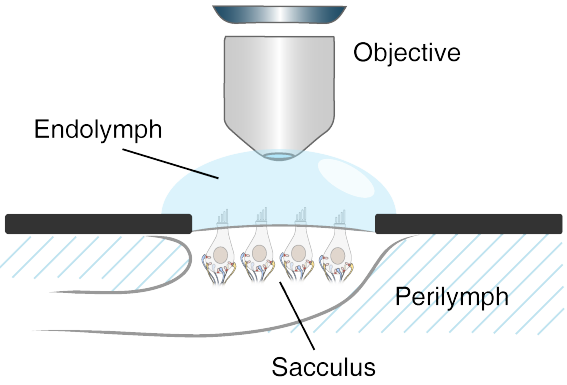


FIGURE 2. HCs in the two-compartment chamber emulating the ionic concentrations of the natural fluid environment of the sacculus. The nerves are immersed in perilymph at the bottom compartment, and the HCs in endolymph at the top. Cells where accessed from the top for mechanical stimulation allowed by the working distance of the lens.

Optical recordings

Recordings were performed using an upright optical microscope (Olympus BX51WI) with a water-immersion objective (Olympus LUMPlanFL N 60X, NA:1.00) mounted on an optical table (Technical Manufacturing). The setup was constructed inside an acoustically isolated chamber (Industrial Acoustics) to avoid introducing external perturbations to the highly sensitive HCs. 16 bit TIFF images at a resolution of 108.3 nm/px were recorded with a high speed camera (ORCA-Flash4.0 CMOS) at 1000 frames per second (fps).

For one set of experiments, motion of the hair bundles was tracked according to previously developed methods, in order to verify the response of cells to stimuli and integrity of the sample. Bright field images were recorded at 1 ms of exposure and processed using custom made MATLAB scripts. Specifically, the center of gravity of the hair bundle was determined from the intensity profile along a row of pixels, obtaining its position in each frame of the recording. Plots of bundle position over time then provided traces of its motion.

Subsequently, the focal plane was modified to obtain consecutive recordings of planes along the z axis. In this manner, a z stack of the HC was obtained, paying special attention to the planes near the basal pole of the cell. The standard deviation of each pixel's intensity,

$$\sigma = \sqrt{\mathrm{E}\left[(\mathbf{I}_k - \mu)^2\right]} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(\mathbf{I}_k - \mu)^2}, \tag{1}$$

was calculated along the time course of the recording, giving an "activity" heat map based on light intensity fluctuations.

$$\mu = \mathrm{E}\left[\mathbf{I}_k\right] = \frac{1}{N}\sum_{k=1}^{N}\mathbf{I}_k, \tag{2}$$

represents the temporal mean matrix of the image sequence. $\mathbf{I}_k$ is the matrix representation of a frame intensity map, a 2D matrix composed of the grayscale levels of each pixel. k is the frame number that assumes values from $0, 1, ..., N$, with N the total number of images from the sequence recorded during the experiment.

With this information, "hotspots" of activity are determined from the matrix σ, thus defining specific spatial regions of interest in the recording. Subsequently, a light intensity trace was obtained by averaging the values of the pixels from the region of interest for each frame.

Mechanical stimulus

In addition to recordings obtained during spontaneous oscillations of the hair bundle, we introduced external perturbations aimed at triggering the synaptic response in the HC. Mechanical stimulation was imposed using a glass probe, fabricated from a borosilicate glass capillary that had been pulled with a micropipette puller (Sutter Instruments). Probes were mounted on a piezoelectric actuator (Piezosystem Jena PA 4/12) and positioned to attach to either the kinocilium or the tallest row of stereocilia, as shown in Fig. 3.

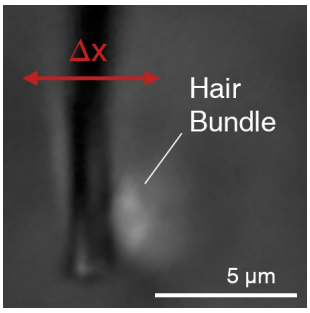


FIGURE 3. Bright field image of a single hair bundle with a glass fiber attached taken with a 60X lens. Focal plane is set at the tip of the stereocilia so as to locate the probe on the kinocilium side. The introduced displacement Δx is perpendicular to the tip of the probe.

RESULTS AND DISCUSSION

Mechanical fluctuations within the hair cell soma

During spontaneous oscillations of the hair bundle, we observed that particular regions of the HC soma show periodic light intensity fluctuations. To study the influence of the hair bundle movement on areas located deeper into the tissue, we therefore analyzed lower focal planes by recording z stacks that maintain the same region of interest.

We first tracked the movement of the hair bundle near the tips of the stereocilia, then in successive recordings along the z axis, we obtained the light intensity trace for the same region of interest that provided the hair bundle oscillation for the first focal plane (Fig. 4). Interestingly, the two following planes did not show detectable fluctuations in light intensity, implying that the hair bundle displacements were not optically influencing the full recording, nor directly mechanically perturbing the full cell soma. The same result was obtained for planes near the basal pole of the cell. On the other hand, the third focal plane presented an oscillating light intensity trace (blue).

As this observation was localized to only one section of the soma, it is unlikely to reflect a direct optical projection of the hair bundle oscillation. Further, the frequency of the somatic movement did not replicate the bundle light trace. These results imply that localized somatic mechanical fluctuation may accompany hair bundle oscillation.

The next set of measurements focused on several hotspots found in the basal pole of the HCs, close to the perimeter of the cell soma. As discussed in the introduction, neuronal innervation is located in this area; hence, optical recordings of this region allow the possibility of establishing an association between spontaneous bundle oscillations and mechanical fluctuations in the synaptic region. Comparisons were made between regions of interest inside and outside the cell, as shown in Fig. 5(A) for the same recording. Hotspots (II) showed oscillatory trends in the light trace with higher frequency when compared to the hair bundle movement. The spot (I) outside the cell showed no activity, as expected. Figure 5(B) depicts the light intensity traces of two very close spots within the cell, in an area that is far from the hair bundle z projection. The trace IV corresponds to a hotspot that showed intensity oscillations, while displacing the region of interest slightly lower, resulted in no activity (III). As demonstrated in these experiment, this technique enables very selective tracing of separate elements of interest within a recording.

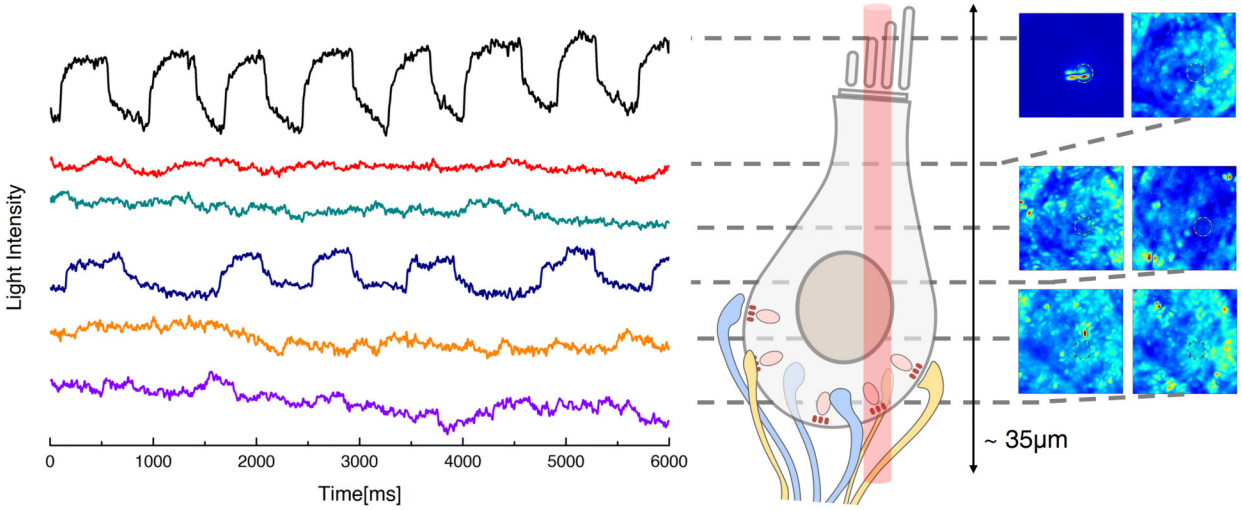


FIGURE 4. Light intensity traces from the same region of interest, obtained at different focal planes along the z axis of a HC. The region of interest is located at the tip hair bundle, tracking its spontaneous oscillations in the first focal plane. For each plane, a sequence of images was recorded and the standard deviation calculated, as shown in the pictures on the right. Mechanical activity was observed within the cell at a specific depth (blue line), with no oscillations detected in the other traces.

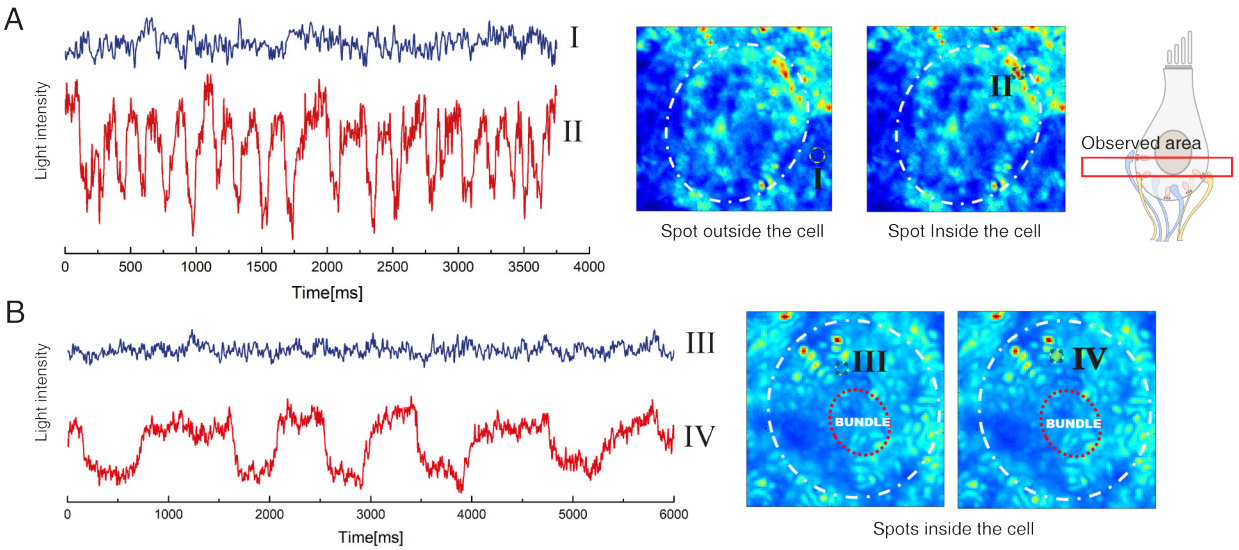


FIGURE 5. (A) Light traces and standard deviations of image sequences of a HC with a spontaneously oscillating hair bundle. The observed imaging plane was centered near the fiber innervation area. The blue line corresponds to spot I outside the cell, and the red one corresponds to spot II, inside the cell within a hot area. Oscillatory behavior was observed in spot II. (B) Two spots in close proximity were compared, selected from a region within the HC that is far from the hair bundle projection. Spot III (in blue) showed no activity, while spot IV (in red) exhibited oscillatory behavior.

Mechanical stimulation of the hair bundle

The use of a glass probe mounted on a piezoelectric stimulator allowed us to introduce controlled mechanical deflections of specific amplitude and frequency. The probe was micromanipulated to be next to a single hair bundle, and visual confirmation was utilized to ensure the glass fiber was in contact with only one hair bundle. The tip of the probe was then attached to either the kinocilium or the tallest row of stereocilia. An oscillatory 10 Hz sinusoidal signal of

$\sim$ 60 nm amplitude was sent to the hair bundle. Recordings consisted of a 5 s stimulation period. Figure 6 shows the tracked position of the hair bundle coupled with the glass probe. As expected, the bundle was entrained by the applied signal, adjusting its oscillations to a 10 Hz frequency. The plots of σ obtained from deeper focal planes during the application of the stimulus yielded clear hot areas (Fig. 6). The most prominent signals were observed at the lower area of the cell body (trace III), far from the projection of the glass probe. Hotspots were located near the edges of the HC, and light intensity traces presented frequency distributions similar to the sinusoidal deflection applied to the bundle.

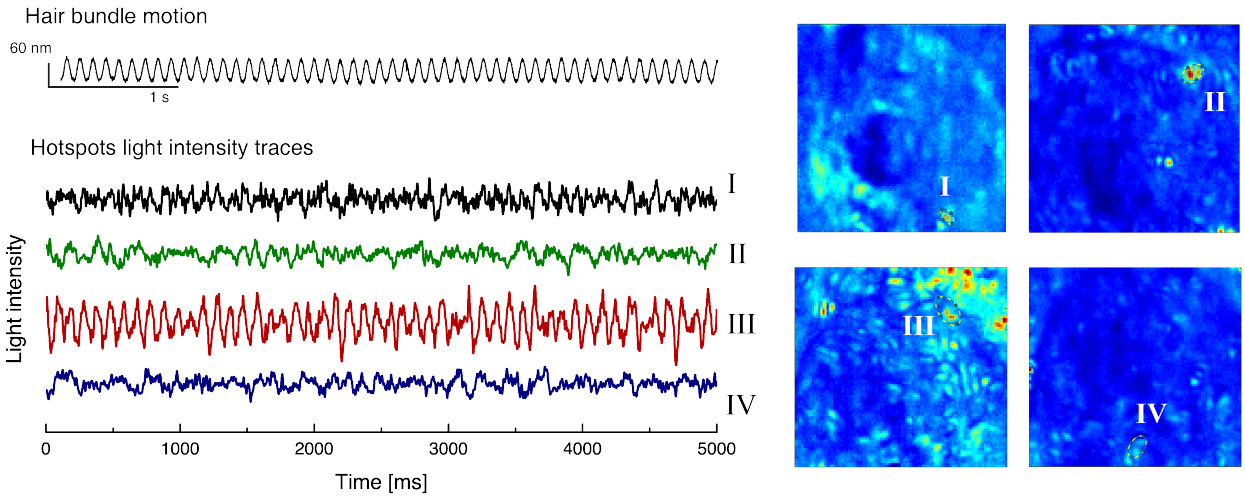


FIGURE 6. Tracking a hair bundle motion attached to a glass probe, light traces and standard deviations of image sequences of a HC obtained during mechanical stimulation of the hair bundle. The probe introduced a 10 Hz sinusoidal oscillatory displacement, perpendicular to the rows of stereocillia, producing displacements of 60 nm. Light intensity traces correspond to regions of interest from consecutive depths within the cell, starting below the cuticular plate. Most prominent hot areas where found at the level of trace III located closer to the basal pole of the cell. Further, the light intensity fluctuations in trace III matched the frequency spectrum of the stimulus. Few hotspots with scarce activity were also found on the top and bottom ends of the cell body.

CONCLUSION

This work proposes a dye-free optical technique aimed at elucidating active motility along various components along the pathway of information flow in the auditory system. HCs were subjected to stimulation through introduction of mechanical deflection of the hair bundles. We recorded images of HCs at high speed to study physical alterations within the cell. Light intensity fluctuations, attributed to mechanical perturbations of tissue within the cells, were detected near the perimeters of the cell somae at localized spots at different focal planes along the z axis. These results indicate that widefield microscopy, applied to semi-intact preparations of the sensory epithelia, provides a possible means of assessing active internal processes of the HC.

Further improvement of these techniques could pave the way forward for developing less invasive techniques to study the encoding of auditory information. In particular, a large number of experimental studies have shown that neuronal activity is accompanied by fast and temporary mechanical changes. These include changes in axonal radius [7, 8], pressure [9] optical properties [10] and the release and subsequent absorption of a small amount of heat and shortening of the axon at its terminus when the action potential arrives. Future work entails the application of these techniques to probe the afferent spike trains in the auditory system.

ACKNOWLEDGMENTS

First author thanks Dr. Justin Faber for his help and guidance on the glass probe making and application of the stimulus. This research was funded by the U.S. Army Research Office under the grant W911NF1910179.

REFERENCES

1. B. L. Robinson and D. McAlpine, "Gain control mechanisms in the auditory pathway," Current Opinion in Neurobiology **19**, 402–407 (2009), sensory systems.
2. A. J. Hudspeth, "Making an effort to listen: Mechanical amplification in the ear," Neuron **59**, 530–545 (2008).
3. M. A. Vollrath, K. Y. Kwan, and D. P. Corey, eng"The micromachinery of mechanotransduction in hair cells," Annual review of neuroscience **30**, 339–365 (2007).
4. E. Glowatzki, L. Grant, and P. Fuchs, "Hair cell afferent synapses," Current Opinion in Neurobiology **18**, 389–395 (2008), sensory systems.
5. R. D. Rabbitt and W. E. Brownell, "Efferent modulation of hair cell function," Current Opinion in Otolaryngology & Head and Neck Surgery **19** (2011).
6. C.-H. J. Lin and D. Bozovic, "Effects of efferent activity on hair bundle mechanics," Journal of Neuroscience **40**, 2390–2402 (2020), https://www.jneurosci.org/content/40/12/2390.full.pdf.
7. L. B. Cohen, "Changes in neuron structure during action potential propagation and synaptic transmission." Physiological Reviews **53**, 373–418 (1973), pMID: 4349816, https://doi.org/10.1152/physrev.1973.53.2.373.
8. I. Tasaki and P. Byrne, "Rapid structural chances in nerve fibers evoked by electric current pulses," Biochemical and Biophysical Research Communications **188**, 559–564 (1992).
9. P. H. Barry, "Volume flows and pressure changes during an action potential in cells ofchara australis," The Journal of Membrane Biology **3**, 313–334 (1970).
10. I. Tasaki and P. M. Byrne, en"The origin of rapid changes in birefringence, light scattering and dye absorbance associated with excitation of nerve fibers," Jpn. J. Physiol. **43 Suppl 1**, S67–75 (1993).
11. E. C. Keen and A. J. Hudspeth, "Transfer characteristics of the hair cell's afferent synapse," Proceedings of the National Academy of Sciences **103**, 5537–5542 (2006), https://www.pnas.org/content/103/14/5537.full.pdf.
12. D. E. Featherstone, "Intercellular glutamate signaling in the nervous system and beyond," ACS Chemical Neuroscience **1**, 4–12 (2010).
13. M. Beurg, R. Fettiplace, J.-H. Nam, and A. J. Ricci, "Localization of inner hair cell mechanotransducer channels using high-speed calcium imaging," Nature Neuroscience **12**, 553–558 (2009).
14. J. Neef, N. T. Urban, T.-L. Ohn, T. Frank, P. Jean, S. W. Hell, K. I. Willig, and T. Moser, "Quantitative optical nanophysiology of ca2+ signaling at inner hair cell active zones," Nature Communications **9**, 290 (2018).
15. J. Lock, I. Parker, and I. Smith, "A comparison of fluorescent ca2+ indicators for imaging local ca2+ signals in cultured cells," Cell Calcium **58** (2015), 10.1016/j.ceca.2015.10.003.
16. A. El Hady and B. B. Machta, "Mechanical surface waves accompany action potential propagation," Nature Communications **6**, 6697 (2015).

On Natural Selection of Outer Hair Cell Electro-Mechanical Properties

Richard D. Rabbitt[1, a)] and Tamara Bidone[2, b)]

[1]*Biomedical Engineering, Otolaryngology and Neuroscience Program, University of Utah, 36 S Wasatch Drive, Salt Lake City, UT 84112.*
[2]*Biomedical Engineering, Pharmaceutics & Pharmaceutical Chemistry and SCI Institute, University of Utah, 36 S Wasatch Drive, Salt Lake City, UT 84112.*

a) Corresponding author: r.rabbitt@utah.edu
b) tamarabidone@sci.utah.edu

Abstract. The biophysical and mechanical properties of outer hair cells (OHCs) vary systematically with location along the cochlear tonotopic map following monotopic relationships that apply universally across all mammalian species studied to date. To reasonable approximation, OHCs located at a specific characteristic frequency (CF; ω_{cf}, rad-s^{-1}) all have the same length, linear capacitance, nonlinear capacitance, basolateral conductance, axial stiffness and mechano-electrical transduction (MET) current kinetics. Such correlations do not arise in nature by chance, but in the Darwinian view arise through natural selection. We hypothesized that OHC biophysical and mechanical properties evolved simply to maximize frequency tuning and OHC electro-mechanical power efficiency. To test this hypothesis, we analyzed a linear toy model consisting of an OHC working against a simulated cochlear load. Analytical methods were used to optimize tuning and efficiency at the same time by adjusting: prestin expression (non-linear capacitance), cell length (linear capacitance), compliance (stiffness), voltage-gated ion-channel expression (basolateral conductance), and MET channel kinetics (gain and adaptation). Results demonstrate these factors are interrelated, and when optimized together predict key relationships between OHC properties and CF observed experimentally. To overcome viscous dissipation and optimize tuning, the MET current must be gated by both displacement and velocity, and have a magnitude quantitatively related to prestin expression, basolateral conductance and OHC stiffness. To maximize efficiency at CF, the resistive-capacitive corner frequency (ω_{rc}, rad-s^{-1}) must be below CF ($\omega_{rc}/\omega_{cf}\sim0.35$), prestin expression must yield a nonlinear capacitance roughly equal to the linear capacitance (C; $C_n/C\sim0.9$) and the OHC axial stiffness (k_o, N-m^{-1}) must exceed the stiffness of the cochlear load (stiffness per OHC without OHCs; $k/k_o\sim0.2$). If any of these optimized parameters are changed, tuning and power efficiency drop precipitously in the toy model. Based on this finding, we further hypothesize that an important role for tonic firing of olivo-cochlear efferent (MOC) neurons is to optimize the tuning, and that changing MOC activity would have a much stronger effect on power output than predicted by a reduction in the voltage modulation alone.

INTRODUCTION

Outer hair cell (OHC) length, ionic conductance, mechano-electrical transduction current, passive capacitance, axial stiffness and prestin expression all vary monotonically along the logarithmic cochlear tonotopic map. Cells located at the 0.2 kHz characteristic frequency (CF, ω_{cf}) place have a length of ~80 μm while cells located at the 10 kHz CF place have a length of ~40 μm [1-3]. The reduction in length causes an increase in OHC whole-cell stiffness (k_o) with CF [4], which has log-linear scaling similar to the stiffness of the cochlear partition (k) [5]. The correlation between OHC length and CF is universal across mammalian species, but precisely why the correlation exists is not completely known. The change in length is much too small to support a purely mechanical hypothesis, suggesting electro-mechanical properties of OHC might drive the correlation. Passive capacitance (C) and peak nonlinear capacitance (C_n) scale roughly in proportion to the membrane surface area [6], while ion channel expression and

membrane conductance underlying the passive resistive-capacitive RC corner (ω_{rc}) increase dramatically with CF [7]. Capacitance and conductance cause the RC corner frequency to increase with CF, which plays an important role in amplification by shifting the timing of electro-mechanical force production [8, 9]. It has been hypothesized previously that OHC biophysical properties vary with CF in a way that optimizes electro-mechanical power conversion through a type of impedance matching [10, 11], but it is still not known precisely how key parameters are interrelated to achieve the match. Here, using a simple toy model of an OHC coupled to a simulated cochlear load, we present evidence that the nondimensional ratios $\frac{\omega_{rc}}{\omega_{cf}}$, $\frac{C_n}{C}$ & $\frac{k}{k_o}$ are dictated by impedance matching, and are mutually-dependent constants that apply universally across all mammalian species independent of CF.

A SIMPLE TOY MODEL

To illustrate the general principles governing impedance matching of OHCs to the cochlear load, we introduce a simple toy model consisting of a one-dimensional piezoelectric outer hair cell working against a spring-mass-damper load. It has been shown previously for axisymmetric, homogeneous deformations under constant intracellular volume, the linear piezoelectric shell equations can be reduced to one dimension relating the OHC axial strain S to the axial force F and transmembrane voltage V according to:

$$S = \kappa F + \delta V, \tag{1}$$

where κ (N^{-1}) is the axial strain compliance and δ (V^{-1}) is the axial piezoelectric strain coefficient. The whole-cell displacement X in this toy model is the strain S times the length ℓ. The lateral wall constitutive parameters κ and δ vary with voltage and stress, and are approximated in the present analysis as constants independent of cell length and CF location within the cochlea (c.f. prestin expression per unit membrane area is assumed constant and strain compliance is constant). For simplicity, we assume the speed of the prestin-motor complex is dictated by the force working against the motor in the membrane environment and is included as part of F. Upper case denotes the frequency domain. Eq. 1 gives the force from the prestin-motor complex as $F = k_o X - \frac{\delta}{\kappa} V$, where the axial whole-cell OHC stiffness is $k_o = 1/\kappa\ell$ (N-m-1) [4].

Treating the cochlear load as a simple spring-mass-damper system:

$$\left((k - m\omega^2) + bj\,\omega\right)X = F_p - F, \tag{2}$$

where ω (rad-s^{-1}) is stimulus frequency, k (N-m^{-1}) is the stiffness of the cochlear load, m is the effective mass and b is the viscous damping, all per OHC. F_p (N) is the force per OHC arising from the sound-induced dynamic pressure across the cochlear partition. Combining 1-2

$$k_p\left(1 - \left(\frac{\omega}{\omega_p}\right)^2 + 2\,\zeta_p j\left(\frac{\omega}{\omega_p}\right)\right)X - \frac{\delta}{\kappa}V = F_p, \tag{3}$$

where $\omega_p = \sqrt{(k + k_o)/m}$ (rad-s^{-1}) is the passive undamped natural frequency of the cochlear partition in the absence of prestin. The stiffness $k_p = k + k_o$ (N-m^{-1}) of the cochlear partition per OHC is the addition of the OHC stiffness k_o and parallel cochlear structures k. The viscous drag parameter $\zeta_p = b\omega_p/2k_p$ is the nondimensional damping coefficient associated with the passive cochlear partition in the absence of OHC amplification.

From Kirchhoff's law in the frequency domain, for small perturbations, the intracellular voltage modulation V is related to the current input I by

$$j\omega CV + (\delta/\kappa)U + GV = I\,, \tag{4}$$

where $U = j\,\omega\,X$ is the velocity, C is the passive electrical membrane capacitance and G is the linearized conductance from all ion channels. The linearized mechano-electrical transduction (MET) current is assumed be a function of cell

displacement and velocity according to $I = \left(\alpha + \frac{\sigma}{j\,\omega}\right) U$, where σ and α are MET velocity and displacement gains respectively. The electrode current (A, if present) adds to the MET current. It's important to note that these expressions have been linearized for small perturbations in velocity, force, voltage and current about a holding potential, and do not include nonlinear saturation of the MET current, nonlinear saturation of prestin, or strain stiffening of the OHC cytoskeleton.

Combining Eq. 1-4, it can be shown that the nonlinear whole-cell capacitance measured at low frequencies under zero load is $C_n = \ell\delta^2/\kappa$ (pF) and the passive RC corner frequency is $\omega_{rc} = G/C$ (rad-s^{-1}). For large displacements, not considered here, δ is strain-dependent, and the MET gains σ and α are stimulus amplitude dependent. Considering only the linear case and multiplying Eq. 4 by complex conjugate of voltage in the frequency domain and taking half of the real part gives the electrical power consumed by the prestin-motor complex:

$$P \approx \frac{\delta}{2\kappa} Re(UV^*). \tag{5}$$

The * denotes the complex conjugate. Note from Eq. 5 that power conversion is maximized when the phase of the velocity is aligned with the phase of the voltage – which only occurs when the OHC biophysical parameters are optimized for power conversion efficiency.

Optimum MET phase, prestin expression, RC corner and stiffness.

Solving Eqns. 1-4 in the frequency domain for force driven OHC velocity provides:

$$U = \frac{-\omega\left(\frac{A\delta}{\kappa} + F_p(G + iC\omega)\right)}{i(G+iC\omega)(\omega(-m\omega+ib)+k+k_o)-\frac{\delta}{\kappa^2}(-\alpha\kappa\omega+\delta\omega+i\kappa\sigma)}. \tag{6}$$

Resonance occurs when the denominator goes to zero, which we define as the characteristic frequency (CF) of the toy OHC-cochlea model. In the absence of dissipation, ω_{cf} (rad-s^{-1}) is

$$\omega_{\mathrm{cf}} = \sqrt{\frac{k_o\left(1+\frac{C_n}{C}\right)+k}{m}}, \tag{7}$$

Note, prestin expression increases $\frac{C_n}{C}$, the effective stiffness of the OHC $k_o\left(1+\frac{C_n}{C}\right)$ and ω_{cf}, over and above that present from passive mechanics alone.

We hypothesize nature has optimized the MET current magnitude and phase to maximize the quality Q of the resonance at CF. Quality is theoretically infinite in this toy linear model when the denominator is zero, giving the optimum MET displacement gain:

$$\sigma = \frac{\kappa}{\delta}\left(G k_o \frac{C_n}{C} - bC{\omega_{cf}}^2\right) \tag{8}$$

and optimum velocity gain:

$$\beta = \frac{\kappa}{\delta} bG. \tag{9}$$

Assuming this is the case, the total magnitude of the MET current squared is:

$$\langle I_{MET}\rangle^2 = \left(\frac{\kappa}{\delta\omega}\right)^2 \left(\left(2\zeta G\,(k + k_o\,)\right)^2 + \left(2C\zeta\omega(k + k_o\,) - \frac{C_n}{C} G\,k_o\,\right)^2\right). \tag{10}$$

Optimization of the MET current requires the phase ϑ (rad) of the MET current relative to the OHC displacement to be:

$$\mathrm{Cot}(\vartheta) = -\frac{\omega_{\mathrm{cf}}}{\omega_{\mathrm{rc}}} - \frac{1}{2\zeta_p}\frac{C_n}{C}\sqrt{\left(\frac{k}{k_o}+1\right)\left(\frac{C_n}{C}+\frac{k}{k_o}+1\right)}. \tag{11}$$

It is important to note that the optimum phase depends on *four key nondimensional numbers*: $\frac{\omega_{\mathrm{cf}}}{\omega_{\mathrm{rc}}}$ which is the ratio of the characteristic frequency to the passive RC corner frequency, $\frac{C_n}{C}$ which is the ratio of the nonlinear capacitance under zero external load to the linear capacitance, $\frac{k}{k_o}$ which is the ratio of the cochlear partition stiffness per OHC to the passive OHC axial stiffness and ζ_p which is the passive nondimensional damping coefficient.

OHC electromotility in this toy model perfectly cancels damping of the cochlear load at frequency ω_{cf} providing the MET gains are set according to Eqs 8-9 and 11, but not all parameter combinations result in the same MET current magnitude or electro-mechanical power-conversion efficiency. Efficiency is maximized by finding parameters that minimize the MET current at CF, yet maintain theoretically infinite Q resonance in the linear model. For the optimization, we approximate linear and nonlinear capacitance as proportional to cell length using $C = \xi\ell$ and $C_n = \xi_n\ell$ (ξ *and* ξ_n taken as constants) and whole-cell axial stiffness as inversely proportional to cell length using $k_o = 1/\kappa\ell$. Minimizing $\langle I_{MET}\rangle^2$ with respect to conductance and cell length determines the relationship between the passive RC corner frequency $\omega_{rc} = C/G$ and characteristic frequency ω_{cf} as:

$$\left(\frac{\omega_{\mathrm{rc}}}{\omega_{\mathrm{cf}}}\right)^2 \rightarrow \frac{r(r^2+r+1)}{(r+1)^4}. \tag{12}$$

where the stiffness ratio is $r = \frac{k}{k_o}$. Minimizing $\langle I_{MET}\rangle^2$ with respect to prestin expression and cell length provides the optimum nonlinear capacitance:

$$\frac{C_n}{C} \rightarrow 2\,\zeta_p\frac{(r+1)^3}{r^2+r+1}, \tag{13}$$

Combining equations 11-12, the optimum OHC stiffness k_o relative to the partition stiffness k per OHC is:

$$\frac{1}{r} \rightarrow 2\zeta_p\left(\frac{\omega_{\mathrm{cf}}}{\omega_{\mathrm{rc}}}\right)^2\left(\frac{C}{C_n}\right) - 1. \tag{14}$$

RESULTS AND DISCUSSION

The perfect impedance match

The relationships provided in Eqns. 11-14 define the *perfect impedance match* of the toy OHC to the cochlear load. All four relationships depend on three intrinsic nondimensional numbers ($\frac{\omega_{\mathrm{rc}}}{\omega_{\mathrm{cf}}}$, $\frac{C_n}{C}$, $\frac{k}{k_o}$) which in OHCs are directly controllable through gene expression, and one extrinsic nondimensional number ζ_p which depends primarily on external factors. The key observation is that setting any one of the three intrinsic nondimensional numbers compels specific values for the other two. Nature has the freedom to independently vary channel expression, prestin expression and structural protein expression, thus making these nondimensional numbers independent variables in the mathematical sense. Hence, by separation of variables, we must conclude that $\frac{\omega_{\mathrm{rc}}}{\omega_{\mathrm{cf}}}$, $\frac{C_n}{C}$, and $\frac{k}{k_o}$ are all constants independent of CF or cell length, at least for the toy model considered here. If, for example, we set $\frac{C_n}{C}$=0.89 based on experimental data from guinea pig reported by Corbitt et al. [12], the impedance match defined by Eq. 11-14 requires $\frac{\omega_{\mathrm{rc}}}{\omega_{\mathrm{cf}}}$=0.36, $\frac{k}{k_o}$=0.21 and ϑ=-0.27. The prediction that $\frac{\omega_{\mathrm{rc}}}{\omega_{\mathrm{cf}}}$=0.36 based on guinea pig is in close agreement with independent estimates in rat and gerbil [7].

Breaking the impedance match by altering any of the intrinsic nondimensional numbers is predicted to lead to a dramatic loss in the efficiency of electro-mechanical power conversion, and a concomitant reduction in the power delivered to the viscous load. To illustrate the effect of each ratio, we varied the basolateral conductance to change ω_{rc}, the level of prestin expression to vary C_n and the OHC stiffness to vary k_o. Simulations were carried out for a ℓ=56 µm long OHC at the 3 kHz place ($\omega_{cf} = 3000 * 2 * \pi$ rad-s^{-1}). The axial OHC stiffness was set using $k_o = 1/\kappa\ell$ based on the axial strain compliance κ=1.96x10^6 (N^{-1}) reported previously by Iwasa [4]. With this, all remaining parameters in the impedance matched condition come from $\frac{C_n}{C}$=0.89, $\frac{\omega_{rc}}{\omega_{cf}}$=0.36, $\frac{k}{k_o}$=0.21, ζ_p=0.32 and ϑ=-0.27.

The optimum impedance matched MET current is illustrated in Figure 1 in the form of magnitude (A) and phase (B) re: OHC displacement. Relative OHC power output (Eq. 5) to the viscous load is indicated as equally space isoclines on a log scale, with the maximum shown in the center at 3kHz (dashed gray line). All results in the present report were generated in response to a constant amplitude sinusoidal force driving the simulated cochlear partition. Dashed isoclines indicate regions where the OHC absorbs power from the mechanical driving force, while solid isoclines indicate regions of active electro-mechanical amplification that add power over and above that available from the driving force. As expected, the power output increases as the MET in increased to an optimum level (Figure 1A), but, perhaps counterintuitively, the power begins to decline as the MET current is further increased. The decline at high MET currents occurs because the MET displacement and velocity gains (Eq. 8 & 9) exceed values that force the denominator in Eq. 6 to zero, thereby breaking the impedance match. Exceeding these gains introduces inefficiency and causes the OHC to work against itself and the extrinsic load. Optimization of the MET phase (Figure 1B) simply aligns the timing of OHC force production with velocity, which is required for peak power output (Eq. 5). The optimum phase arises from Eqs. 8-9 and accounts for the fact that the voltage is phase shifted as a result of delivering power to the viscous load [8, 10].

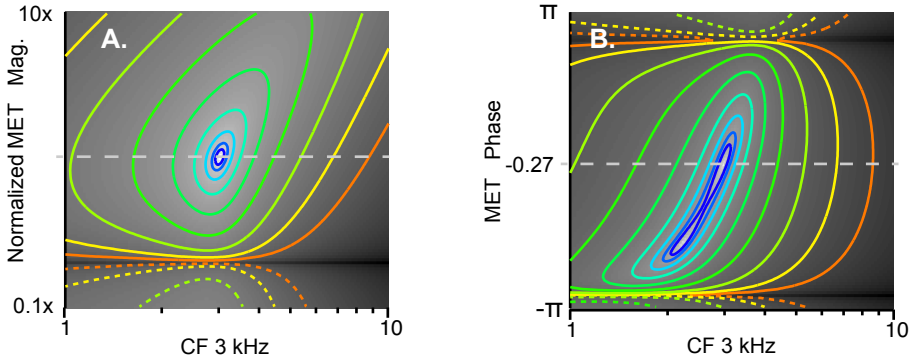


Figure 1. Optimum MET magnitude and phase for a 56 µm long OHC at 3kHz CF. Eq. 8 -9 provide the real and imaginary components of the MET current required to zero the denominator of Eq. 6 and perfectly cancel losses from viscous drag and electrical conductance. Isoclines show power output to the viscous load equally spaced on a log scale. A) Increasing or decreasing the magnitude of the MET current without changing the phase reduces power output and B) increasing or shifting the MET phase relative to cell displacement also decreases the power output.

The influence of changing nonlinear capacitance (prestin), the passive RC corner (conductance and passive capacitance) and passive OHC stiffness (cytoskeleton) on OHC power output to the viscous load is shown for the toy model in Figure 2 for the same 56 µm long cell. Figure 2A shows how the power output to the viscous load declines if C_n is above or below the optimum value. The peak power output occurs in the toy model at $\frac{C_n}{C}$=0.89, consistent with a report in OHCs by Corbitt et al. [12]. The magnitudes of nonlinear capacitance C_n and the corresponding piezoelectric strain coefficient δ (note: $C_n = \ell\delta^2/\kappa$) are set by the level of prestin expression and the resting potential. Reduced power output for low C_n arises from the reduction in electromotive force, while reduced power output for high C_n arises from piezoelectric stiffening of the cell, which breaks the impedance match.

Figure 2B shows how the power output declines if the passive RC corner frequency (note: $\omega_{rc} = G/C$) is shifted relative to the optimum value. The peak power output occurs at $\frac{\omega_{rc}}{\omega_{cf}}$=0.36, consistent with a report by Johnson et al. [7] in OHCs, and computed here from Eq. 12 using exactly the same parameters that predicted $\frac{C_n}{C}$=0.89. A decrease

in conductance G and decrease in ω_{rc} results in a larger voltage modulation, but also shifts the phase of voltage modulation away from velocity and thereby results in decreased power output. An increase in conductance has the dual inhibitory effect of decreasing the voltage modulation and shifting the phase away from the impedance matched value. Results further suggest this dual mechanism is responsible for the dramatic reduction of mechanical vibrations [13] reduction in spiral ganglion neuron sensitivity [14, 15] caused by activation of the medial olivocochlear (MOC) efferent system. MOC synaptic inputs inhibit OHCs through nicotinic acetylcholine receptors at the base of the cell that trigger opening of Ca2+ activated K+ channels [16] thereby increasing ω_{rc}. We hypothesize that tonic MOC efferent firing adjusts ω_{rc} in order to achieve maximize power output through impedance matching, while phasic activation dramatically inhibits amplification by decreasing the receptor potential modulation and disrupting the impedance match.

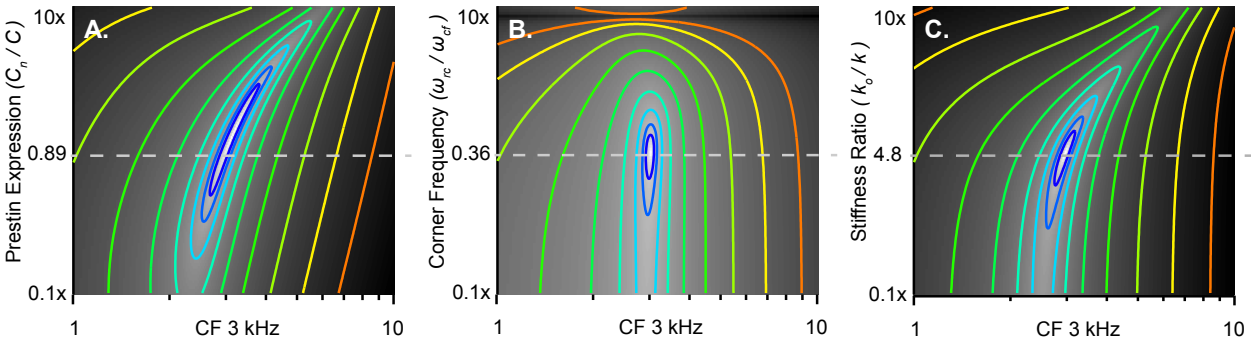


Figure 2. Optimum MET magnitude and phase for a 56 µm long OHC at 3kHz CF. Eq. 8 -9 provide the real and imaginary components of the MET current required to zero the denominator of Eq. 6 and exactly cancel viscous drag and electrical conductive loss. Isoclines show power output to the viscous load equally spaced on a log scale. Dashed isoclines indicate the mechanical power input to the toy model exceeds the mechanical power output to the viscous load. A) Increasing or decreasing the magnitude of the MET current without changing the phase reduces power output and B) increasing or shifting the MET phase relative to cell displacement also decreases the power output.

Impedance matching in the toy model also requires the ratio of the stiffness of a single OHC k_o to the stiffness of the cochlear partition per OHC k (absent OHCs) to be independent of CF location and to have a constant value of ~4.8. The stiffness of the in-tact cochlear partition along its length has a log-linear relation [5], as does the relationship between OHC length and CF [1-3]. Combined, these data suggest the stiffness ratio $\frac{k}{k_o}$ is indeed a constant independent of CF location, as predicted here by impedance matching.

SUMMARY AND CONCLUSIONS

Although the toy model presented here is very simple, results compel the hypothesis that the nondimensional numbers $\frac{\omega_{rc}}{\omega_{cf}}$, $\frac{C_n}{C}$ and $\frac{k}{k_o}$ are nearly constant across all mammalian species and characteristic frequencies, and are precisely related to each other to maximize electro-mechanical power conversion and cochlear amplification. Present results support the hypothesis that OHCs match the properties of the cochlear partition through genes controlling cell size (passive capacitance and stiffness, ω_{rc} and k_o), cytoskeletal architecture (stiffness, k_o), nonlinear capacitance (prestin, C_n) and basolateral ion channel conductance (ion channels, ω_{rc}). Results further support the hypothesis that the amplitude and phase of the bundle driven MET current varies with CF to match the electro-mechanical impedance of the OHC to the cochlear partition. The matched system precisely aligns the phase of prestin force production with the phase of the viscous drag, and in doing so dramatically increases cochlear sensitivity and sharpens frequency tuning. Results support the additional hypothesis that tonic activity of the MOC efferent system serves to optimize the impedance match and cochlear amplification, while phasic MOC activity rapidly controls cochlear sensitivity by disrupting the impedance match.

ACKNOWLEDGMENTS

The authors thank Nidhi Soley for comments on portions of the manuscript. Funding was provided by the National Institutes on Deafness and other Communication Disorders grant R01 DC006685.

REFERENCES

1. Fettiplace, R., Hair Cell Transduction, Tuning, and Synaptic Transmission in the Mammalian Cochlea. Compr Physiol, 2017. 7(4): p. 1197-1227. https://www.ncbi.nlm.nih.gov/pubmed/28915323
2. Dannhof, B.J., B. Roth, and V. Bruns, Length of hair cells as a measure of frequency representation in the mammalian inner ear? Naturwissenschaften, 1991. 78(12): p. 570-3. https://www.ncbi.nlm.nih.gov/pubmed/1780010
3. Pujol, R., et al., Correlation Between the Length of Outer Hair Cells and the Frequency Coding of the Cochlea, in Auditory Physiology and Perception, Y. Cazals, K. Horner, and L. Demany, Editors. 1992, Pergamon. p. 45-52. https://doi.org/10.1016/B978-0-08-041847-6.50011-3
4. Iwasa, K.H. and M. Adachi, Force generation in the outer hair cell of the cochlea. Biophys J, 1997. 73(1): p. 546-55. https://www.ncbi.nlm.nih.gov/pubmed/9199816
5. Naidu, R.C. and D.C. Mountain, Measurements of the stiffness map challenge a basic tenet of cochlear theories. Hear Res, 1998. 124(1-2): p. 124-31. https://www.ncbi.nlm.nih.gov/pubmed/9822910
6. Farrell, B., R. Ugrinov, and W. Brownell, Frequency dependence of admittance and conductance of the outer hair cell, in Auditory Mechanisms, Processes and Models, A. Nuttall, et al., Editors. 2006, World Scientific: New Jersey. p. 231-2.
7. Johnson, S.L., et al., Prestin-driven cochlear amplification is not limited by the outer hair cell membrane time constant. Neuron, 2011. 70(6): p. 1143-54. https://www.ncbi.nlm.nih.gov/pubmed/21689600
8. Nam, J.H. and R. Fettiplace, Optimal electrical properties of outer hair cells ensure cochlear amplification. PLoS One, 2012. 7(11): p. e50572. https://www.ncbi.nlm.nih.gov/pubmed/23209783
9. Agarwal, N. and S. Ramamoorthy, Balance in the feedback loop component of the mammalian cochlear amplifier. J. Applied Physics, 2020. 128(3): p. 034701. https://doi.org/10.1063/5.0010802
10. Rabbitt, R.D., The cochlear outer hair cell speed paradox. Proc Natl Acad Sci U S A, 2020. https://www.ncbi.nlm.nih.gov/pubmed/32848062
11. Rabbitt, R.D., et al., Power efficiency of outer hair cell somatic electromotility. PLoS Comput Biol, 2009. 5(7): p. e1000444. https://www.ncbi.nlm.nih.gov/pubmed/19629162
12. Corbitt, C., et al., Tonotopic relationships reveal the charge density varies along the lateral wall of outer hair cells. Biophys J, 2012. 102(12): p. 2715-24. https://www.ncbi.nlm.nih.gov/pubmed/22735521
13. Murugasu, E. and I.J. Russell, The effect of efferent stimulation on basilar membrane displacement in the basal turn of the guinea pig cochlea. J Neurosci, 1996. 16(1): p. 325-32. https://www.ncbi.nlm.nih.gov/pubmed/8613799
14. Guinan, J.J., Jr. and K.M. Stankovic, Medial efferent inhibition produces the largest equivalent attenuations at moderate to high sound levels in cat auditory-nerve fibers. J Acoust Soc Am, 1996. 100(3): p. 1680-90. https://www.ncbi.nlm.nih.gov/pubmed/8817894
15. Wiederhold, M.L. and N.Y. Kiang, Effects of electric stimulation of the crossed olivocochlear bundle on single auditory-nerve fibers in the cat. J Acoust Soc Am, 1970. 48(4): p. 950-65. https://www.ncbi.nlm.nih.gov/pubmed/5480390
16. Rohmann, K.N., et al., Activation of BK and SK channels by efferent synapses on outer hair cells in high-frequency regions of the rodent cochlea. J Neurosci, 2015. 35(5): p. 1821-30. https://www.ncbi.nlm.nih.gov/pubmed/25653344

Low-Level Distortion Product Otoacoustic Emissions in Lizards are Influenced by Spontaneous Activity of the Inner Ear

Yuttana Roongthumskul[1, a)]

[1]Department of Physics, Faculty of Science, Chulalongkorn University, Bangkok, Thailand.
[a)]*Corresponding author: yuttana.r@chula.ac.th*

Abstract. Distortion product otoacoustic emissions (DPOAEs) from lizards can display a rapid variation in intensity level upon changes in the distortion frequency. The level functions of both cubic distortion products elicited by low- to moderate-level stimuli exhibit pronounced notches across range of the distortion frequencies. The emission level varies more smoothly with the distortion frequency at higher stimulus intensity. To understand the variation in DPOAE level, we measured DPOAEs from tokay geckos and demonstrated that the level of the distortion signals, evoked by primary tones with intensity levels below approximately 60 dB SPL was strongly correlated with the amplitude spectrum of spontaneous otoacoustic emissions (SOAEs). The DPOAE intensity level was significantly enhanced when the distortion frequency coincided with the center frequency of an SOAE peak. Manipulations of the body temperature revealed that peaks in the SOAE amplitude spectrum and the DPOAE level function displayed systematic shifts to lower frequencies of comparable magnitudes upon cooling. Our results suggest a common underlying mechanism of low-level DPOAEs and spontaneous activity in the inner ears of tokay geckos.

INTRODUCTION

In response to two-tone stimuli at frequencies f_1 and f_2, the inner ears can produce acoustic energy at frequencies corresponding to linear combinations of f_1 and f_2. These intermodulation distortions can be detected externally as distortion product otoacoustic emissions (DPOAEs). When evoked by low- to moderate-level stimuli, the intensity level of the cubic distortion products observed at $2f_1$-f_2 and $2f_2$-f_1 can vary drastically upon changes in the stimulus frequency. In lizards, the DPOAE level functions can exhibit multiple peaks and notches across range of the distortion frequencies [1-3]. At higher stimulus intensity, however, the emission level varies more smoothly with the distortion frequency [1-2].

The distinct characteristics of the DPOAE level functions elicited by low- and high-level stimuli suggest multiple mechanisms that involve in the generation of DPOAEs. While high-level distortions are likely contributed by the passive nonlinearity of the inner ear, low-level DPOAEs potentially involve the active processes performed by hair cells. This has been partly evidenced by the alterations of low-level DPOAE intensity upon physiological insults to the inner ear, such as hypoxia [1]. Moreover, previous investigations suggest that low-level DPOAEs can be amplified when the primary tone's or distortion frequency coincide with the center frequency of a peak in the spontaneous otoacoustic emission (SOAE) amplitude spectrum, measured in the absence of stimulus [1].

To illustrate whether low-level distortions can be influenced by the SOAE activity from the inner ears of lizards, we recorded the pressure signals from the ears of tokay geckos in the presence of two-tone stimuli. The intensity level of DPOAEs measured across range of the distortion frequencies was compared to the SOAE amplitude spectrum. We found that the level of DPOAEs evoked by primary tones with large frequency differences varied consistently with the SOAE amplitude spectrum. Finally, we decreased the body temperature of tokay geckos to reveal the correlation between the SOAE activity and the level of DPOAEs evoked by two-tone stimuli at small frequency differences.

METHODS

Tokay geckos (*Gekko gecko*) were anesthetized with an intraperitoneal injection of 20 mg/kg sodium pentobarbital (Nembutal; Akorn Pharmaceuticals, Lake Forest, IL). The body temperature of the animal was controlled by a heating pad at approximately 28 ºC. Measurements of otoacoustic emissions were performed with a probe microphone (ER-10B+; Etymotic Research, Elk Grove Village, IL), connected to an end of a plastic coupler. The other end of the coupler was sealed around the outer ear with vacuum grease. Pressure signals were acquired at a sampling interval of 20 μs. Each SOAE recording was performed for 100 seconds. Acoustic stimulations for DPOAE measurements were delivered by two speakers (MF1; Tucker Davis Technologies, Alachua, FL) connected to the microphone via two plastic tubes. Each DPOAE recording was acquired for 5 seconds.

An artifact-rejection algorithm was applied to each recording. The maximal acoustic pressure levels in each 0.2-s nonoverlapping window and their SDs were calculated. Any windows whose peak pressure exceeded thrice the SD was excluded from further analysis. A finite-time Fourier transform with a Hanning window was calculated for each of the remaining sections. The amplitude spectrum was obtained by averaging the magnitudes of the spectra. A peak in the SOAE amplitude spectrum or a DPOAE level function was defined as a point whose magnitude exceeded those of the adjacent troughs by more than 1 dB.

RESULTS

In the absence of acoustic stimuli, the SOAE amplitude spectrum typically displayed several peaks whose center frequencies ranged from approximately 0.5 kHz to 5.5 kHz. Upon an application of two pure tones at frequencies f_1 and f_2, the amplitude spectrum of the recorded acoustic pressure illustrated an SOAE activity perturbed by the stimuli superposed by DPOAEs at the distortion frequencies, as well as the stimuli at f_1 and f_2 (Fig. 1a). Comparison to the SOAE spectrum recorded in a quiet environment revealed that the level of the spontaneous activity at frequencies near those of the stimuli was greatly attenuated (Fig. 1a). The SOAE activity within the range of frequencies sufficiently far from the stimuli remained unperturbed.

We first investigated the variation of the $2f_1$-f_2-DPOAE level across range of the distortion frequencies by applying a pure tone at a frequency, $f_1 = 3.4$ kHz, corresponding to the center frequency of an SOAE peak. The other stimulus frequency, f_2, was varied from 3.42 kHz to 4.80 kHz at a 20-Hz increment. The intensity levels of both primary tones were fixed at 50 dB SPL. Upon varying f_2, the frequency range displaying the SOAE suppression remained unaltered (Fig. 1a). For this stimulus paradigm, we thus averaged the emission amplitude spectra recorded at different values of f_2, with the amplitudes at the distortion frequencies excluded from the analysis, to illustrate the perturbed SOAE activity (Fig. 1b – 1c).

For small differences in the primary tone frequencies, f_2-$f_1 < 0.7$ kHz, the distortion was elicited at $2f_1$-f_2 greater than 2.7 kHz. Comparison between the perturbed SOAE activity and the SOAE spectrum recorded in the absence of stimuli indicated that these distortion signals emerged at frequencies within the range of SOAE suppression. The level of $2f_1$-f_2-DPOAE varied smoothly with the distortion frequency and displayed a broad peak (Fig. 1b). The center frequency of the peak, however, did not coincide with a peak in the spontaneous activity.

On the other hand, $2f_1$-f_2-DPOAE whose frequency was less than 2.7 Hz displayed a more rapid variation in its intensity level across range of the distortion frequencies with peaks and notches observed (Fig. 1b). These distortion signals were elicited at frequencies outside of the range of SOAE suppression. We found that when the value of $2f_1$-f_2 was near the center frequency of an SOAE peak, the distortion level significantly exceeded that of the spontaneous activity (Fig. 1d). On the contrary, the level of DPOAE at a frequency corresponding to a trough in the SOAE spectrum were indistinguishable from the SOAE level. This resulted in a level function of DPOAE that varied consistently with the SOAE amplitude spectrum.

Similar effects were observed when the primary tone's frequency f_2 was fixed at 2.66 kHz, corresponding to the center frequency of an SOAE peak, and the other stimulus frequency f_1 was varied from 1.26 kHz to 2.64 kHz. The level of $2f_2$-f_1-DPOAE was enhanced when the distortion frequency was near the center of a peak in the perturbed SOAE activity (Fig. 1c).

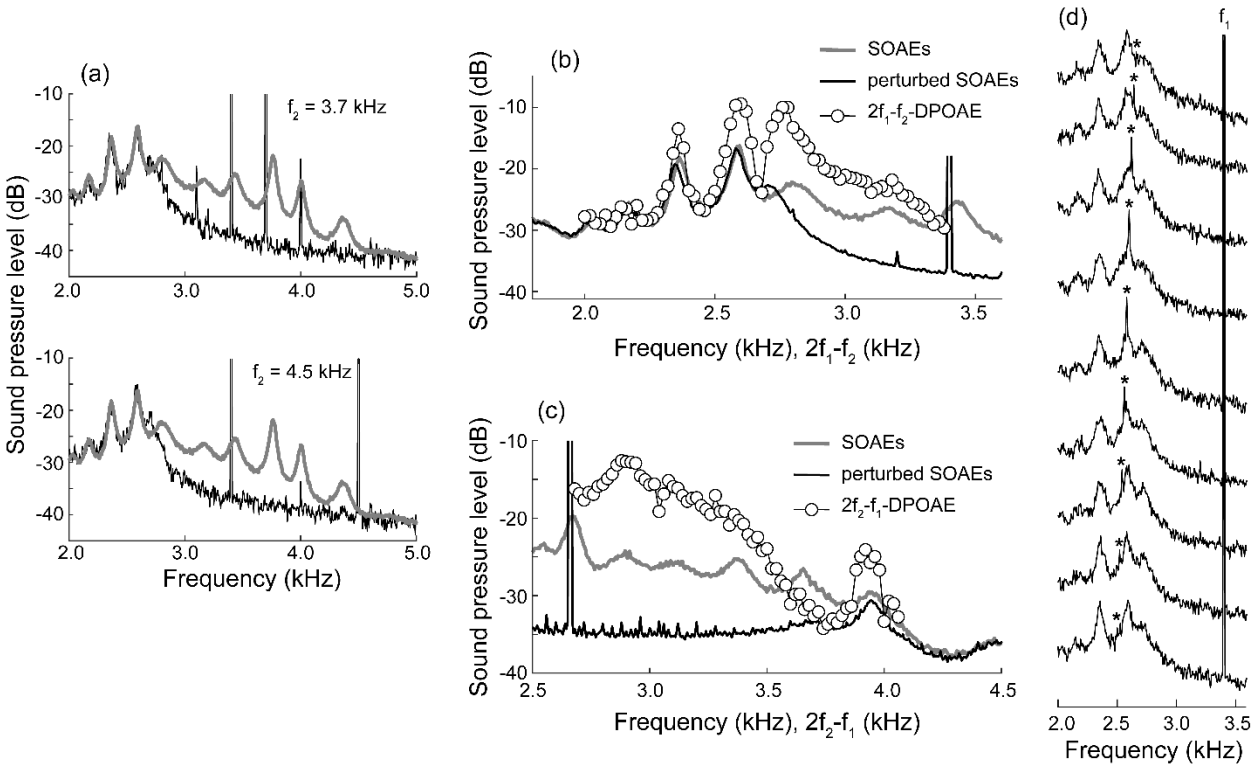


FIGURE 1. (a) The emission amplitude spectrum recorded in the presence of a two-tone stimulus at f_1 = 3.4 kHz, f_2 = 3.7 kHz (*black line, top panel*) or f_1 = 3.4 kHz, f_2 = 4.5 kHz (*black line, bottom panel*) reveals that the perturbed SOAE activity is attenuated from the SOAE level recorded in a quiet environment (*gray lines*). Below 2.7 kHz, the SOAE activity is not affected by the stimuli. (b) The level of $2f_1$-f_2-DPOAE (*open circles*) is plotted as a function of distortion frequency, $2f_1$-f_2. Peaks in the distortion level is observed in the unperturbed part of the SOAE activity and coincide with SOAE peaks. The perturbed SOAE activity (*black line*) is obtained from averaging the emission spectra recorded with f_1 = 3.4 kHz and f_2 varied from 3.42 kHz to 4.8 kHz. (c) Same as (b), showing the level of $2f_2$-f_1-DPOAE (*open circles*) plotted as a function of distortion frequency, $2f_2$-f_1. The perturbed SOAE activity (*black line*) is obtained from averaging the emission spectra recorded with f_2 = 2.66 kHz and f_1 varied from 1.26 kHz to 2.64 kHz. (d) Examples of the emission spectra used in the analysis of (b). From top to bottom, the stimulus frequency f_2 ranges from 4.14 kHz to 4.3 kHz, with f_1 fixed at 3.4 kHz. When the $2f_1$-f_2-DPOAE (*asterisks*) is superposed at a frequency corresponding to an SOAE peak, its level is significantly enhanced above the SOAE level. In all panels, the stimulus intensity level was 50 dB SPL.

Next, we focused on the intensity level of DPOAEs evoked by primary tones with small frequency differences. The stimulus frequency, f_1, were increased from 0.5 kHz to 5.0 kHz at a 50-Hz increment, while f_2 was varied correspondingly at a frequency ratio fixed at f_2/f_1 = 1.1. The intensity levels of both primary tones were fixed at 50 dB SPL. For all stimulus frequencies, the distortions at $2f_1$-f_2 and $2f_2$-f_1 occurred within the frequency range of SOAE suppression (Fig. 2a). We found that the level functions of both $2f_1$-f_2- and $2f_2$-f_1-DPOAEs displayed peaks and pronounced notches across range of the distortion frequencies (Fig. 2b – 2c). However, the center frequencies of the peaks did not coincide with those of SOAE peaks, recorded in the absence of stimulus (Fig. 2d).

We then decreased the body temperature of tokay geckos from 28.6 ºC to 24.3 ºC. At each temperature level, DPOAEs and SOAEs were recorded. Consistent with previous studies, all SOAE peaks exhibited systematic shifts toward lower frequencies upon cooling (Fig. 2d). For each SOAE peak, we determined the difference in its center frequencies at two levels of body temperature and found that the magnitude of the shift increased progressively with the SOAE center frequency. The frequency shift ranged from approximately 50 Hz for SOAE peaks centering below 2.0 kHz to 250 Hz for an SOAE peak centering at 4.8 kHz (Fig. 2e).

A reduction in the body temperature also elicited systematic shifts of the level functions of both $2f_1$-f_2- and $2f_2$-f_1-DPOAEs (Fig. 2b – 2c). The magnitudes of the shifts were determined for the peaks in the DPOAE level functions.

Consistent with the response of SOAEs, the frequency shift increased with the peak frequency and was comparable to those of SOAE peaks centering within the same frequency range (Figure 2(e)).

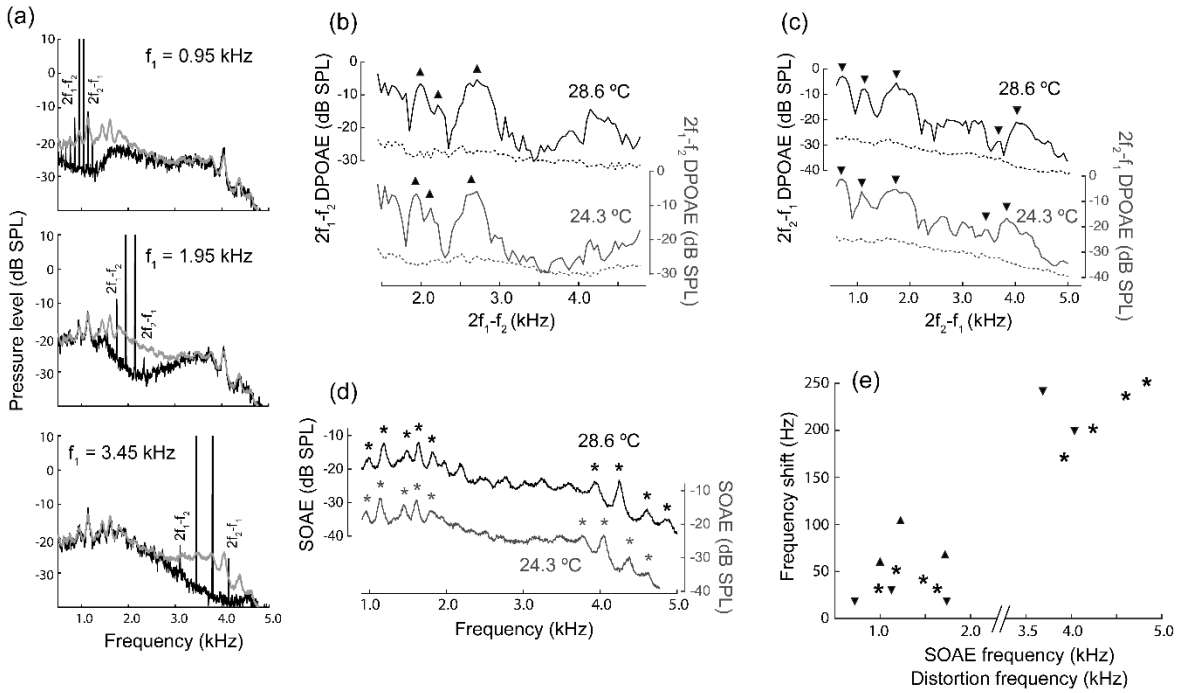


FIGURE 2. (a) The emission spectra (*black lines*) are recorded with two-tone stimuli applied at frequencies f_1 = 0.95 kHz, 1.95 kHz, and 3.45 kHz (from top to bottom), with the ratio, f_2/f_1 = 1.1. The stimulus intensity level was 40 dB SPL. The emission spectra show that DPOAEs are superposed at frequencies within the range of SOAE suppression. Each panel shows the SOAE amplitude spectrum recorded in the absence of stimuli (*gray line*). (b) The intensity levels of $2f_1$-f_2-DPOAE as a function of distortion frequency display multiple peaks (*upright triangles*), shifting toward lower frequencies upon reducing the body temperature from 28.6 ºC (*top*) and 24.3 ºC (*bottom*). The level of DPOAEs recorded at 24.3 ºC over the range of frequencies above 3.0 kHz does not significantly exceed the noise floor (*dashed lines*). (c) Similar characteristics can be observed in the intensity level of $2f_2$-f_1-DPOAE. The peaks in the DPOAE level (*upside down triangles*) shift toward lower frequencies upon cooling. (d) The SOAE spectra recorded at 28.6 ºC (*top*) and 24.3 ºC (*bottom*) also display a frequency shift. *Asterisks* indicate SOAE peaks used in the analysis. (e) The magnitudes of frequency shifts of SOAE peaks (*asterisks*), $2f_1$-f_2-DPOAE peaks (*upright triangles*), and $2f_2$-f_1-DPOAE peaks (*upside down triangles*) are comparable. Note that the plot between 2.0 kHz and 3.5 kHz is omitted due to the lack of discernible peaks in the DPOAE level functions and SOAE spectra.

DISCUSSION

Our results demonstrated the association between the spontaneous activity of the inner ear and the level of DPOAEs evoked by low-intensity stimuli. For DPOAEs whose frequencies were far from those of the primary tones, the distortion level could be enhanced when the distortion frequency was within the bandwidth of an SOAE peak. On the other hand, distortion signals evoked by primary tones with small frequency differences displayed a level function that did not correlate with the SOAE spectrum. However, the DPOAE level functions and SOAE peaks displayed shifts in their frequencies of comparable magnitudes in response to changes in the body temperature level. This suggests that the generation of low-level DPOAEs potentially involve a mechanism that underlies the SOAE production.

Hair cells in the inner ears of lizards have been previously modelled by an array of active nonlinear oscillators, arranged in order of their characteristic frequencies [4-6]. Adjacent oscillators are coupled via an elastic element and viscous fluid. Previous numerical studies of the model suggest that SOAEs could result from synchronization of the

autonomous movements of nearby oscillators which gives rise to a small number of clusters, each contributing to a peak in the SOAE amplitude spectrum.

The chain of coupled nonlinear oscillators potentially provides a theoretical framework that describes the association between DPOAEs and SOAEs. Upon an application of two pure tones, the oscillators with characteristic frequencies near those of the stimuli are entrained to the stimuli and produce distortion signals. Mechanical coupling between the oscillators could be responsible for transferring the distortion energy toward the cluster whose oscillation frequency coincides with the distortion frequency. We postulate that an amplification of the distortion signal might be observed if the oscillators operate near a verge of an instability.

The shift of SOAE peaks in response to a decrease in the level of body temperature may result from a systematic decrease of the oscillators' characteristic frequencies. However, it remains unclear how the level functions of DPOAEs recorded at small primary frequency ratios can be influenced by the alteration of the clusters' oscillation frequencies.

ACKNOWLEDGMENTS

This work is supported by Office of the Permanent Secretary, Ministry of Higher Education, Science, Research and Innovation, Thailand (No. RGNS63-017).

REFERENCES

1. G. A. Manley, C. Köppl, and B.M. Johnstone, J. Acoust. Soc. Am. 93, 2820-33 (1993).
2. G. Taschenberger, L. Gallo, and G. A. Manley, Hear. Res. 91, 87-92 (1995).
3. C. Bergevin, D. Freeman, J. C. Saunders, and C. A. Shera, J. Comp. Physiol. A 194, 665-83 (2008).
4. A. Vilfan and T. Duke, Biophys. J 95, 4622-30 (2008).
5. M. Gelfand, O. Piro, M. O. Magnasco, and A. J. Hudspeth, Plos One 0011116 (2010).
6. B. Epp, H. Wit, and P. van Dijk, AIP Conf. Proc. 1703, 090025 (2015).

Synchronization and Sensitive Signal Detection by Coupled Systems of Chaotic Hair Cells

Justin Faber[1, a)] and Dolores Bozovic[1, 2, b)]

[1)] *Department of Physics & Astronomy, University of California, Los Angeles, California 90095, USA.*
[2)] *California NanoSystems Institute, University of California, Los Angeles, California 90095, USA.*

[a)] *Electronic mail: faber@physics.ucla.edu*
[b)] *Electronic mail: bozovic@physics.ucla.edu.*

Abstract. Auditory and vestibular systems display mechanical coupling of active, oscillating components. The strength and extent of this coupling varies across species and organs, depending on the specific task of each system. It is believed that this mechanical coupling is essential to achieve sufficient sensitivity, frequency selectivity, and temporal resolution for reliable signal detection. These characteristics depend heavily on the ability of the active oscillators to synchronize with each other. It has previously been shown that spontaneous otoacoustic emissions may arise through a dynamical systems phenomenon known is as frequency clustering, where active oscillators synchronize with those of similar characteristic frequency, forming several distinct clusters. In this work, we demonstrate that nonisochronicity enhances the stability of these clusters, yielding larger, sharper spectral peaks. Further, we show that nonisochronicity enhances the ability of the system to detect weak signals. We propose that auditory and vestibular systems utilize nonisochronicity as a mechanism of facilitating synchronization among oscillating components to enhance signal detection.

INTRODUCTION

The auditory and vestibular systems are astounding signal detectors. They respond to sounds, vibrations, and accelerations that produce motion on the order of angstroms, comparable to or below the levels of intrinsic noise in the system [1]. Humans are able to resolve two tones that differ by only 0.1% in frequency [2]. We are also able to resolve two acoustic impulses that are separated by only 10 microseconds, where the stimulus waveform is presented simultaneously into both ears [3]. These characteristics of sensitivity, frequency selectivity, and temporal resolution are essential for complex tasks such as localizing sound sources and comprehending speech. How these systems reliably achieve such signal detection properties remains an open question, and the physics of hearing continues to be an active area of research [4].

Mechanical detection of sound, vibration, and acceleration is performed by sensory hair cells. These specialized cells get their name from the hair-like stereovilli that protrude from their apical surfaces. This cluster of interconnected stereovilli is collectively called the hair bundle, and performs the essential task of converting external mechanical signals into electrically conveyed information [5, 6, 7]. The external signal pivots the hair bundle, modulating the tension in the connective tip links between adjacent stereovilli. This tension modulates the opening probability of the mechanically sensitive ion channels embedded in tips of the stereovilli and connected to the tip link. The hair bundles are further mechanically coupled to each other through overlying structures, with the strength and extent of coupling varying with the species and the role of the sensory organ.

Auditory detection has been shown to require an active, energy-consuming process in order to overcome the dissipation in the surrounding fluid and to achieve such remarkable signal detection [8]. This active process manifests itself in several ways, one of which is the occurrence of autonomous oscillations of hair bundles in several species [9, 10, 11]. These spontaneous oscillations were shown to violate the fluctuation dissipation theorem, proving them to be active [12]. While the role of these spontaneous oscillation *in vivo* is unknown, prior studies have suggested that they could be utilized to amplify external signals and thereby improve signal detection [13].

Another manifestation of the active process is the emission of faint tones from the external ears of living animals [14]. These spontaneous otoacoustic emissions (SOAEs) exhibit several sharp peaks in their power spectra and are metabolically sensitive, indicating that they utilize an energy consuming process. SOAEs are ubiquitous across the vertebrates and serve as one of the probes for studying the active process of hearing. Although they have beed used as a diagnostic for hearing-related disorders, the mechanism by which SOAEs arise remains controversial [15, 16, 17]. One theory suggests that they are locally generated by the spontaneous activity of hair cells, through a phenomenon known as frequency clustering [18, 19]. Dynamical-systems theory shows that frequency clustering can occur when active oscillators with frequency dispersion are coupled together. The oscillators synchronize with one of several distinct clusters, forming sharp peaks in the power spectrum of the collective motion.

The dynamics of active hair cells have been modeled using the normal form equation of the Hopf bifurcation. This simple, empirical model reproduces many of the experimentally observed characteristics of the auditory and vestibular systems, such as spontaneous oscillations, the nonlinear response function, as well as the extreme sensitivity and frequency selectivity [20, 21]. In order to maintain these signal detection properties, prior studies have assumed the individual oscillators are poised near the Hopf bifurcation. This raises the question of how the system maintains proximity to this bifurcation in the presence of noise, changing environments, and ultimately heterogeneity in the parameters of the oscillators. To circumvent this issue, feedback mechanisms have been proposed in order to keep the system poised at criticality [22], or to expand the parametric range of the critical behavior [23]. Another issue with being in the immediate vicinity of the Hopf bifurcation is the dynamical-systems phenomenon known as critical slowing down. The closer the system gets to the bifurcation, the more unstable it becomes, requiring a long return time to its natural state after cessation of external signals. This seems inconsistent with the high temporal resolution exhibited by the auditory system.

To answer both of these questions, we have previously proposed that hair cells are not poised in the immediate vicinity of the Hopf bifurcation, but are rather positioned sufficiently far in the oscillatory regime so as to avoid critical slowing down. We demonstrated that active Hopf oscillators could achieve high sensitivity and the requisite temporal resolution when poised in the chaotic regime, as chaotic systems respond very rapidly and are extremely sensitive to external perturbations [24]. We also demonstrated, experimentally, the presence of an underlying chaotic attractor in the spontaneous oscillations of living, active hair bundles. Further, when we modulated the chaoticity of the innate bundle motion, the signal detection properties changed in the same manner as in our numerical model.

For Hopf oscillators, chaos arises in the presence of noise when the system is nonisochronous. When a limit cycle oscillator exhibits the same characteristic frequency for every amplitude of oscillation, it is called isochronous. Nonisochronous systems, by contrast show a characteristic frequency that depends on the amplitude of oscillation. Using numerical simulations and an analytic calculation, we previously demonstrated that the level of chaos increases with increasing nonisochronicity, with the two being directly proportional for weakly nonisochronous systems. This dynamical instability, which gives rise to chaos, also enhances the ability of the system to entrain to external signals, improving its sensitivity and temporal resolution [25]. However, because nonisochronous oscillators are so easily entrained, they are not frequency selective at the level of individual oscillators.

We extended this study from the individual oscillator to the coupled system and explored the role of nonisochronicity in a model of the sacculus [26]. Coupling was assumed to be strong, and we therefore modeled the system with mean-field coupling and significant frequency dispersion in the characteristic frequencies. We demonstrated that nonisochronicity of the active hair cells in this model gives rise to enhanced synchronization. The inter-cell synchronization resulted in significant improvement in sensitivity and frequency selectivity of the system and minor improvement in the temporal resolution. Further, we demonstrated experimentally that spontaneously oscillating hair bundles of the bullfrog sacculus synchronize when mechanically coupled with artificial membranes, even with significant frequency dispersion between neighboring oscillators. We proposed that this system achieves such robust synchronization because it is comprised of nonisochronous oscillators.

In the current work, we extend our focus to an array of coupled Hopf oscillators with a tonotopic arrangement of the characteristic frequencies. This serves as a simple model for auditory organs, such as the cochlea and the basilar papilla. Significant frequency dispersion and dissipative coupling can lead to portions of the system oscillating incoherently or not oscillating at all. However, when we introduce nonisochronicity through random selection of the cubic parameter, the system stabilizes, synchronization is enhanced, and the oscillations grow larger. This results in more pronounced SOAE peaks. Further, we show that this enhanced synchronization results in significant improvement in the sensitivity and frequency selectivity, with slight improvement in the response time. We propose that nonisochronicity is a ubiquitous phenomenon to auditory and vestibular systems that is utilized to stabilize synchronization between oscillatory components and to enhance signal detection.

NUMERICAL MODEL OF COUPLED HAIR CELLS

The dynamics of the j^{th} hair cell in the coupled system are described using the complex variable, $z_j(t) = x_j(t) + iy_j(t)$, and are assumed to be governed by the normal form equation for the supercritical Hopf bifurcation,

$$\frac{dz_j}{dt} = (\mu + i\omega_j)z_j - (\alpha + i\beta_j)|z_j|^2 z_j + (d_R + id_I)(z_{j+1} + z_{j-1} - 2z_j) + \eta_j(t) + F(t). \tag{1}$$

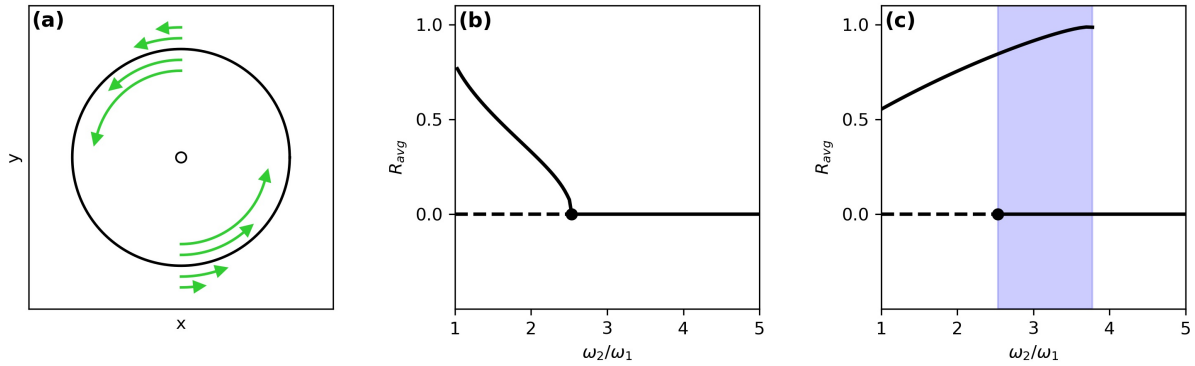


FIGURE 1. (a) Illustration of a nonisochronous oscillator. The large black circle represents a stable limit cycle, while the point in the center represents an unstable fixed point. The green arrows illustrate the different velocities that occur at different radii. (b) Bifurcation diagram of two coupled Hopf oscillators with $d_R = 2$ and $\beta_1 = \beta_2 = 7$. R_{avg} represents the average amplitude of the two oscillators. (c) Bifurcation diagram of two coupled Hopf oscillators with $d_R = 2$, $\beta_1 = -7$, and $\beta_2 = 7$. The purple shading indicates the region of bistability, where both the amplitude-death state and the oscillatory, synchronized state are stable.

x_j represents the hair-bundle position, while y_j represents a scaled version of the hair-bundle velocity [18]. μ controls the proximity of the system to the Hopf bifurcation. In the absence of coupling, $\mu > 0$ results in active, oscillatory behavior, while $\mu > 0$ yields quiescent, fixed-point dynamics. The natural frequency of oscillation at the Hopf bifurcation ($\mu = 0$) is given by ω_j. In isochronous systems, $\beta_j = 0$, and the characteristic frequency is constant for any amplitude of oscillation. However, in the nonisochronous cases, the phase dynamics in the absence of noise, stimulus, and coupling is given by $\frac{d\theta_j}{dt} = \omega_j - \beta_j r_j^2$, where $r_j^2 = x_i^2 + y_i^2$. In the oscillatory regime, the limit-cycle amplitude is $r_j = \sqrt{\frac{\mu}{\alpha}}$. Using this, we define the limit-cycle frequency to be $\Omega_j = \omega_j - \frac{\beta_j \mu}{\alpha}$.

The nonlinearity is controlled by α and β_j, with the latter governing the degree of nonisochronicity. In previous studies, β_j has often been set to zero, rendering the system isochronous. However, we choose to study the more general case where β_j may be non-zero. External acoustic stimulus acting on the hair bundles is represented by the real variable, $F(t)$, where the forcing is simultaneously presented to all of the oscillators. To represent white-noise stimulus or biological noise, we use the stochastic variables, $\eta_j(t)$, where each oscillator experiences a different realization of noise, with independent real and imaginary parts. The correlation functions are $\langle Re(\eta_j(t))Re(\eta_j(t'))\rangle = \langle Im(\eta_j(t))Im(\eta_j(t'))\rangle = 2D\delta(t-t')$, where D is the noise strength of the system. We use nearest-neighbor, complex coupling, where d_R and d_I represent the dissipative and elastic coupling strengths, respectively [18].

Throughout the study, we keep $\mu = \alpha = 1$, poising the individual oscillators in their active, oscillatory state. We set $d_I = -2$ and vary d_R, thereby exploring the effects of dissipation on the dynamics of the system. We choose a tonotopic arrangement of the characteristic frequencies in order to model auditory organs. The limit-cycle frequencies increase exponentially from $\Omega_1 = 5$ to $\Omega_{50} = 25$. To incorporate nonisochronicity into the system, all β_j values are independently chosen from a Gaussian distribution centered at 0 with a standard deviation of σ_β. All simulations are performed without additive noise except when stated otherwise. In the noiseless case, the differential equations are solved using the 4th-order Runge-Kutta method with time steps of 10^{-3}. The 2nd-order Runge-Kutta method is used for simulations that include stochastic noise.

RESULTS

We explore the effects of nonisochronicity on the dynamics of the coupled system. When an individual oscillator is nonisochronous, its frequency of oscillation depends on its amplitudes, meaning that a deviation away from the stable limit cycle will result in the system either speeding up or slowing down in frequency [Fig. 1a]. The degree of nonisochronicity is controlled by β_j, with larger magnitudes yielding more nonisochronous behavior. This property causes an oscillator to more easily become entrained to an external stimulus, as it can adjust its amplitude to tune its frequency, thus minimizing differences between its characteristic frequency and that of the applied stimulus. For the same reason, nonisochronicity can enhance synchronization between multiple oscillators.

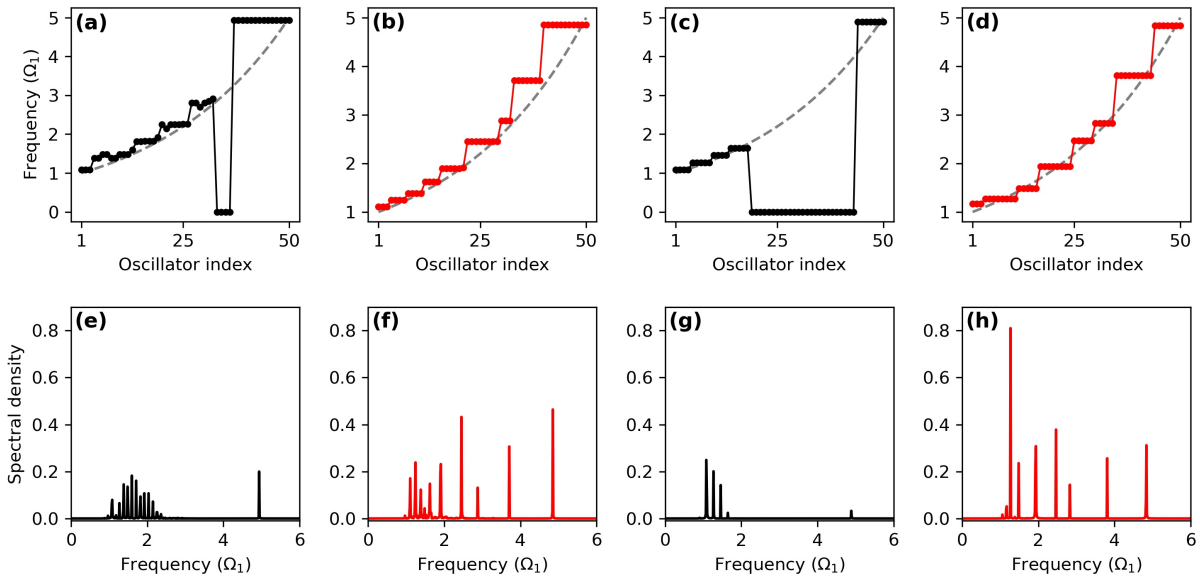


FIGURE 2. (a)-(d) Spontaneous oscillation frequency as a function of oscillator location in the coupled system. The grey, dashed curves indicate each oscillators limit-cycle frequency, Ω_j. Quiescent oscillators are denoted by a frequency of zero. Isochronous systems are shown in black, while nonisochronous systems are shown in red. Model parameters are as follows: (a) $\sigma_\beta = 0$, $d_R = 0.5$, (b) $\sigma_\beta = 2$, $d_R = 0.5$, (c) $\sigma_\beta = 0$, $d_R = 2$, (d) $\sigma_\beta = 2$, $d_R = 2$. (e)-(h) Power spectral densities corresponding to the systems in (a)-(d), respectively.

We first study the dynamics of just two coupled nonisochronous oscillators. With two oscillators in their active regime ($\mu > 0$), the system will occupy one of three states [27]. For weak coupling and small dispersion of the two characteristic frequencies, the system occupies the incoherent state, where unsynchronized dynamics occur. For sufficiently high coupling strength and small frequency dispersion, the oscillators synchronize. However, for high coupling strength and large frequency dispersion, the system occupies the amplitude-death state, and the motion is quenched. If this state occurs at the origin of phase space, then the β_j values will have no effect on the stability of this state as they appear in the cubic term, which is irrelevant near the origin. As we increase the frequency dispersion in a system of two oscillators with identical β_j, we see that the amplitude of oscillation is reduced and eventually crosses a bifurcation, beyond which the amplitude-death state is the only stable solution [Fig. 1b]. However, when we introduce heterogeneity in the β_j parameters, the synchronized state persists at larger frequency dispersions and even coexists with the amplitude-death state [Fig. 1c].

Normally, coherence and cooperation between active oscillators is improved by similarities in the model parameters. However, in this case, heterogeneity of the β_j parameter enhances synchronization. This can be explained by the fact that coupling tends to pull the dynamics inwards, closer to the origin, thereby reducing the average amplitude of oscillation [26]. If the two oscillators have identical β_j, then the coupling will have the same effect on them. It will either raise or lower the oscillation frequencies by the same amount, which may not improve synchronization. However, if the oscillators have different β_j, then the coupling will shift the oscillation frequencies by different amounts and possibly in different directions, causing them to converge towards the middle. With the introduction of heterogeneity, the system is therefore more likely to show enhanced synchronization.

Next, we study the effects of nonisochronicity on the dynamics of a chain of 50 Hopf oscillators with complex coupling and a tonotopic arrangement of the characteristic frequencies. When the oscillators are isochronous, we observe states where part of the system oscillates incoherently or does not even oscillate at all [Fig. 2a, c]. The number of oscillators in this amplitude-death state grows with increasing dissipative coupling strength, d_R. However, when we introduce nonisochronicity into the system, the cluster state is stabilized. The oscillators become better entrained to their neighbors, and they all avoid the amplitude-death state [Fig. 2b, d]. As a result of the clusters becoming more stable, the peaks in the power spectra also become more pronounced [Fig. 2e-h].

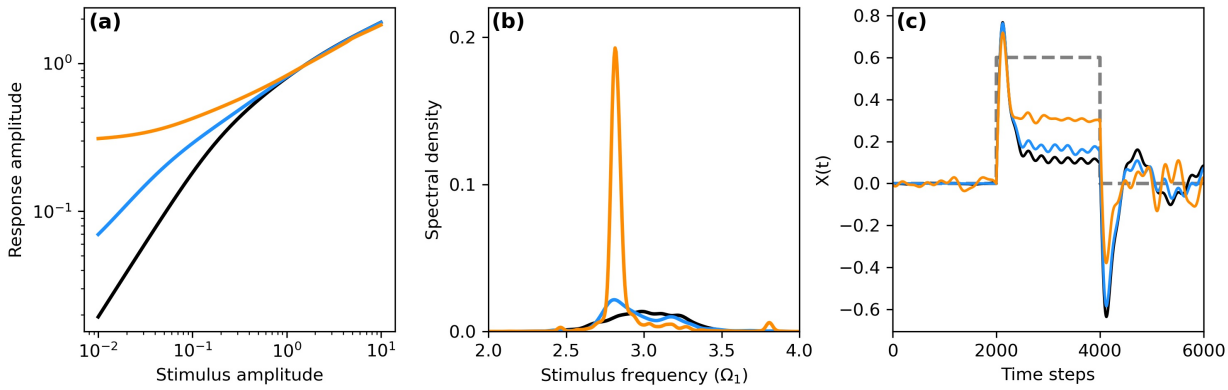


FIGURE 3. (a) Phase-locked amplitude of the 50-oscillator coupled system in response to sinusoidal stimulus. The stimulus frequency was chosen to be Ω_{32}, which is near the center of the region prone to amplitude death. Phase-locked responses to the stimulus were calculated from oscillators in the range $j = 30$ to $j = 34$ and averaged. (b) Spectral response of the 32nd oscillator to white Gaussian noise stimulus with $D = 0.003$. (c) Average response of the system to a large step-like stimulus with an amplitude of $f_0 = 10$. This response is the average over all oscillators and over 100 different, randomly-selected initial conditions of the system. The grey, dashed lines illustrate the stimulus waveform. For all panels, black, blue, and orange curves correspond to $\sigma_\beta = 0$, 0.5, and 2, respectively.

To explore the effects of nonisochronicity on the system's ability to detect external signals, we apply a sinusoidal stimulus, $F(t) = f_0 \sin(\omega_s t)$, at a frequency well within the detection range of the system, at various stimulus amplitudes. We calculate the average phase-locked response of the oscillator at the stimulus frequency and its two neighbors on either side, since they are likely to entrain to the stimulus as well. For large stimulus amplitudes ($f_0 > 1$), the oscillators are forced to entrain, and nonisochronicity has no effect on the sensitivity. However, for weaker stimulus ($f_0 < 1$), the nonisochronous system detects the external signals more easily, with an improvement of up to an order of magnitude in amplitude [Fig. 3a]. This increase in sensitivity is a direct result of the stabilizing effect that nonisochronicity has on the synchronization of the coupled system. When active oscillators are synchronized, background noise from each oscillator is averaged out, and the signal-to-noise ratio increases [28]. This background noise could be intrinsic to each of the oscillators, or as in the present case, it could be a byproduct of the many other active components oscillating at frequencies not relevant to the signal of interest.

Next, we measure the frequency selectivity by providing a white Gaussian noise stimulus to the system and looking at the spectral response of one of the oscillators. This response yields a fair approximation of the frequency selectivity of the nonlinear system, provided that the stochastic stimulus is weak [29]. For the same reasons as discussed previously, we find that frequency selectivity is greatly improved when nonisochronicity is included. The synchronized state is stabilized, providing a narrow band-pass filter to the external noise [Fig. 3b].

Lastly, we provide a large, step-like stimulus, $F(t) = f_0\Big(\Theta(t-\tau_1) - \Theta(t-\tau_2)\Big)$, where Θ is the Heaviside step function. This type of stimulus is used to characterize the response time and approximate the temporal resolution to brief stimuli. We calculate the average response over all of the oscillators, $X(t) = \frac{1}{N}\Sigma_{j=1}^{N} x_j(t)$. After averaging the response over many randomly-selected initial conditions, we observe an exponential decay. Nonisochronicity causes the system to respond slightly faster, consistent with a numerical model of the sacculus [26]. Further, the amplitude at which the plateau occurs increases for increasing σ_β, indicating that nonisochronicity also makes the system more responsive to step-like stimulus waveforms [Fig. 3c].

DISCUSSION

Auditory and vestibular systems exhibit a wide range of behaviors that can be mapped onto different classes of systems studied using bifurcation theory and nonlinear dynamics [30, 31, 32]. The active hair cells of these sensory systems have been shown to display Hopf bifurcations, saddle-node on an invariant circle (SNIC) bifurcations [33], and the quasiperiodic transition to chaos [34], among others [27, 35]. Individual hair bundles display quiescent, fixed-point

dynamics as well as oscillatory behavior. Active hair bundles have also been observed to display a bursting behavior, where they flicker between oscillatory and quiescent states. When mechanically coupled, hair bundles likewise exhibit a wide range of dynamical behaviors. Complete synchronization has been observed, as well as partial synchronization in the form of cluster states and chimera states [36]. The otolithic membrane in the amphibian sacculus imposes strong coupling, pushing the system into a quiescent, amplitude-death state [27]. How these different dynamical states and bifurcations shape the response of the system to achieve reliable signal detection remains an open question.

In the present work, we demonstrated that heterogeneity in the nonisochronicity parameters of active oscillators, leads to a bistable regime where the amplitude-death state and the synchronized, oscillatory state coexist. We further showed that this results in stabilization of the cluster state in the coupled system, and more pronounced spontaneous otoacoustic emissions in the numerical model of auditory organs. Lastly, we demonstrated that this stabilized cluster state proves to be a more reliable signal detector, as it is more responsive to sinusoidal and step-like stimuli, more frequency selective, and slightly faster to respond to an abrupt stimulus.

The function of each sensory organ likely determines the most favorable dynamical state. Saccular hair cells are responsible for detecting low-frequency vibrations, as well as accelerations of the animal. Such an organ may not need much frequency selectivity. However, the human cochlea requires sharp frequency selectivity, as it is essential for complex tasks, such as comprehending speech. Achieving sensitivity, frequency selectivity, and temporal resolution sufficient for the demands of each sensory organ is a remarkable feat, as these biological systems are faced with inherent noise, changing environments, and imprecise structural parameters. One would expect that such heterogeneity of the dynamical system parameters would weaken synchronization and harm the performance of the detector. However, our numerical model shows that heterogeneity in nonisochronicity actually stabilizes synchronization and enhances signal detection. We propose that nonisochronicity is a ubiquitous feature of auditory and vestibular systems and is utilized as a tuning mechanism in order to detect signals reliably.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of NSF Biomechanics and Mechanobiology under Grant No. 1916136.

REFERENCES

1. A. J. Hudspeth, "Integrating the active process of hair cells with cochlear function," Nat. Rev. Neurosci. **15**, 600 (2014), https://doi.org/10.1038/nrn3786.
2. M. F. Spiegel and C. S. Watson, "Performance on frequency-discrimination tasks by musicians and nonmusicians," J. Acoust. Soc. **76**, 1690 (1984), https://doi.org/10.1121/1.391605.
3. B. Leshowitz, "Measurement of the two-click threshold," J. Acoust. Soc. **49**, 462 (1971), https://doi.org/10.1121/1.1912374.
4. T. Reichenbach and A. J. Hudspeth, "The physics of hearing: fluid mechanics and the active process of the inner ear," Rep. Prog. Phys. **77**, 076601 (2014), https://doi.org/10.1088/0034-4885/77/7/076601.
5. M. LeMasurier and P. G. Gillespie, "Hair-cell mechanotransduction and cochlear amplification," Neuron **48**, 403 (2005), https://doi.org/10.1016/j.neuron.2005.10.017.
6. M. A. Vollrath, K. Y. Kwan, and D. P. Corey, "The micromachinery of mechanotransduction in hair cells," Annu. Rev. Neurosci. **30**, 339 (2007), https://doi.org/10.1146/annurev.neuro.29.051605.112917.
7. D. Ó Maoiléidigh and A. J. Ricci, "A bundle of mechanisms: Inner-ear hair-cell mechanotransduction," Trends Neurosci. **42**, 221 (2019), https://doi.org/10.1016/j.tins.2018.12.006.
8. A. J. Hudspeth, "Making an effort to listen: mechanical amplification in the ear," Neuron **59**, 530 (2008), https://doi.org/10.1016/j.neuron.2008.07.012.
9. M. E. Benser, R. E. Marquis, and A. J. Hudspeth, "Rapid, active hair bundle movements in hair cells from the bullfrog's sacculus," J. Neurosci. **16**, 5629 (1996), https://doi.org/10.1523/JNEUROSCI.16-18-05629.1996.
10. P. Martin, D. Bozovic, Y. Choe, and A. J. Hudspeth, "Spontaneous oscillation by hair bundles of the bullfrog's sacculus," J. Neurosci. **23**, 4533 (2003), https://doi.org/10.1523/JNEUROSCI.23-11-04533.2003.
11. A. C. Crawford and R. Fettiplace, "The mechanical properties of ciliary bundles of turtle cochlear hair cells," J. Physiol. **364**, 359 (1985), https://doi.org/10.1113/jphysiol.1985.sp015750.
12. P. Martin, A. J. Hudspeth, and F. Jülicher, "Comparison of a hair bundle's spontaneous oscillations with its response to mechanical stimulation reveals the underlying active process," Proc. Natl. Acad. Sci. U.S.A. **98**, 14380 (2001), https://doi.org/10.1073/pnas.251530598.
13. P. Martin and A. J. Hudspeth, "Active hair-bundle movements can amplify a hair cell's response to oscillatory mechanical stimuli," Proc. Natl. Acad. Sci. U.S.A. **96**, 14306 (1999), https://doi.org/10.1073/pnas.96.25.14306.
14. D. T. Kemp, "Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea," Arch. Oto-Rhino-Laryngol. **224**, 37 (1979), https://doi.org/10.1007/BF00455222.

15. C. A. Shera, "Mechanisms of mammalian otoacoustic emission and their implications for the clinical utility of otoacoustic emissions," Ear Hear. **25**, 86 (2004), https://doi.org/10.1097/01.AUD.0000121200.90211.83.
16. C. Bergevin, G. A. Manley, and C. Köppl, "Salient features of otoacoustic emissions are common across tetrapod groups and suggest shared properties of generation mechanisms," Proc. Natl. Acad. Sci. U.S.A. **112**, 3362 (2015), https://doi.org/10.1073/pnas.1418569112.
17. Y. Roongthumskul, D. Ó Maoiléidigh, and A. J. Hudspeth, "Bilateral spontaneous otoacoustic emissions show coupling between active oscillators in the two ears," Biophys. J. **116**, 2023 (2019), https://doi.org/10.1016/j.bpj.2019.02.032.
18. A. Vilfan and T. Duke, "Frequency clustering in spontaneous otoacoustic emissions from a lizard's ear," Biophys. J. **95**, 4622 (2008), https://doi.org/10.1529/biophysj.108.130286.
19. F. Fruth, F. Jülicher, and B. Lindner, "An active oscillator model describes the statistics of spontaneous otoacoustic emissions," Biophys. J. **107**, 815 (2014), https://doi.org/10.1016/j.bpj.2014.06.047.
20. V. M. Eguiluz, M. Ospeck, Y. Choe, A. J. Hudspeth, and M. O. Magnasco, "Essential nonlinearities in hearing," Phys. Rev. Lett. **84**, 5232 (2000), https://doi.org/10.1103/PhysRevLett.84.5232.
21. A. Kern and R. Stoop, "Essential role of couplings between hearing nonlinearities," Phys. Rev. Lett. **91**, 128101 (2003), https://doi.org/10.1103/PhysRevLett.91.128101.
22. S. Camalet, T. Duke, F. Jülicher, and J. Prost, "Auditory sensitivity provided by self-tuned critical oscillations of hair cells," Proc. Natl. Acad. Sci. U.S.A. **97**, 3183 (2000), https://doi.org/10.1073/pnas.97.7.3183.
23. A. R. Milewski, D. Ó Maoiléidigh, J. D. Salvi, and A. J. Hudspeth, "Homeostatic enhancement of sensory transduction," Proc. Natl. Acad. Sci. U.S.A. **114**, E6794 (2017), https://doi.org/10.1073/pnas.1706242114.
24. J. Faber and D. Bozovic, "Chaotic dynamics enhance the sensitivity of inner ear hair cells," Sci. Rep. **9**, 18394 (2019), https://doi.org/10.1038/s41598-019-54952-y.
25. J. Faber and D. Bozovic, "Noise-induced chaos and signal detection by the nonisochronous hopf oscillator," Chaos **29**, 043132 (2019), https://doi.org/10.1063/1.5091938.
26. J. Faber, H. Li, and D. Bozovic, "Synchronization and chaos in systems of coupled inner-ear hair cells," Phys. Rev. Res. **3**, 013266 (2021), https://doi.org/10.1103/PhysRevResearch.3.013266.
27. K.-J. Kim and K.-H. Ahn, "Amplitude death of coupled hair bundles with stochastic channel noise," Phys. Rev. E **89**, 042703 (2014), https://doi.org/10.1103/PhysRevE.89.042703.
28. K. Dierkes, F. Jülicher, and B. Lindner, "A mean-field approach to elastically coupled hair bundles," Eur. Phys. J. E **35**, 37 (2012), https://doi.org/10.1140/epje/i2012-12037-6.
29. V. Z. Marmarelis, "Coherence and apparent transfer function measurements for nonlinear physiological systems," Ann. Biomed. Eng. **16**, 143–157 (1988), https://doi.org/10.1007/BF02367386.
30. A. B. Neiman, K. Dierkes, B. Lindner, L. Han, and A. L. Shilnikov, "Spontaneous voltage oscillations and response dynamics of a hodgkin-huxley type model of sensory hair cells," J. Math. Neurosci. **1**, 11 (2011), https://doi.org/10.1186/2190-8567-1-11.
31. S. Ji, D. Bozovic, and R. Bruinsma, "Amphibian sacculus and the forced kuramoto model with intrinsic noise and frequency dispersion," Phys. Rev. E **97**, 042411 (2018), https://doi.org/10.1103/PhysRevE.97.042411.
32. Y. Roongthumskul, J. Faber, and D. Bozovic, "Dynamics of mechanically coupled hair-cell bundles of the inner ear," Biophys. J. **120**, 205 (2021), https://doi.org/10.1016/j.bpj.2020.11.2273.
33. L. Fredrickson-Hemsing, C. E. Strimbu, Y. Roongthumskul, and D. Bozovic, "Dynamics of freely oscillating and coupled hair cell bundles under mechanical deflection," Biophys. J. **102**, 1785 (2012), https://doi.org/10.1016/j.bpj.2012.03.017.
34. J. Faber and D. Bozovic, "Chaotic dynamics of inner ear hair cells," Sci. Rep. **8**, 3366 (2018), https://doi.org/10.1038/s41598-018-21538-z.
35. J. D. Salvi, D. Ó. Maoiléidigh, B. A. Fabella, M. Tobin, and A. J. Hudspeth, "Control of a hair bundle's mechanosensory function by its mechanical load," Proc. Natl. Acad. Sci. U.S.A. **112**, E1000 (2015), https://doi.org/10.1073/pnas.1501453112.
36. J. Faber and D. Bozovic, "Chimera states and frequency clustering in systems of coupled inner-ear hair cells," Chaos **31**, 073142 (2021), https://doi.org/10.1063/5.0056848.

Nonlinear Effects Basal to the Best Place Manifest in the Reticular Lamina's Response due to its Low Impedance Relative to the Basilar Membrane

George Samaras,[a] Haiqi Wen,[b] and Julien Meaud[c]

George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology
801 Ferst Dr NW, Atlanta, GA 30318

[a] *Electronic mail: gsamaras3@gatech.edu*
[b] *Electronic mail: hwen@gatech.edu*
[c] *Corresponding author: julien.meaud@me.gatech.edu*

Abstract. Based on a plethora of basilar membrane vibration measurements, it had been presumed that cochlear nonlinearity due to outer hair cell (OHC) forces is limited to the peak region of a given stimulus. However, recent experiments showing two-tone suppression of the reticular lamina (RL) indicate that OHCs provide electromotile feedback over a much broader basal region. In this work, a computational model of the cochlea containing organ of Corti (OoC) structures, a fluid domain coupled to the basilar membrane (BM), and an electrical model to represent OHC activity is calibrated to experimental data for the mechanical and electrical responses to a pure tone and then used to make two-tone suppression predictions. These predictions are in line with recent experimental measurements, capturing extended nonlinearity in the RL response compared to that of the BM. By comparing the OHC force of pure tone and two-tone model predictions, we examine the root cause of broader nonlinearity seen in the response of OoC structures at the top end of the OHCs.

INTRODUCTION

Two-tone suppression (TTS) is a phenomenon that arises due to cochlear nonlinearity, and specifically refers to the reduction in the cochlear response due to a pure tone stimulus presented alone compared to when that same stimulus is presented in the presence of a second tone. In this case, the original stimulus is referred to as a probe tone and the second as a suppressor tone. The underlying mechanism of this phenomenon is saturation of the mechano-electrical transduction (MET) channels of the outer hair cells (OHCs), whereby the mechanical response to the suppressor tone saturates the OHCs, reducing the potency of their feedback to the probe tone.

Experimentally, suppression of the basilar membrane (BM) is well-studied and established. Ruggero et al. [1] demonstrates the effect of a variety of probe and suppressor frequency and level combinations in chinchillas. Computational work has enriched understanding of the phenomenon [2, 3, 4]. With recent developments in experimental techniques allowing for extensive micromechanical measurements, characterization of suppression in the response of the reticular lamina (RL) has been made possible. Dewey et al. [5] recently showed that, in the mouse cochlea, high-frequency suppressors (peaking basal to probe best place (BP)) can suppress the RL response but not the BM response at sub-BP locations. This challenges the notion that OHCs only provide active feedback in the immediate region near a particular response's BP, and instead indicates that OHC feedback extends to a broader basal region. Nevertheless, this extended feedback does not manifest in BM suppression, indicating that the mechanics of the cochlea are such that OHC feedback at basal locations affects structures at the top end of the OHCs only.

In this study, a computational model of the gerbil cochlea with a high level of physiological detail is used to make predictions of two-tone suppression for a range of suppressor frequencies and levels, and to assess whether these predictions are in line with recent experimental findings. From there, the flexibility afforded by the computational environment allows further exploration of the underlying force that underpins the suppression phenomenon.

MODEL

Physiologically Based Model

The gerbil model used in this study is heavily motivated by the detailed physiology of the cochlea [6]. As seen in Figure 1A, it contains structural degrees of freedom at both ends of the organ of Corti (OoC) complex, with the BM

transverse and TM (both radial and transverse) motion represented. The BM is coupled to a fluid domain that is broken into two ducts to represent the scala tympani (ST) and scala vestibuli (SV) respectively, yielding the characteristic BM traveling wave as seen in Figure 1B. Detailed micromechanics relate motion of the structural DOFs to the OHC hair bundle (HB) which regulates the mechanoelectric transduction (MET) current. As such, an electrical model containing electrical longitudinal cables is tied to the structural DOFs, to yield voltages across the OHCs (Figure 1C). The model's nonlinearity stems from the nonlinear relationship between the HB displacement and the MET current. Somatic electromotility is also represented, via a reciprocal relation between the OHC somatic current and the OHC velocity, which causes OHCs to contract as they depolarize. The fast Fourier transform (FFT) is performed on steady-state time-domain model responses to the two-tone stimuli, to yield probe and suppressor responses in the frequency domain.

The model was previously calibrated based on pure tone measurements of the ST pressure and local cochlear microphonics (LCM) [6] at the 23.5 kHz best place. Further calibration using pure tone mechanical measurements of the BM and RL [7] at the 26 kHz best place resulted in the model used in this study, which is identical to the one in [6] other than higher damping in the TM-shear degree of freedom ($c_{tms} = 2.4 \cdot 10^{-2}$ N-s/m^2, previously $3 \cdot 10^{-3}$ N-s/m^2) and higher MET conductance to account for loss of sensitivity due to higher damping. The motivation for increased TM-shear damping was to eliminate a sharp notch in the RL response in the near-BF region, as this notch is not evident in the experimental data.

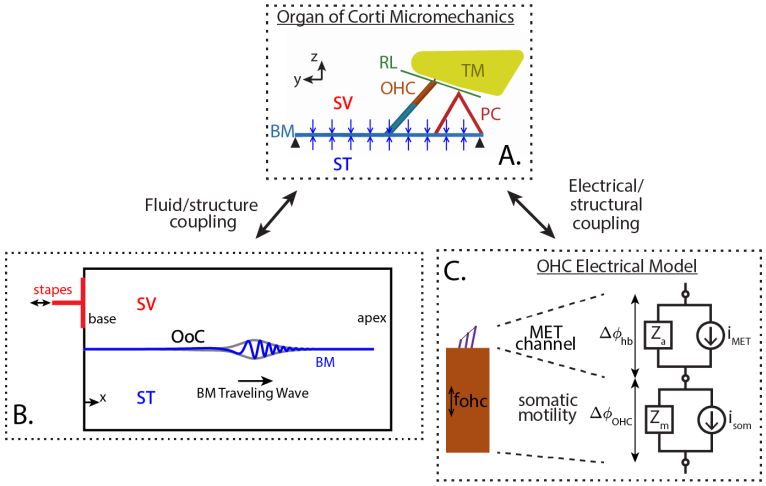


FIGURE 1. 3-D Box model with mechanical, acoustic, and electrical domains. The BM structure is coupled to the ST and SV fluid (A), resulting in a longitudinal BM traveling wave (B). The OHCs in the OoC provide cochlear activity via the OHC electrical model (C).

RESULTS

Pure tone model predictions

Figure 2 shows a side-by-side comparison of the experimental and model responses for the BM and RL displacements and the LCM. The nonlinearity in the BM model response is limited to near-BF frequencies (Fig. 2A-B), in contrast to that of the RL and LCM response, which extends to sub-BF regions (Fig. 2C-D, 2E-F). These characteristics are in line with experimental data. Similarly, the LCM model response is in good agreement with the experimental data as well, with both the degree and the range of nonlinearity being captured. The LCM serves as an indicator of cochlear activity, and hence has broader nonlinearity as well, closer to that of the RL than the BM. In Fig. 2B and 2C, the passive model response (MET transduction eliminated) is also shown (this is representative of postmortem experiments). The sub-BF 80-dB BM response (Fig. 2B) is the same as the passive response. In contrast, the sub-BF 80-dB RL response (Fig. 2D) is above the postmortem response. This is consistent with the postmortem experimental results in [7].

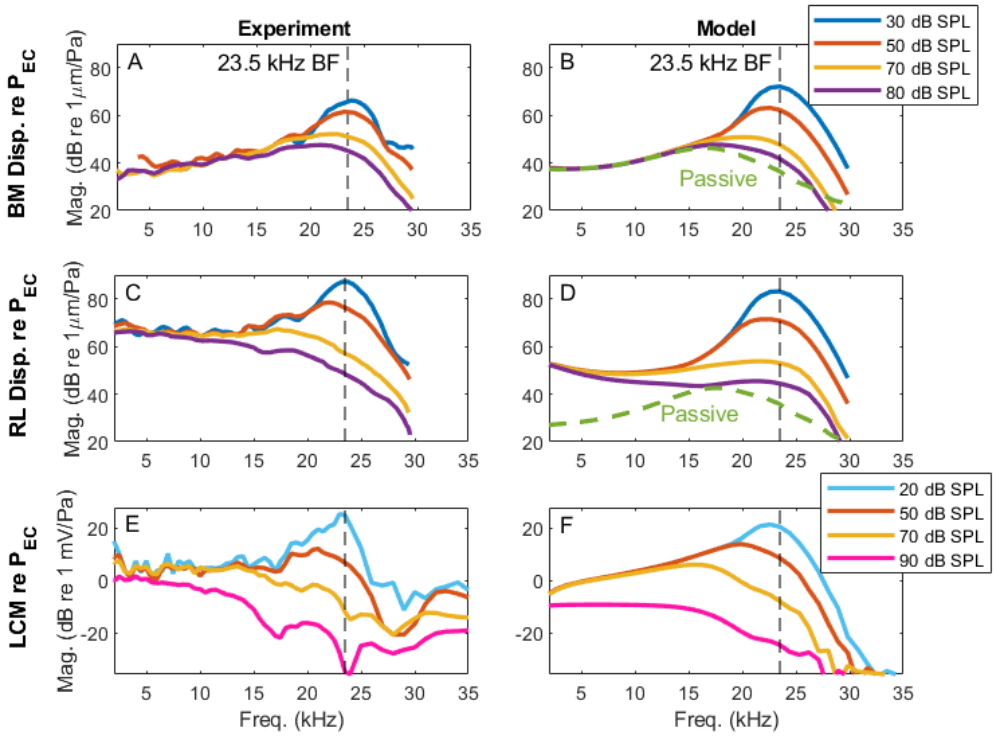


FIGURE 2. Experimental data (left) for RL and BM at 26 kHz BP (from He et al. [7], shifted to 23.5 kHz using scaling symmetry) and LCM at 23.5 kHz (from Bowling et al. [6]). 23.5 kHz Pure Tone model response RL, BM, and LCM displacement (right). The passive model response (green dashed line) is also shown for BM and RL.

Model predicts two-tone suppression of sub-BF probe responses of the RL but not of the BM

Using this model, a sweep of suppressor frequency and level was conducted to assess sub-BF suppression in the BM, RL and LCM responses. As shown in Figure 3 for a single level and frequency combination, when the suppressor peaks sub-BF (Fig. 3A), only RL and LCM suppression is observed (Figs. 3B-C), with the BM response remaining unchanged relative to its pure tone response (Fig. 3A). To explore this phenomenon more broadly, both the suppressor frequency and level were swept (5 - 25 kHz in 0.4 kHz steps, 30, 50, 70, 80 dB SPL). Suppression was measured in a 10 kHz, 30 dB SPL probe response, at the 23.5 kHz BP (0.23 cm from base), which is far basal to the 10 kHz peak (0.47 cm from base) and well outside the region of BM nonlinearity. To calculate suppression, the level of the probe response was subtracted from the level of the pure tone response at the 23.5 kHz BP.

Figure 4 shows the results of this sweep, and is in qualitative agreement with the Dewey et al. [5] experimental data. The BM (Fig. 4A) is essentially unsuppressed (with a maximum suppression of 0.91 dB) for all suppressor frequencies and levels. On the other hand, the RL and LCM probe responses (Fig. 4B-C) are suppressed at all suppressor frequencies for 70 and 80 dB SPL suppressors, and at suppressor frequencies above 0.5 octaves (relative to the 23.5 kHz BF) for 30 and 50 dB SPL suppressors. This is because the suppressor response is very low at the 23.5 kHz BP for suppressors below that frequency, since they would peak apically in relation to the probe. Finally, peak suppression in the RL and LCM occurs at a suppressor frequency close to the BF of the measurement location (23.5 kHz BP) for the lowest suppressor level (30 dB SPL), and shifts to lower frequency with increasing suppressor level. This trend is consistent with the experimental findings in [5] and is due to the fact that, as suppressor level is increased, suppressors of lower frequency peak at the 23.5 kHz BP (where suppression is measured).

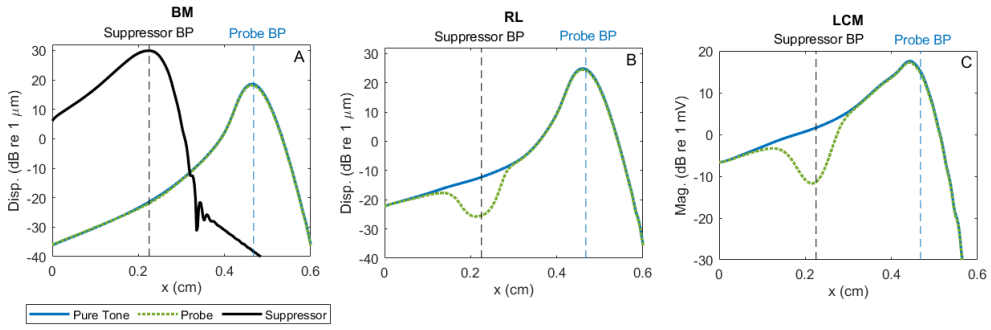


FIGURE 3. 30 dB SPL, 10 kHz pure tone BM, RL, and LCM probe responses (blue) are shown compared to 30 dB SPL, 10 kHz probe responses (green dashed) in the presence of a 70 dB SPL, 23.5 kHz suppressor (shown in black for BM). Suppression of the RL and LCM response observed near the peak of the suppressor response.

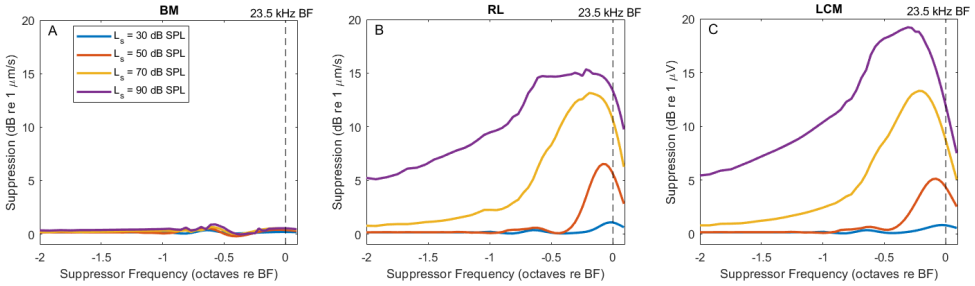


FIGURE 4. Suppression for BM, RL, and LCM displacement (left, middle, right, respectively) shown for a sweep of suppressor frequency and suppressor level. Suppression measured at the 23.5 kHz BP for a probe of 10 kHz.

TTS of RL is due to saturation of the OHC force and not due to a change in the phase of the force relative to RL velocity

In Figure 5, the magnitude and phase of the electromotile force applied by the OHCs on the BM and RL is shown for pure tone and two-tone for the same case as shown in Figure 3. A 10 kHz, 30 dB SPL probe is suppressed by a 23.5 kHz, 70 dB SPL suppressor, and the forces at the probe frequency are shown compared to the pure tone forces. For the phase, green and red zones indicate regions of power addition and removal respectively, to the BM and RL by the OHCs, allowing for interpretation of where power is added to the BM traveling wave and where it is removed, and how that changes in the presence of a suppressor.

The presence of the suppressor reduces the amplitude of the force on both the BM and RL at the 23.5 kHz BP (Fig. 5A, 5B). However, this only suppresses the response of the RL and not the BM as seen along the vertical dashed lines in Figure 3. This indicates that while the force applied by the OHCs on the BM and RL changes comparably due to the presence of a suppressor, this changing force magnitude can only affect the RL response. Regarding the phase of the force, the model predicts that in response to a pure tone, the OHCs deliver power to the BM in a region located mostly basal to the peak (Fig. 5C), in line with prior computational work [8]. For the RL, the OHCs deliver power at at all locations (Fig. 5D). A change in the phase of the force of the BM relative to its velocity is also observed near the region of the suppressor peak, and this slightly shifts the onset of the region where OHCs deliver power to the BM towards the apex, compared to the pure tone. Neither the change in the magnitude nor the phase of the force applied on the BM relative to its velocity is able to impact the BM response. On the other hand, the suppressed force amplitude on the RL at the 23.5 kHz BP, which is similar in reduction to that of the force amplitude on the BM, is able to suppress the RL response. The phase of the OHC force applied on the RL relative to the RL velocity remains unchanged compared to the pure tone response, so any phase effect as a driver of suppression can be precluded. This analysis indicates that the pure tone force at the 23.5 kHz best place is able to influence the RL's, but not the BM's

response. Hence, the change in the force amplitude due to a suppressor, while comparable between the BM and RL, is only reflected in the RL response, while the BM remains unaffected.

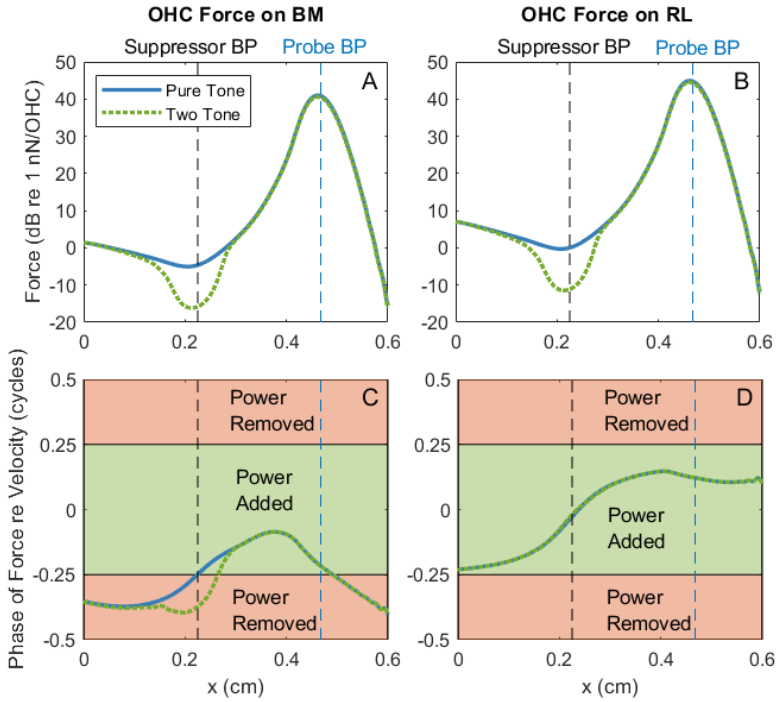


FIGURE 5. Electromotile force applied by the OHCs on BM (A) and RL (B) for 10 kHz pure tone and 10 kHz probe response in presence of 23.5 kHz suppressor. Phase of force relative to velocity of BM (C) and RL (D). Green and pink zones indicate zones of power addition and removal respectively.

DISCUSSION

This work utilizes a physiologically-based model of the gerbil cochlea to make predictions about two-tone suppression in the BM, RL, and LCM in response to a variety of suppressor frequencies and levels. These predictions agree qualitatively with recent experiments from Dewey et al. [5]. Through analysis of the OHC electromotile force on the BM and RL for a single suppression case, it is demonstrated that the suppressed OHC force is only able to affect the RL. The mechanism for suppression is not a change in the phase of the force relative to the velocity, but rather a change in the amplitude which is due to OHC saturation caused by the presence of a suppressor.

Other phenomena that arise from cochlear nonlinearity show discrepancies in the effects on the BM and RL as well. For instance, the pure tone experimental data shown from He et. al [7] (shown in Fig. 2) show broader sub-BF nonlinearity for the RL compared to the BM. In addition, the experimental results in [9] show intracochlear distortion products (IDPs) that are higher in magnitude and extend over a broader frequency range for the RL compared to the BM. This ties with the computational findings reported by Bowling et al. [6]. In that work, it is demonstrated that intracochlear distortion products are broadly generated, but the distortion-product otoacoustic emissions (DPOAEs) that they result in are spatially restricted. More specifically, the computational analysis in that work demonstrated that eliminating OHC activity basal to the distortion-product best place does not affect the DPOAE response, but eliminating it near the distortion-product best place reduces it. This is because the DPOAE reverse traveling wave is carried by the BM, and it is unaffected by the elimination of OHC activity in basal regions. The OHC force magnitude in the region of eliminated OHC activity is reduced in a manner similar to Fig. 5 in this work, but that change is not able to affect the BM's response and the DPOAE reverse traveling wave travels to the stapes unimpeded. The

supplementary information in [6] further shows that in basal regions, the OHC force on the tectorial membrane (TM, similar impedance to the RL) is of comparable magnitude to the contribution of the TM impedance, but that the OHC force on the BM is several orders of magnitude lower than the contribution of the BM impedance and the force of the fluid on the BM.

Along similar lines, experimental results to pure tone stimuli in [10, 11] indicate that in response to a pure tone stimulus in the mouse cochlea, there is a prominent DC shift of the TM, RL, and bottom end of the OHCs in addition to the cyclical motion of those structures. The DC effect on the BM, however, is much less pronounced. This suggests that the BM is relatively unaffected by the tonic contraction of the OHCs, which is indicated by the DC shift in the other structures.

The work by Moleti and Sisto [2] uses a 1-D transmission line model with two degrees of freedom (to represent the BM and RL) and nonlinear functions directly representing OHC feedback to make predictions about two-tone suppression. The traveling wave power at the peak region is found to be minimally affected by high-frequency suppressors, even though the RL degree of freedom (DOF) is suppressed at those high frequencies (corresponding to basal suppression in this work). The minimal complexity of their phenomenological model is advantageous in connecting predictions directly to model parameters. The authors claim that the traveling wave kinetic energy is dependent on the mass of both DOFs, with the more massive BM DOF dominating. As such, any suppression of the lighter RL DOF minimally affects the total kinetic energy of the traveling wave. The 2-D representation of OoC micromechanics in this work complements that model by providing greater insight on the motion of OoC structures in two directions (radial, transverse), MET channel nonlinearity to characterize OHC feedback, and the ability to directly analyze the OHC force on the surrounding structures. This allows extensive model validation via prediction of BM, RL and LCM responses. The prediction of suppression in the LCM, as shown earlier in Fig. 4C, is valuable because the LCM is an experimentally obtainable indicator of OHC feedback. The results of both works indicate that the RL responds to OHC feedback in basal regions while the BM does not, and they generally predict the experimental predictions in [5].

In summary, a growing set of experimental data strongly points to a discrepancy in the effect of OHC-driven nonlinear phenomena on the BM and other OoC structures at the top of the OHCs. While OHC feedback is only able to affect the BM in the peak region, it affects other structures (such as the RL and TM) over a much broader region. Computational efforts explore this further, providing analysis of power and force on various OoC structures. The implications for cochlear mechanics are that the difference in the impedance of various OoC structures plays a key role in ensuring that the BM traveling wave is unimpeded in reaching its BP. Even in the presence of a basal disruption such as a suppressor, the BM traveling wave is unaffected. Future modeling work might explore whether this large difference in impedance is a prerequisite for the characteristic nonlinearity in the BM response that is associated with 'normal' cochlear function.

ACKNOWLEDGMENTS

This research is funded by NIH grant R01 DC016114.

REFERENCES

1. M. A. Ruggero, L. Robles, and N. C. Rich, "Two-tone suppression in the basilar membrane of the cochlea: mechanical basis of auditory-nerve rate suppression," Journal of neurophysiology **68**, 1087–1099 (1992).
2. A. Moleti and R. Sisto, "Suppression tuning curves in a two-degrees-of-freedom nonlinear cochlear model," The Journal of the Acoustical Society of America **148**, EL8–EL13 (2020).
3. W. Zhou and J.-H. Nam, "Probing hair cell's mechano-transduction using two-tone suppression measurements," Scientific reports **9**, 1–16 (2019).
4. J. Meaud and K. Grosh, "Effect of the attachment of the tectorial membrane on cochlear micromechanics and two-tone suppression," Biophysical journal **106**, 1398–1405 (2014).
5. J. B. Dewey, B. E. Applegate, and J. S. Oghalai, "Amplification and suppression of traveling waves along the mouse organ of corti: evidence for spatial variation in the longitudinal coupling of outer hair cell-generated forces," Journal of Neuroscience **39**, 1805–1816 (2019).
6. T. Bowling, H. Wen, S. W. F. Meenderink, W. Dong, and J. Meaud, "Intracochlear distortion products are broadly generated by outer hair cells but their contributions to otoacoustic emissions are spatially restricted," Scientific Reports **11**, 13651 (2021).
7. W. He, D. Kemp, and T. Ren, "Timing of the reticular lamina and basilar membrane vibration in living gerbil cochleae," eLife **7**, e37625 (2018).
8. E. de Boer, "The "inverse problem"solved for a three-dimensional model of the cochlea. i. analysis," The Journal of the Acoustical Society of America **98**, 896–903 (1995).

9. T. Ren and W. He, "Two-tone distortion in reticular lamina vibration of the living cochlea," Communications biology **3**, 1–8 (2020).
10. W. He and T. Ren, "The origin of mechanical harmonic distortion within the organ of corti in living gerbil cochleae," Communications biology **4**, 1–11 (2021).
11. J. B. Dewey, A. Altoè, C. A. Shera, B. E. Applegate, and J. S. Oghalai, "Cochlear outer hair cell electromotility enhances organ of corti motion on a cycle-by-cycle basis at high frequencies in vivo," Proceedings of the National Academy of Sciences **118** (2021).

Forms of longitudinal coupling in the organ of Corti

Stephen Elliott,[1, a)] Riccardo Marrocchio,[1] and Karl Grosh[2, 3]

[1)] *Institute of Sound and Vibration Research, University of Southampton, Highfield Campus, Southampton, SO17 1BJ, UK*
[2)] *Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI 48109, USA*
[3)] *Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI 48109, USA*

[a)] *Corresponding author: s.j.elliott@soton.ac.uk*

Abstract. The effect of different forms of longitudinal coupling on the active response of the cochlea are analysed using an elemental approach, based on a previous finite element model of the guinea pig cochlea. The forms of coupling in the organ of Corti include mechanical longitudinal coupling along the basilar membrane (BM) and the tectorial membrane (TM). An intuitive and efficient method is used for representing the longitudinal coupling in the fluid, by decomposing the 3D fluid coupling into a local near-field mass loading of the organ of Corti and 1D fluid coupling to represent long range fluid-structure coupling. As has been shown in previous studies, mechanical longitudinal coupling in the TM, together with 3D fluid coupling, have the greatest effect on the predicted response and are most important in obtaining a frequency response that is both tall and broad, as observed in experimental data. The 3D fluid coupling is also necessary to obtain a phase variation that is consistent with experimental measurements. The effect of BM longitudinal coupling is not as significant as that due to the TM in determining the response. The overall BM admittance can be readily calculated using the elemental method, and since this is the only aspect of the organ of Corti dynamics that couples into the fluid, it is a useful indicator of its overall behaviour. Whereas in the locally reacting case, the BM velocity at a given position depends only on the pressure difference at the same point, when longitudinal coupling is introduced, the BM velocity also depends on the pressure difference some distance away. For excitation at a single point and at a given frequency, the BM admittance thus has a distribution along the cochlea, with a form that depends on excitation frequency and on the assumed types of longitudinal coupling. It is also shown that the broadening of the cochlear response due to longitudinal coupling is accompanied by an extension of the frequency range over which power is supplied by the BM into the fluid.

INTRODUCTION

Different approaches can be taken to modelling the mechanics of the cochlea. One approach considers the detailed physical processes involved, using a realistic geometry, in order to understand how the various elements combine together to give the observed behaviour. In another approach, the details of these processes are disregarded, and a simplified, but potentially more insightful, description of the behaviour is sought, typically involving wave propagation. This paper is the first of two which try to make a connection between these two approaches by reformulating a previous detailed model of the physical processes within the organ of Corti, OC, to provide such a simplified interpretation. In this first paper, the model is briefly described, initially emphasising the individual forms of longitudinal coupling in the OC. The fluid coupling is then divided into a 1D component and a component due to the near field fluid pressure, with the latter being associated with a modified OC model. This can then be combined with the 1D fluid coupling to derive both the overall response and the properties of the waves that could propagate in such a model. There are many mechanisms for longitudinal coupling in the OC. Table I lists the approximate lengthscale of several of these mechanisms when analysed in isolation. It should be emphasised, however, that the importance of each of these mechanisms depends on the way in which they react together within the OC. The bending response of a sandwich depends in a rather complicated way on the properties of the individual layers (Richardson et al., 2008). The coupling along the tectorial membrane, TM, has been singled out as being particularly important in determining the overall response (Meaud and Grosh, 2010; Lukashkin et al., 2010).

TABLE I. The lengthscale of the individual mechanisms of longitudinal coupling in the OC at CF, in isolation.

Element	Mechanism	Lengthscale	Quantity	Reference
BM	Orthotropic plate	$20\,\mu$m	wavelength	Meaud and Grosh (2010)
TM	Viscoelastic	$60\,\mu$m	wavelength	Meaud and Grosh (2010)
OHC	Electrical	$40\,\mu$m	Decay rate	Teal and Ni (2016)
Fluid	Near field acoustic	$120\,\mu$m	Decay rate	Ni and Elliott (2015)

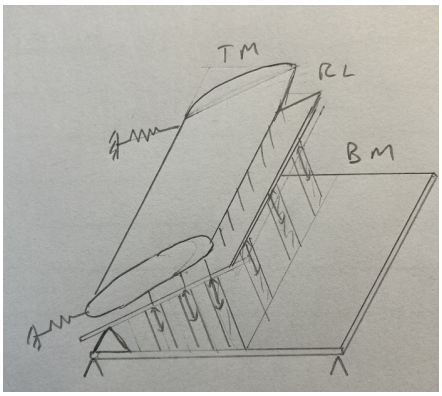


FIGURE 1. Micromechanical model of the organ of Corti.

This paper uses a micromechanical model based on that of (Ramamoorthy et al., 2007; Meaud and Grosh, 2010; Sasmal and Grosh, 2019) and is illustrated in Fig. 1 . There are three mechanical degrees of freedom, associated with the transverse motion of the basilar membrane, BM, the rotational motion of the reticular lamina, RL, and the shear motion of the tectorial membrane, TM. The BM and TM are assumed to be coupled longitudinally, with that along the BM being modelled as an orthotropic plate and that in the TM using a viscoelastic model.

MECHANICAL AND ELECTRICAL COUPLING

Although this micromechanical model was formulated using the finite element method, it is more convenient here to use a simplified elemental approach (Elliott and Ni, 2018)in which the cochlea is divided into N sections, and the distributions of complex BM velocity $v(x)$ and pressure difference $p(x)$, at a given frequency, can be represented by two vectors:

$$\mathbf{p} = [p(1), p(2), ..., p(N)]^T, \tag{1a}$$

$$\mathbf{v} = [v(1), v(2), ..., v(N)]^T. \tag{1b}$$

The fluid coupling in the chambers is then described by the matrix equation:

$$\mathbf{p} = \mathbf{Z_{fc}}\mathbf{v}, \tag{2}$$

where $\mathbf{Z_{fc}}$ is the fluid coupling matrix. The micromechanical response of the BM is given by:

$$\mathbf{v} = \mathbf{v_s} - \mathbf{Y_{bm}}\mathbf{p}, \tag{3}$$

where $\mathbf{v_s}$ is the excitation vector due to the stapes, whose motion drives the system and $\mathbf{Y_{bm}}$ is the BM admittance matrix, which is no longer diagonal if the OC is longitudinally coupled. When both the BM admittance and fluid coupling are considered, the overall response of the BM velocity is thus:

$$\mathbf{v} = [\mathbf{I} + \mathbf{Y_{bm}}\mathbf{Z_{fc}}]^{-1}\mathbf{v_s}. \tag{4}$$

So, having calculated the elements of the matrices $\mathbf{Y_{bm}}$ and $\mathbf{Z_{fc}}$, the distribution of the BM velocity and the pressure difference along the cochlea can be readily calculated.

FLUID COUPLING

The fluid coupling in the 3D cochlea can be described using the analytic wavenumber model of (Steele and Taber, 1979), in which:

$$Z_{fc}(k) = i\omega\rho Q(k). \tag{5}$$

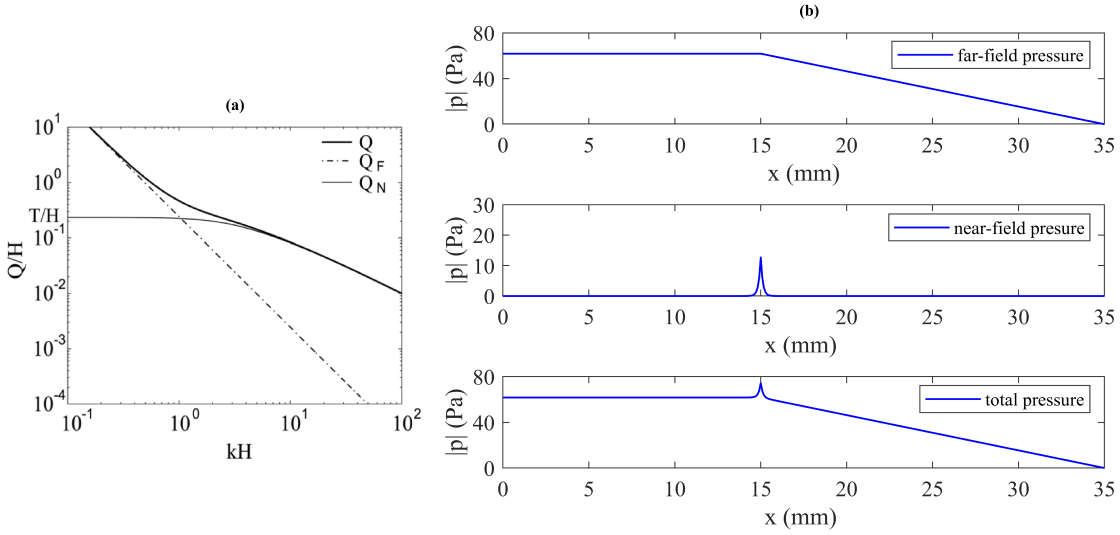


FIGURE 2. The separation of the 3D fluid coupling into 1D and near field pressure components, in (a) the wavenumber domain and (b) the spatial domain. Fig. (a) is reprinted with permission from (Elliott et al., 2011). Copyright 2011, Acoustic Society of America.

It is convenient to divide the pressure distributions due to the 3D fluid coupling into two components (Elliott et al., 2011), one of which is that due to 1D fluid coupling, which has a simple analytical form, and the other due to the near field pressure distribution, as shown in Fig 2. The longitudinal variation of the near field pressure is only locally distributed around the excitation position and is well approximated by an exponentially decaying distribution (Ni and Elliott, 2015), the decay length of which is shown in Table I. The fluid coupling impedance matrix can thus be expressed as:

$$\mathbf{Z_{fc}} = \mathbf{Z_{1D}} + \mathbf{Z_{nf}}. \tag{6}$$

The coupling term in Eq. (4) can now be written as:

$$\mathbf{Y_{bm}Z_{fc}} = \mathbf{Y'_{bm}Z_{1D}}, \tag{7}$$

where

$$\mathbf{Y'_{bm}} = \mathbf{Y_{bm}}[\mathbf{I} + \mathbf{Z_{nf}Z_{1D}^{-1}}]. \tag{8}$$

The near field fluid coupling can thus be incorporated into a modified BM admittance that interacts only with the 1D fluid coupling to give the overall response of the cochlea. This approach may be viewed as a generalisation of the way that (Neely, 1985) incorporated an additional mass in the micromechanical model to account for the fluid loading on the BM. It has the advantage that the peak in the modified BM admittance curve is close to the characteristic frequency, CF.

OVERALL RESPONSE

The BM admittance matrix can be derived from the micromechanical model described above using finite difference methods (Marrocchio, 2022), including the longitudinal coupling in each of the elements. By taking individual columns of this admittance matrix, the frequency response of the BM admittance can be calculated not only for the velocity at the position that the pressure is applied, x_0, but also at positions some distance, x', from this excitation position, and the spatially distributed BM admittance can be written as:

$$Y_{bm}(\omega, x_0, x_0 + x') = \frac{v(x_0 + x')}{p(x_0)}. \tag{9}$$

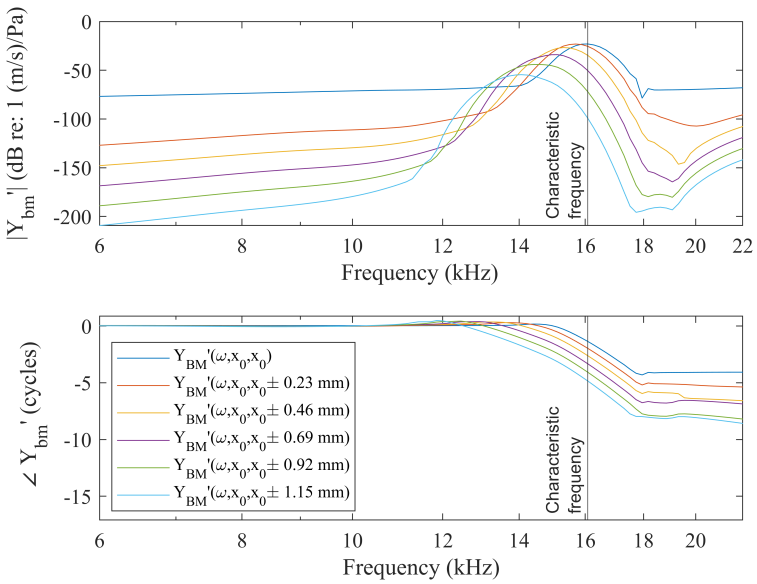


FIGURE 3. The frequency variation of the admittance corresponding to an excitation due to a pressure difference applied at the position x_0 only and the BM velocity measured at various positions around x_0, i.e. $x = x_0 + x'$.

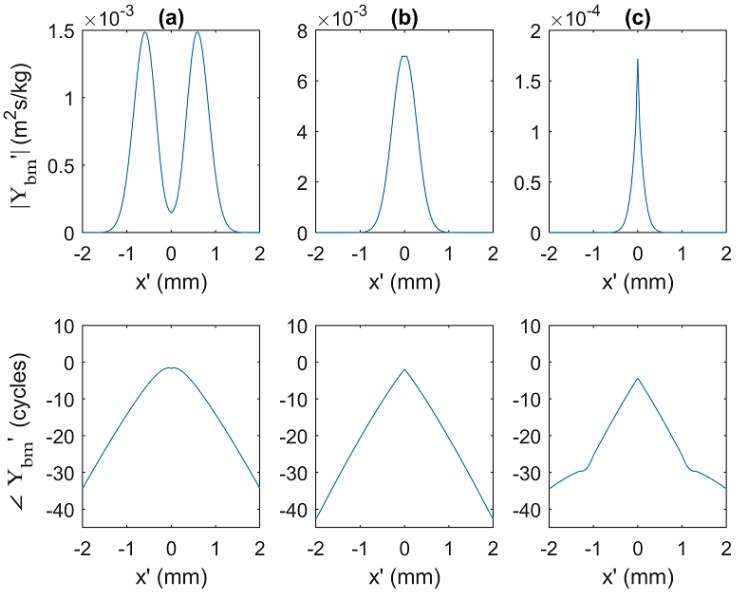


FIGURE 4. The magnitude (top) and phase (bottom) of the distribution of BM admittance along the cochlea at (a) 0.9 CF, (b) CF and (c) 1.1 CF, derived from the response shown in Fig. 3.

This is plotted as a function of the frequency in Fig. 3 for different values of x' and with parameters representative of 4 mm along the guinea pig cochlea, as detailed in (Sasmal and Grosh, 2019), with 3D fluid coupling and mechanical longitudinl coupling in the TM only. Fig. 4 shows an alternative way of displaying the BM admittance; in terms of the magnitude and phase of its spatial distribution around x', plotted for three discrete frequencies. Since all the forms of longitudinal coupling considered here are the same whether x' is positive or negative, these plots are symmetrical about $x' = 0$. It should be noted that to illustrate their shape, the vertical scales are different for the results at different frequencies in 3, with the magnitude of the admittance being much smaller at 1.1CF than at CF, for example.

Fig. 5 shows the frequency response of the BM velocity in the coupled cochlea, calculated from the elemental model, in the case above and also with only a locally-reacting OC and 1D fluid coupling. It can be seen that although the locally-reacting model does a reasonable job of predicting the coupled BM responses, these are not as tall or broad as the responses predicted with TM coupling and 3D fluid coupling. This observation is quantified in Table II, in which the amplification of the active cochlear model at CF, compared with the passive case, is shown. It is clear that somewhat greater amplification can be achieved with TM longitudinal coupling compared with only the locally reacting case and that the amplification is only slightly increased if longitudinal coupling in both the BM and TM are included. The width of the coupled response is quantified by the $Q_{10\,\mathrm{dB}}$ values also shown in Table II. This value is relatively large, indicating a sharp peak, in the locally reacting case with 1D fluid coupling, but decreases significantly when longitudinal coupling is included, particularly when 3D fluid coupling is combined with longitudinal coupling along the TM, as noted by (Meaud and Grosh, 2010). The amplification and $Q_{10\,\mathrm{dB}}$ of the model with 3D fluid coupling and both BM and TM longitudinal coupling are similar to those measured experimentally by (Zheng et al., 2007).

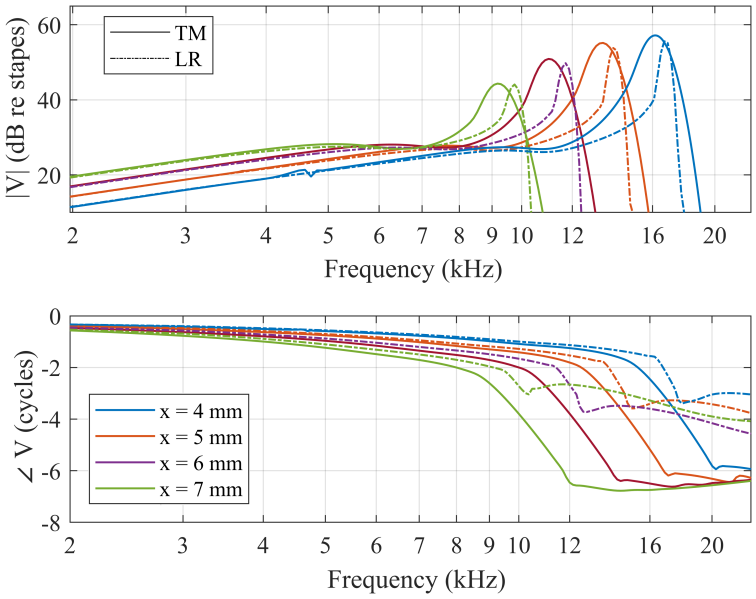


FIGURE 5. The BM velocity predicted by the coupled elemental model of the cochlea with longitudinal coupling in the TM and 3D fluid coupling, TM, and with only a locally reacting organ of Corti and 1D fluid coupling, LR.

TABLE II. The characteristic frequency, CF, amplification of the active cochlea over the passive case, in dB, and the $Q_{10\,\mathrm{dB}}$ measure of the sharpness of the coupled response, in the model when it is locally reacting, with 1D or 3D fluid coupling, and when there is only mechanical longitudinal coupling in either the BM or the TM alone or both TM and BM, with 3D fluid coupling. Also shown are the results from the measurements reported by (Zheng et al., 2007).

Type of longitudinal coupling	CF	Amplification	$Q_{10\,\mathrm{dB}}$
Locally reacting	17.0 kHz	27 dB	8.2
BM only	16.6 kHz	34 dB	8.3
TM only	16.2 kHz	32 dB	6.2
BM and TM	16.1 kHz	35 dB	6.8
BM and TM	EL 16.1 kHz	35 dB	6.8
Measurements	16.5 kHz	35 dB	6.5

POWER TRANSFER

In a locally reacting model of the BM admittance, the sign of the real part of admittance determines whether power at any one location is being transferred from the fluid to the BM or vice versa. This is more difficult to determine from the BM admittance alone in the non locally-reacting model described above, but can still be calculated by solving the coupled problem and using the resulting BM motion, as in Fig. 5, to calculate the fluid pressure and hence the real part of the product of this with the BM velocity to give the power transferred from the BM to the fluid. Fig. 6 shows the power transferred from the BM to the fluid as a function of frequency from the simulations above with TM longitudinal coupling and 3D fluid coupling and with a locally reacting BM and 1D fluid coupling. It is clear that the frequency range over which the travelling wave is amplified by the power transfer from the BM to the fluid is significantly greater in the longitudinally coupled case than in the locally reacting case.

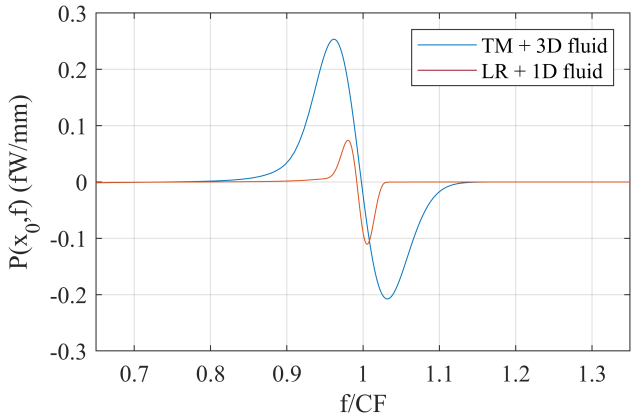


FIGURE 6. Power transfer per unit length from the BM to the fluid as a function of frequency, f, normalized with respect to the characteristic frequency, CF, calculated from the elemental model of the cochlea with longitudinal coupling in the TM and 3D fluid coupling (in blue) and without longitudinal coupling in the OC and 1D fluid coupling (in red).

CONCLUSIONS

The main result of this paper is the derivation of the spatial distribution of the BM admittance in the active cochlea with longitudinal coupling, when it is forced at a single point, as shown in Fig. 4. This is be used in an elemental model to determine the coupled response of the cochlea, with a lower computational cost with respect to a detailed finite element model. It is instructive to compare the BM admittance distribution, which extends over about 1 mm at CF, with the spatial extent of the various individual forms of longitudinal coupling in isolation, in Table I, which generally extend over much shorter distances. This shows that it is not straightforward to predict the dynamics of the combined cochlear system from the characteristics of these individual components. The power transfer from the BM to the fluid is also derived, showing that the frequency range over which power is transferred to the travelling wave depends on the type of longitudinal coupling. The near field component of the fluid coupling is included in a modified BM admittance and thus associated with the OC dynamics, rather than those of the fluid coupling. The residual fluid coupling is then only 1D. This allows a simple formulation of the dispersion equation, from which the wavenumbers can be derived for the waves that can propagate in the cochlear model, using the Fourier transform of the spatial distribution of the BM admittance (Marrocchio, 2022).

ACKNOWLEDGMENTS

This work is supported by the DigiTwin project (EP/R006768/1).

REFERENCES

Stephen J. Elliott and Guangjian Ni. An elemental approach to modelling the mechanics of the cochlea. *Hearing Research*, 360:14–24, mar 2018. ISSN 03785955. doi:10.1016/j.heares.2017.10.013. URL https://linkinghub.elsevier.com/retrieve/pii/S0378595517303465.

Stephen J. Elliott, Ben Lineton, and Guangjian Ni. Fluid coupling in a discrete model of cochlear mechanics. *The Journal of the Acoustical Society of America*, 130(3):1441–1451, sep 2011. ISSN 0001-4966. doi:10.1121/1.3607420. URL http://asa.scitation.org/doi/10.1121/1.3607420.

Andrei N. Lukashkin, Guy P. Richardson, and Ian J. Russell. Multiple roles for the tectorial membrane in the active cochlea. *Hearing Research*, 266 (1-2):26–35, jul 2010. ISSN 03785955. doi:10.1016/j.heares.2009.10.005. URL https://linkinghub.elsevier.com/retrieve/pii/S0378595509002500.

Riccardo Marrocchio. *Understanding the micromechanics of the cochlea*. PhD thesis, University of Southampton, 2022.

Julien Meaud and Karl Grosh. The effect of tectorial membrane and basilar membrane longitudinal coupling in cochlear mechanics. *The Journal of the Acoustical Society of America*, 127(3):1411–1421, mar 2010. ISSN 0001-4966. doi:10.1121/1.3290995. URL http://asa.scitation.org/doi/10.1121/1.3290995.

Stephen T. Neely. Mathematical modeling of cochlear mechanics. *The Journal of the Acoustical Society of America*, 78(1):345–352, jul 1985. ISSN 0001-4966. doi:10.1121/1.392497. URL http://asa.scitation.org/doi/10.1121/1.392497.

Guangjian Ni and Stephen J. Elliott. Comparing methods of modeling near field fluid coupling in the cochlea. *The Journal of the Acoustical Society of America*, 137(3):1309–1317, mar 2015. ISSN 0001-4966. doi:10.1121/1.4908242. URL http://asa.scitation.org/doi/10.1121/1.4908242.

Sripriya Ramamoorthy, Niranjan V. Deo, and Karl Grosh. A mechano-electro-acoustical model for the cochlea: Response to acoustic stimuli. *The Journal of the Acoustical Society of America*, 121(5):2758–2773, may 2007. ISSN 0001-4966. doi:10.1121/1.2713725. URL http://asa.scitation.org/doi/10.1121/1.2713725.

Guy P. Richardson, Andrei N. Lukashkin, and Ian J. Russell. The tectorial membrane: one slice of a complex cochlear sandwich. *Current Opinion in Otolaryngology & Head & Neck Surgery*, 16(5):458–464, oct 2008. ISSN 1068-9508. doi:10.1097/MOO.0b013e32830e20c4. URL https://journals.lww.com/00020840-200810000-00012.

Aritra Sasmal and Karl Grosh. Unified cochlear model for low- and high-frequency mammalian hearing. *Proceedings of the National Academy of Sciences*, 116(28):13983–13988, jul 2019. ISSN 0027-8424. doi:10.1073/pnas.1900695116. URL https://pnas.org/doi/full/10.1073/pnas.1900695116.

Charles R. Steele and Larry A. Taber. Comparison of WKB calculations and experimental results for three-dimensional cochlear models. *The Journal of the Acoustical Society of America*, 65(4):1007–1018, apr 1979. ISSN 0001-4966. doi:10.1121/1.382570. URL http://asa.scitation.org/doi/10.1121/1.382570.

Paul D. Teal and Guangjian Ni. Finite element modelling of cochlear electrical coupling. *The Journal of the Acoustical Society of America*, 140 (4):2769–2779, oct 2016. ISSN 0001-4966. doi:10.1121/1.4964897. URL http://asa.scitation.org/doi/10.1121/1.4964897.

Jiefu Zheng, Niranjan Deo, Yuan Zou, Karl Grosh, and Alfred L. Nuttall. Chlorpromazine Alters Cochlear Mechanics and Amplification: In Vivo Evidence for a Role of Stiffness Modulation in the Organ of Corti. *Journal of Neurophysiology*, 97(2):994–1004, feb 2007. ISSN 0022-3077. doi:10.1152/jn.00774.2006. URL https://www.physiology.org/doi/10.1152/jn.00774.2006.

Wave motion in the longitudinally coupled cochlea

Riccardo Marrocchio,[1, a)] Karl Grosh,[2, 3, b)] and Stephen Elliott[1, c)]

[1)] *Institute of Sound and Vibration Research, University of Southampton, Highfield Campus, Southampton, SO17 1BJ, UK*

[2)] *Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI 48109, USA*

[3)] *Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI 48109, USA*

[a)] *Corresponding author: Riccardo.Marrocchio@soton.ac.uk*

[b)] *grosh@umich.edu*

[c)] *s.j.elliott@soton.ac.uk*

Abstract. A detailed model of the physical processes occurring within the organ of Corti can be compared with experimental data and can give an indication of the mechanisms of its active behaviour. On the other hand, a model that involves wave propagation in the cochlea can give more insight into its coupled response. In principle many kinds of wave can propagate in the cochlea, but it is a single travelling wave that mainly determines its coupled response, which can be characterized by the frequency variation of its complex wavenumber. In the case of a cochlear model using a locally reacting basilar membrane (BM) and 1D fluid coupling, this wavenumber can be calculated explicitly. When longitudinal coupling, due to the mechano-electrical structure of the Organ of Corti or other forms of fluid coupling, are introduced however, additional wave types are possible, and it is not so straightforward to calculate the wavenumber of the main travelling wave. This paper is the second of a series, in which a method is presented of deriving the wavenumber distribution associated to different wave types, based on an elemental model of the cochlea. This allows an investigation of the effect of different forms of longitudinal coupling on the wave motion. In general, the main travelling wave dominates the BM response, and when the model is active, the imaginary party of its wavenumber is positive in a frequency region just before the characteristic frequency. The extent of this active region depends on the form of longitudinal coupling assumed in the organ of Corti.

INTRODUCTION

It is usually assumed that the properties of the cochlea can be explained in terms of the propagation of a single travelling wave, which is generated by the interaction between the inertia of the fluid in the chambers and the dynamics of the basilar membrane (De Boer, 1996). The properties of this wave are characterized by a complex wavenumber, whose real part, called the propagation function, determines the wave speed and wavelength while the imaginary part, called the gain function, determines whether the wave is amplified or attenuated by the active processes in the Organ of Corti (Shera, 2007). In the case of a model of the passive cochlea which includes a locally reacting basilar membrane (BM) with 1D fluid coupling, only one wave type is present (De Boer, 1996). This is characterized by a propagation function that starts with a small value at low frequencies, increases up to the characteristic frequency (CF), and then decreases at higher frequency. The rising value of the propagation function indicates that the phase and group velocities are decreasing, so that the wave is slowed down as it propagates aling the cochlea and so its amplitude is concentrated near the CF. The gain function, on the other hand, remains small at low frequencies and increases only just before the CF, reaching a negative and high value at higher frequencies, where the wave is greatly attenuated. When 2D or 3D fluid coupling is included, the models predict an infinite number of wavenumber distributions corresponding to multiple wave types (Steele and Taber, 1979; Taber and Steele, 1981).In this case, only one wavenumber distribution corresponds to a travelling wave, while the others describe non propagation modes, characterized by a large value of the imaginary part at the lowest frequencies, which do not have a significant effect on the coupled response. Using an inverse method to derive the wavenumber distribution from experimental measurements, in (Shera, 2007) it is shown that in the active cochlea, the propagation and gain functions are similar in shape to those of the travelling wave in the passive case, apart from two main differences. The propagation function has a higher peak, indicating a smaller wavelength and the gain function is positive just before the peak of the real part. Shera also noted that a positive value of the imaginary part corresponds to a power transfer to the travelling wave and thus to an amplification of the response. This is in agreement with (Dewey et al., 2019; Dong and Olson, 2013) in which it is shown, from experimental measurements, that the amplification of the cochlear wave occurs in a frequency region just below the CF. In this paper, a method is presented of deriving the wavenumber distribution associated to different types of waves that propagate in the cochlea.

WAVENUMBER DERIVATION

In the wavenumber domain at a given position along the cochlea, the modal pressure $P(k)$, the modal BM velocity, $V(k)$, and the fluid coupling impedance, $Z_{fc}(k)$, are related according to the equation (Taber and Steele, 1981; Elliott et al., 2011):

$$P(k) = Z_{fc}(k)V(k), \tag{1}$$

where k is the complex wavenumber, which is equal to:

$$k = \kappa + i\gamma, \tag{2}$$

where κ is the propagation function, equal to $2\pi/\lambda$, and γ is the gain function. Similarly, the BM response can be written in the wavenumber domain as:

$$V(k) = -Y_{BM}(k)P(k), \tag{3}$$

where $Y_{BM}(k)$ is the BM admittance transformed into the wavenumber domain. Equations (1) and (3) are simultaneously satisfied if a value of k exists for which:

$$Z_{fc}(k) = -\frac{1}{Y_{BM}(k)}, \tag{4}$$

which is the condition for propagation, termed the dispersion equation. In the case of 1D fluid coupling, the fluid coupling impedance can be written as (Ni and Elliott, 2015):

$$Z_{fc}(k) = Z_{1D}(k) = \frac{2i\omega\rho}{k^2 h}, \tag{5}$$

and if the BM admittance is only locally reacting then $Y_{BM}(k)$ is independent of k and can be written simply as Y_{BM}. So Eq. (4) is satisfied if:

$$k = \pm\sqrt{\frac{-2i\omega\rho}{h}Y_{BM}}, \tag{6}$$

which, given Y_{BM}, can be solved for each ω to determine the real and imaginary parts of k as a function of the angular frequency, at a particular position along the cochlea. More generally, however, Eq. (4) is true for 3D fluid coupling and a longitudinally coupled Organ of Corti. The fluid coupling impedance in this case can be written as (Elliott et al., 2011):

$$Z_{fc}(k) = Z_{1D} + Z_{nf}, \tag{7}$$

where $Z_{1D}(k)$ has the simple analytic form of Eq. (5), and $Z_{nf}(k)$ is due to the near field component of the pressure. Substituting into Eq. (4) the dispersion equation can be written as:

$$Z_{1D}(k) = -\frac{1}{Y'_{BM}(k)}, \tag{8}$$

where $Y'_{BM}(k)$ is a modified BM impedance that takes into account both the longitudinal coupling within the OoC and that due to the near field component of the pressure and can be written as:

$$Y'_{BM}(k) = \left(\frac{1}{Y_{BM}(k) + Z_{nf}(k)}\right)^{-1}, \tag{9}$$

where $Y_{BM}(k)$ now depends on k because it includes longitudinal coupling in the OoC. To obtain $Y'_{BM}(k)$ the admittance as a function of the relative position, $Y'_{BM}(x')$, is first derived using the elemental model, as discussed in (Marrocchio, 2022; Elliott et al., 2022). Then, by taking the Fourier transform, the admittance as a function of the real part of the wavenumber can be derived. $Y'_{BM}(k)$ is an even function of k in this case since $Y'_{BM}(x')$ is a complex but even function of x', since all the forms of longitudinal coupling considered here are symmetric about the point of excitation. As an

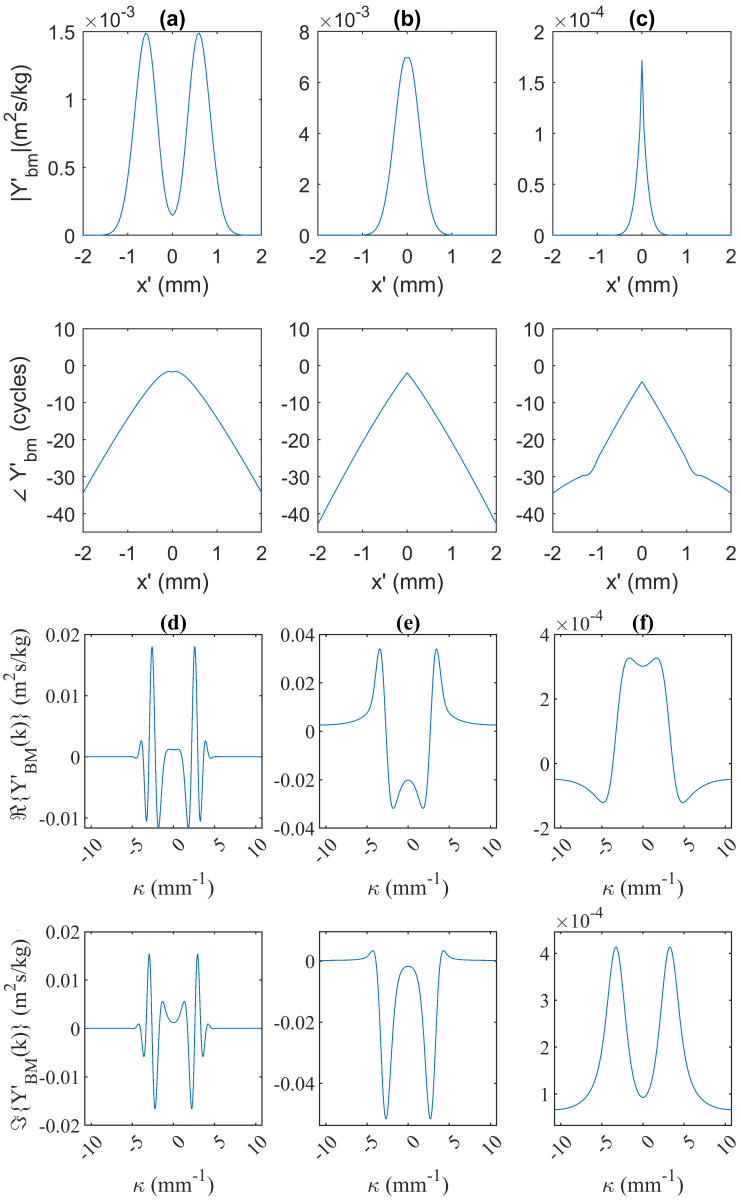


FIGURE 1. (a), (b), (c) The distribution of the modified BM admittance as a function of relative position along the cochlea, $Y'_{BM}(x')$, and (d), (e), (f) its Fourier transform for entirely real wavenumbers, $Y'_{BM}(k)$, at three different frequencies around the characteristic frequency, respectively 0.9 CF, CF and 1.1 CF.

example, the real and imaginary part of $Y'_{BM}(x')$ and $Y'_{BM}(k)$ are plotted in Fig. 1 for $\gamma = 0$ at 0.9CF, CF and CF, in the case where there is only mechanical longitudinal coupling in the tectorial membrane, TM, and 3D fluid coupling. To obtain $Y_{BM}(k)$ for different values of γ, $Y_BM(x')$ is multiplied by $exp(-\gamma|x'|)$ and then the Fourier transform is taken on the resulting function. Using Eq. (5) for $Z_{1D}(k)$ in Eq. (4), the dispersion equation can be written as:

$$1 + \frac{2i\omega\rho}{k^2 h} Y'_{BM}(k) = 0. \tag{10}$$

Both the real and the imaginary part of Eq. (4) must be zero for this equation to be satisfied. Hence, the equation can be solved geometrically by determining the intersecting points of the curves given by the real and imaginary parts of

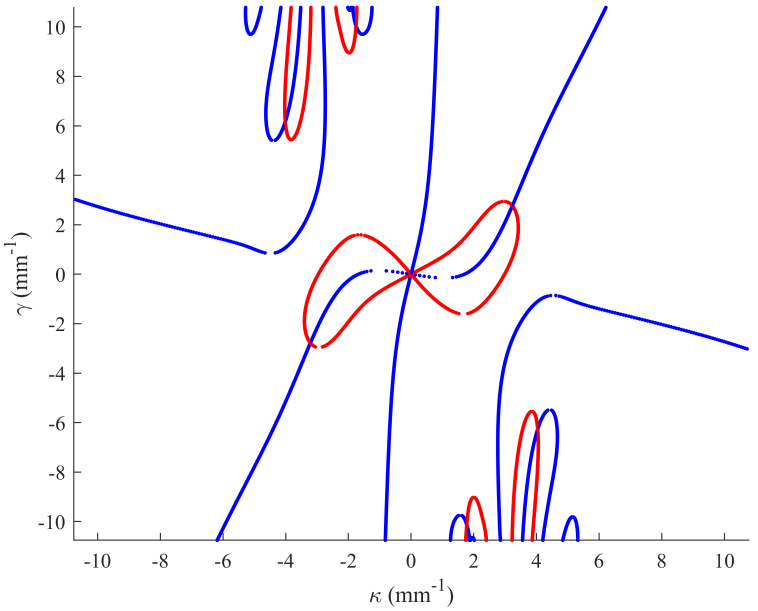


FIGURE 2. Values of the wavenumber for which the real (red) and imaginary (blue) part of Eq. (10) are zero at a frequency corresponding to $\omega = 2\pi \times 13.55\,\text{kHz}$.

this equation:

$$\begin{cases} \Re\{\frac{Y'_{BM}(k)}{k^2}\} = 0 \\ 1 - \frac{2\omega\rho}{h}\Im\{\frac{Y'_{BM}(k)}{k^2}\} = 0 \end{cases} \tag{11}$$

In Fig. 2, for example, the curves corresponding to the values for which the real and imaginary part in Eq. (11) are zero are plotted at a frequency below the CF, corresponding to $\omega = 2\pi \times 13.55\,\text{kHz}$. The intersecting points of these curves are the solutions to Eq. (10). There are multiple solutions, corresponding to different wave types. The solution with the lowest value of γ corresponds to the main travelling wave and it has a positive value of γ, indicating that the wave is being amplified at this frequency. The other solutions corresponds to waves with a large imaginary component and are thus heavily attenuated. By solving Eq. (10) for a set of values of ω, the real and imaginary part of the wavenumber can be determined as a function of the angular frequency.

WAVENUMBER DISTRIBUTION

In Fig. 3 the wavenumber distribution is plotted for the passive, locally reacting case with 1D and 3D fluid coupling and for the active case with TM longitudinal coupling. In the passive, locally reacting case with 1D fluid coupling (Fig. 3a), the real and imaginary part of the wavenumber distribution are similar to those derived in (De Boer, 1996). When 3D fluid coupling is introduced (Fig. 3b) there are multiple wavenumber distribution. The one with a low value of the propagation function at the lowest frequencies is associated to the main travelling wave, which is plotted in the figure. The others are characterized by a great, negative value of the propagation function at low frequencies so that they are only local and do not contribute significantly to the coupled response. In the active case with TM longitudinal coupling (Fig. 3c), the wavenumber distribution of the main travelling wave has a real part with a higher peak and an imaginary part which is positive just before the characteristic frequency, in accordance to the results derived by (Shera, 2007). The positive value of the imaginary part indicates that power is being transferred to the travelling wave, amplifying the response. This is in agreement with the experimental results by (Dewey et al., 2019; Dong and Olson, 2013).

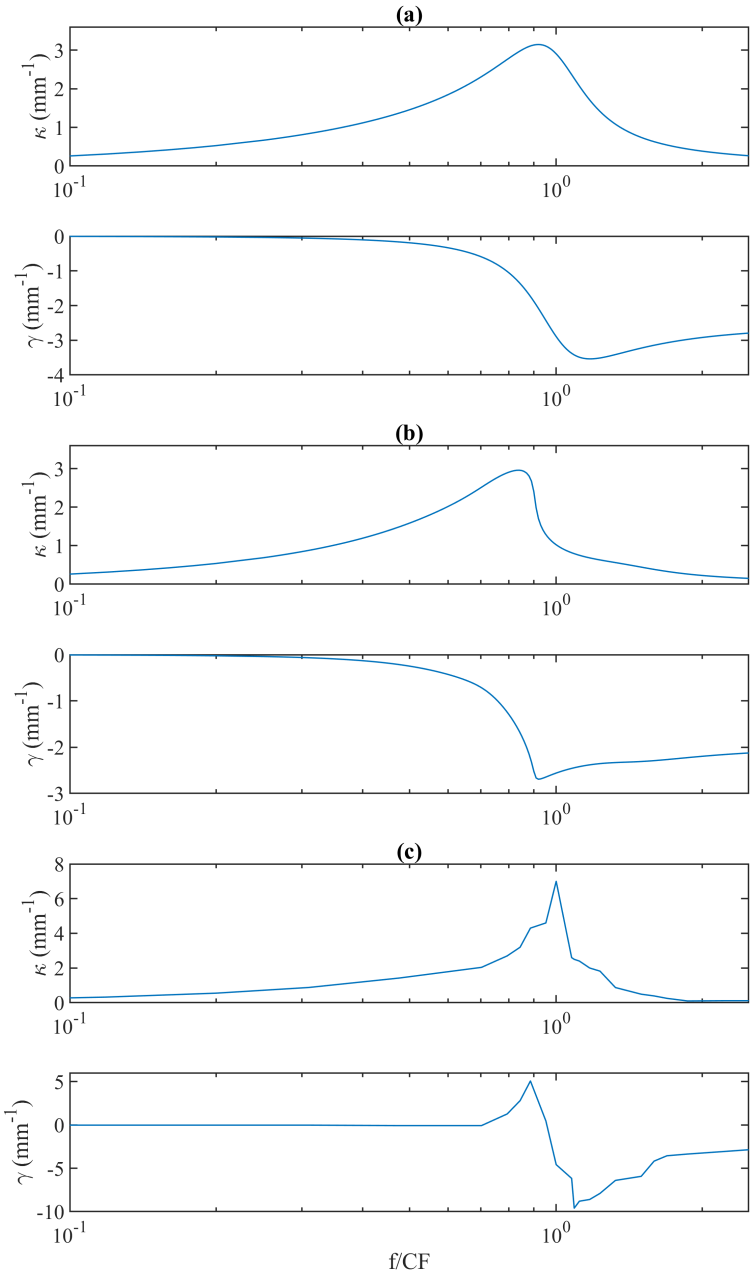


FIGURE 3. Real and imaginary part of the wavenumber as a function of normalized frequency, for (a) the passive case with 1D fluid coupling, (b) the passive case with 3D fluid coupling and (c) the active case with TM longitudinal coupling.

CONCLUSIONS

In this paper, we present a method of deriving the wavenumber distribution associated to the different wave types that can propagate in the cochlea, based on an elemental model of the cochlea mechanics that includes longitudinal coupling in the organ of Corti as part of the 3D fluid coupling. Other methods to derive the wavenumber distribution have been presented by (Shera, 2007; Steele and Taber, 1979). Shera used an inverse method to derive the wavenumber distribution from experimental measurements of the basilar membrane. However, this method can be used to derive only the main travelling wave and not other more heavily damped waves. Steele and Taber used a Lagrangian approach to derive the dispersion equation. Their model includes 3D fluid coupling and longitudinal coupling in the

BM, which is modelled as a plate. By solving the dispersion equation, they obtained a spectrum of roots for the wavenumber and noticed that only one of these corresponds to the main travelling wave, while the others correspond to non-propagating modes. In the method presented in this paper, the BM admittance in the dispersion equation is calculated numerically using the elemental formulation. The wavenumber distribution corresponding to the travelling wave in the case of an active model of the cochlea with TM longitudinal coupling is derived. It is shown that the real part of the wavenumber has a distribution similar to the one for a passive locally reacting model, but with a sharper peak. The imaginary part is positive in a region before the CF, which extension depends on the type of longitudinal coupling.

ACKNOWLEDGMENTS

Riccardo Marrocchio is supported by the DigiTwin project (EP/R006768/1).

REFERENCES

Egbert De Boer. Mechanics of the Cochlea: Modeling Efforts. In R. R. Dallos , P. ; Poppwe, N. A.; Fay, editor, *The Cochlea*, chapter 5, pages 258–317. Springer, 1996 edition, 1996. doi:10.1007/978-1-4612-0757-3_5.

James B Dewey, Brian E Applegate, and John S Oghalai. Amplification and Suppression of Traveling Waves along the Mouse Organ of Corti: Evidence for Spatial Variation in the Longitudinal Coupling of Outer Hair Cell-Generated Forces. *The Journal of Neuroscience*, 39(10):1805–1816, mar 2019. ISSN 0270-6474. doi:10.1523/JNEUROSCI.2608-18.2019. URL https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.2608-18.2019.

Wei Dong and Elizabeth S. Olson. Detection of Cochlear Amplification and Its Activation. *Biophysical Journal*, 105(4):1067–1078, aug 2013. ISSN 00063495. doi:10.1016/j.bpj.2013.06.049. URL https://linkinghub.elsevier.com/retrieve/pii/S000634951300800X.

Stephen J. Elliott, Ben Lineton, and Guangjian Ni. Fluid coupling in a discrete model of cochlear mechanics. *The Journal of the Acoustical Society of America*, 130(3):1441–1451, sep 2011. ISSN 0001-4966. doi:10.1121/1.3607420. URL http://asa.scitation.org/doi/10.1121/1.3607420.

Stephen J. Elliott, Riccardo Marrocchio, and Karl Grosh. Forms of longitudinal coupling in the organ of Corti. In *AIP Conference Proceedings*, 2022.

Riccardo Marrocchio. *Understanding the micromechanics of the cochlea*. PhD thesis, University of Southampton, 2022.

Julien Meaud and Karl Grosh. The effect of tectorial membrane and basilar membrane longitudinal coupling in cochlear mechanics. *The Journal of the Acoustical Society of America*, 127(3):1411–1421, mar 2010. ISSN 0001-4966. doi:10.1121/1.3290995. URL http://asa.scitation.org/doi/10.1121/1.3290995.

Guangjian Ni and Stephen J. Elliott. Comparing methods of modeling near field fluid coupling in the cochlea. *The Journal of the Acoustical Society of America*, 2015. ISSN 0001-4966. doi:10.1121/1.4908242.

Christopher A. Shera. Laser amplification with a twist: Traveling-wave propagation and gain functions from throughout the cochlea. *The Journal of the Acoustical Society of America*, 122(5):2738, 2007. ISSN 00014966. doi:10.1121/1.2783205.

Charles R. Steele and Larry A. Taber. Comparison of WKB calculations and experimental results for three-dimensional cochlear models. *The Journal of the Acoustical Society of America*, 65(4):1007–1018, apr 1979. ISSN 0001-4966. doi:10.1121/1.382570. URL http://asa.scitation.org/doi/10.1121/1.382570.

Larry A. Taber and Charles R. Steele. Cochlear model including three-dimensional fluid and four modes of partition flexibility. *The Journal of the Acoustical Society of America*, 70(2):426–436, aug 1981. ISSN 0001-4966. doi:10.1121/1.386785. URL http://asa.scitation.org/doi/10.1121/1.386785.

Fluid viscosity versus solid damping in a cochlear FEM

Xinyu Zhou[1], Daniel Zackon[1] and Torsten Marquardt[1, a)]

[1]*UCL Ear Institute, University College London, 332 Gray's Inn Road, London, WC1X8EE, United Kingdom*

[a)]Corresponding author: t.marquardt@ucl.ac.uk

Abstract. For reasons of computational efficiency, most cochlear models simulate power dissipation solely by damping inside the solid cochlear partition, disregarding viscous dissipation in the fluid. In recent years, however, simulation incorporating fluid viscosity became reasonably affordable. In this study, we compared the effect of both types of power dissipation using a passive 3D box model implemented in the commercial finite-element modelling (FEM) package COMSOL. Indeed, memory requirements and computation time was approximately 10 time larger when simulating fluid viscosity compared to solid damping only. But the qualitative difference was striking: As expected, when lowering the solid damping in a model without fluid viscosity, the peak of the travelling wave (TW) got higher, moved apically and the apical slope of the TW got much shallower. (I.e., the distance over which the wave amplitude decayed increased substantially.) When lowering the fluid viscosity in a model without solid damping, the peak of the TW also got higher and moved apically, but the slope of decay remained constant. The slope of approximately 30 dB/mm was much steeper than could be achieved with solid damping and compares to the more than 100-dB/octave slopes at the high-frequency side of experimentally measured frequency tuning curves. We conclude that that solid damping cannot replace the computation of viscous dissipation when intending to simulate the sharp TW decay. They have qualitatively different effects.

INTRODUCTION

Fluid-mechanically, the cochlear ducts are small relative to the thickness of the Stokes boundary layer and power dissipation herein are likely to dominate the damping of the travelling wave (TW). But because the simulation of viscous dissipation requires large computational resources, cochlear finite-element models (FEMs) have commonly generated losses in their solid cochlear partition. Although commercial software with thermo-viscous solvers became affordable within the last decade, the mesh density required to resolve the Stokes boundary layers requires a large number of small elements, and therefore, large scale computers. In our experience with COMSOL Multiphysics, just the switching to using its Thermoacoustics Interface increased the requirement for both working memory and computation time by a factor of ten, even with identical mesh density. Thus, we wondered whether expensive simulations with viscous solvers are strictly necessary to simulate a realistic TW.

METHODS

The cochlea was implemented as an uncoiled box with two fluid compartments separated by a 35-mm long solid that represented the cochlear partition (Fig. 3a). The cochlear partition had the density of water, a width of 150 μm at base, 450 μm at the apex and a constant thickness of 10 μm. It was an orthotropic solid with a Young's modulus in the longitudinal direction (E_x) a thousand times smaller than across (E_y). Along the longitudinal location (x), E_y was adjusted to give the model roughly the tonotopy of the human cochlea, resulting for characteristic frequencies above 500 Hz in an octave spacing of roughly 5 mm[1]. It was achieved by adjusting empirically the parameter values in the following formula that gives E_y in Pascals. (The distance from the base (x) is given in meter.):

$$E_y = 5 \cdot 10^7 \left(1 - \frac{x}{0.045}\right)^{4.5} \quad (1)$$

The shear moduli were $G_x = (E_y/Pa)/50$ N/m^2, $G_y = (E_y/Pa)/10$ N/m^2 and $G_z = 1$ N/m^2. The corresponding Poisson ratios were {0.005, 0.3, 0.005}.

A 1.35-mm long compartment with a 0.45-mm wide helicotrema was added to the apical end. The fluid had the mechanical properties of water.

The mesh sectioned the model over the length of the cochlear partition in 700 equal-sized sections. The partition was 1 element thick. The cross-sectional mesh of cochlear partition and fluid compartments was non-uniform to resolve the shear movement within the Stokes boundary layers more finely (see mesh lines in the cross section shown in Fig. 2). Since the model was symmetric (with the cochlear partition and helicotrema symmetrically divided along their midline), all middle surfaces were given a symmetry constraint so that only one-half of the model needed to be solved (450800 elements, 12681661 degrees of freedom).

The acoustic input was applied by defining a perpendicular harmonic displacement to fluid surface above the cochlear partition at longitudinal position zero, thus acting as the oval window (OW). The same fluid surface below the cochlear partition was unconstrained and acted as the round window. Surfaces between fluid and cochlear partition were given a non-slip fluid-structure interaction. The two surfaces next to the solid cochlear partition also had a non-slip constraint. All the outer boundaries of the box had slip constraint because the mesh was here too coarse to resolve the Stokes boundary layers, and thus errors would have occurred when calculating the power flux with a non-slip constrained. Instead, the nodes were used to better resolve the Stokes boundary layers at the cochlear partition, where all of the power dissipation occurs at the TW peak and apical thereof due to the short local wavelength.

Three simulations were run with only viscous dissipation. Both, the bulk and dynamic viscosity of the fluid were set to half, equal and double of that of water at 37°C in these three simulations respectively, while all damping parameter of the solid material were set to zero. Three further simulations were run with the two fluid viscosity parameters set zero, while the value of the six damping parameters (three associated with the tensile and the three with the shear strain of the solid) were adjusted to produce TW envelopes that are most comparable with those obtained with the previous three viscosity settings. All six damping loss factors were here set to the same value, which was 0.0025 in the first, 0.05 in the second and 0.1 in the third simulation. In order to keep the conditions identical, all simulations used the Thermoacoustics Interface, also those with zero viscosity.

The direct stationary solver PARDISO ran on a Dell PowerEdge R910 server (four Xeon CPUs E7- 4870 @ 2.40 GHz, 40 cores total). One of these frequency domain computations took approximately 3 hours, requiring almost all of the available 1 TB RAM.

The time-averaged power flux through the cross section at location x was calculated post-hoc as follows[2]:

$$P_{CS}(x) = \int_0^{L_3}\int_0^{L_2}\frac{1}{T}\int_0^{T} u\left(p - 2\mu\frac{\partial u}{\partial x}\right) - \mu v\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right) - \mu w\left(\frac{\partial w}{\partial x} + \frac{\partial u}{\partial z}\right) dt\, dy\, dz, \tag{2}$$

where T is the period of a cycle, μ is the dynamic viscosity, and u, v, and w denote the velocity components in longitudinal (x), radial (y) and transversal (z) directions respectively. L_2 and L_3 are the y- and z-dimensions of the cross section. For this computation, the node density was locally increased to 1-um resolution utilizing Matlab's *griddedInterpolant* with the modified Akima cubic interpolation method.

RESULTS

Panel (a) and of Fig. 1 shows the magnitude of the travelling waves as a function of longitudinal location relative to the OW displacement for various values of fluid viscosity while the solid damping was zero (solid lines) and various damping values for the solid material of the cochlear partition while the fluid viscosity was zero (dashed lines). As expected, reduction in the loss parameters cause in both cases the TW peak to move further apically, leading also to only moderate increases in peak amplitude. The decay of the TW shows however striking differences: Overall, it becomes obvious that the solid damping cannot achieve the steep slope produced by viscosity, not even at a cost of 10-dB peak reduction. Second, whereas the slope of the decay becomes shallower when reducing solid damping, the value of fluid viscosity appears to have no effect on its steepness. We will investigate the reason for this further below.

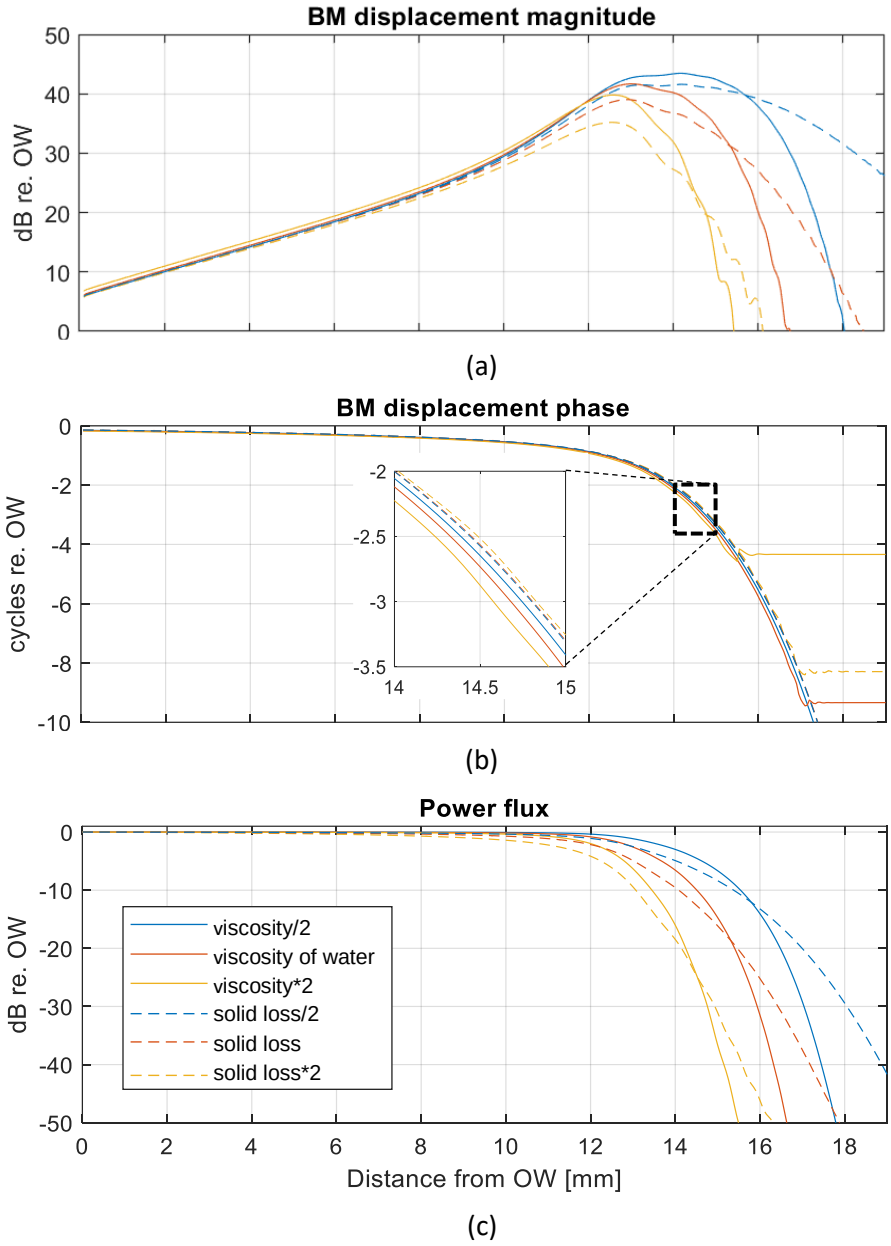


Figure 1. Difference between power dissipation by fluid viscosity (solid lines) and damping within the solid cochlear partition (dashed lines) on magnitude (a), phase (b) and power flux (c) of cochlear TWs. Three parameter values were tested for each type of dissipation (see legend in panel c).

Although at the first glance the type and value of loss does not make a large difference to the TW phase (panel b), the magnified inset reveals that the two types of dissipation have opposite effects: An increase in fluid viscosity slows the TW (i.e., phase accumulation increases), like the TW is facing a higher impedance in apical direction and forces more power to be locally stored in the cochlear partition. On the other hand, an increase in solid damping speeds the wave slightly up, like the increased partition impedance deflects the TW power more along the fluid-filled scalae.

The qualitative difference in power dissipation by solid damping and fluid viscosity becomes most obvious by an analysis of the longitudinal power flux. A comparison of the superimposed graphs in panel (c) reveals that in contract to the case of solid damping (dashed lines), the TWs with viscous dissipation (solid lines) hardly lose any power before their peaks, but that the power even more rapidly dissipates as the wavelength decreases past the peak. Analogous to the slope of the decaying wave amplitude, the dissipation rate per distance (in dB/mm) is largely independent from the amount of fluid viscosity. In contrast, the rate of power dissipation with damping in the solid is inversely related to the material's loss factor.

Panel (c) shows clearly that the power has already considerably decreased before the TW reaches its peak, even more so in case damping in the solid than in the fluid. The reason that the TW amplitude is increasing towards its peak despite the power loss is that the power flux focuses more and more on the cochlear partition as the wavelength decreases. This is illustrated in the left column of panels of Fig. 2. They show the power flux density within cross sections at x-positions, where the total power of the three viscous TWs has dropped by 30 dB relative to the input at the OW. The local wavelengths at these three locations are 0.65, 0.4 and 0.35 mm (as can be inferred from the TW phase slope in panel b). It can be seen that the spatial extent of the power flux decreases roughly linearly with local wavelength (in agreement with Lighthill's analytically derived equations for the two-dimensional fluid velocity field[2]). Because the three cross sections were selected based on equal total power flux, it is obvious that the maximum power densities in these cross sections must differ accordingly. The power density, located close to the cochlear partition, determines its displacement amplitude, which were 20.2 dB, 21.7 dB and 25.6 dB re. OW at these three cross sections (see panel a). In addition, the decrease in cochlear partition stiffness towards the apex contributes to these displacement differences because a higher displacement is required to store the same amount of potential energy at a less stiff location.

The almost identical slopes of the viscous power flux functions at the x-positions (panel c; look where the power has dropped by 30 dB) shows that also the power dissipation is here very similar. (The power drop is 1.28 dB, 1.15 dB and 1.09 dB per 50 µm respectively). To better illustrate this apparent viscosity-independence, the right column of panels in Fig. 2 shows the differences in the spatial distributions of this power dissipation. Dissipation occurs in the Stokes-layer as well as in areas of irrotational-flow fields[3]. Comparison of panel (b) and panel (f) reveals that decreasing the viscosity by four roughly halves their spatial extent and, consequently, reduces the area to a quarter. This is explained by the facts that the thickness of the Stokes-layer is proportional to the square root of the fluid's viscosity and the spatial extent of the irrotational field is proportional to the local wavelength[3], which differs at these two x-positions roughly by a factor of two (see above). At the same time, the labels at the color bars show that the maximum dissipation density differs inversely, roughly also by a factor 4. The reason for the increase in dissipation density (D) is that it growths proportional with the square of the velocity gradient, du/dz:

$$D = \mu \cdot \left(\frac{du}{dz}\right)^2 \tag{3}$$

Two things contribute to the increase in the velocity gradient. First, the reduced spatial extent of the velocity field and, second, the larger displacement of the cochlear partition at the cross section in panel (f), implying that also the longitudinal velocity (u) is 6 dB larger. Taken together, the term du/dz is a factor of four larger in panel (f) than in panel (b). The effect of its squaring in Eq. 3 is reduced by the lower dynamic viscosity (μ) in panel (f), so that the power density panel (f) is four times that in panel (b) and approximately compensates the fourfold increase in area. This explains why the amount of dissipation in these cross sections is similar and also the before-mentioned viscosity-independence of the rate at which the TW decays apically of its peak when dissipation occurs in the fluid.

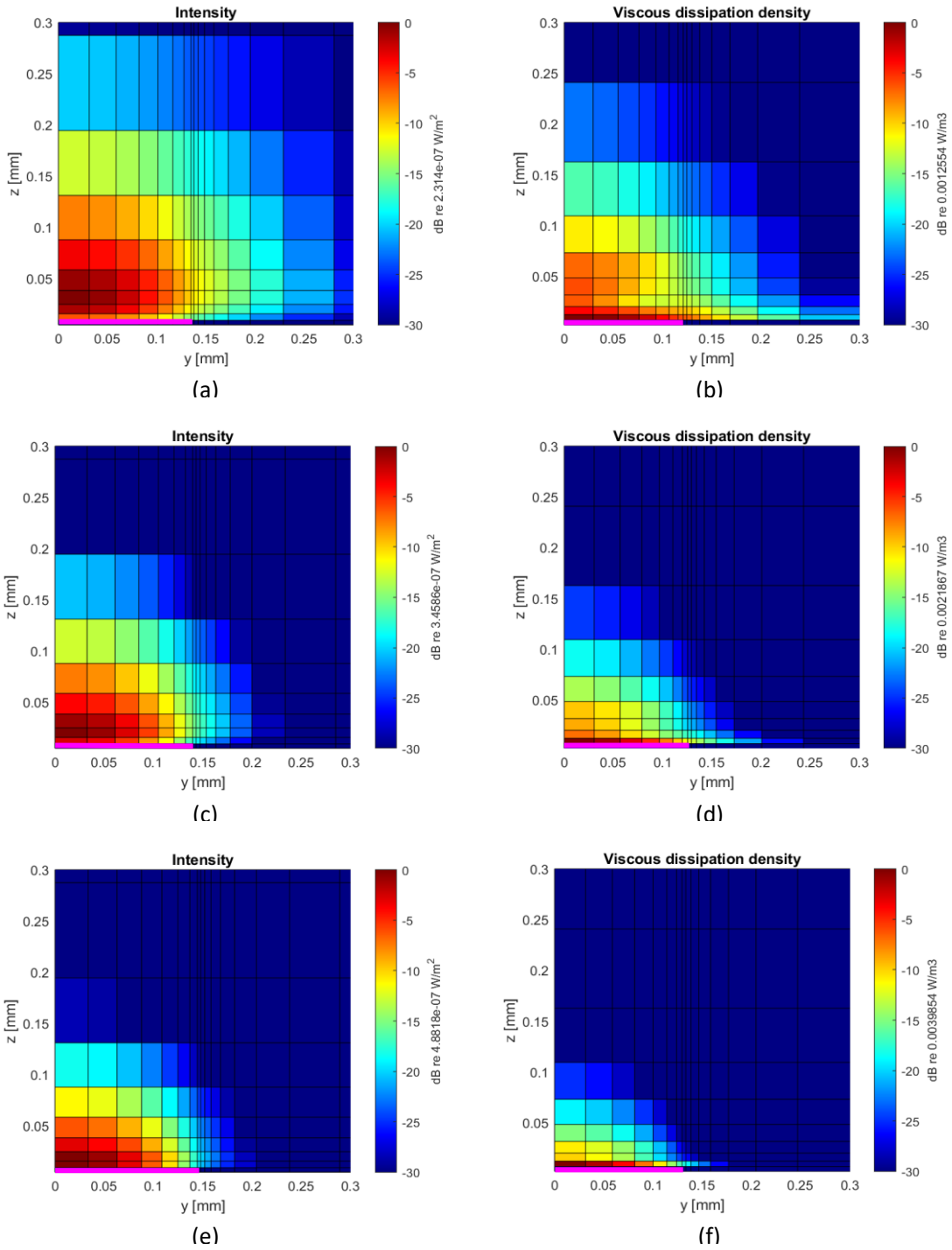


FIGURE 2. Model cross sections showing power flux density (left) and dissipation density (right). Note that only one quadrant of the total cross sections is shown (with the [0,0]-position being the center of the cochlear partition). The panels in the three rows stem from the simulations with the three different viscosity settings. The x-positions of these cross sections are where the power flux has dropped by 30 dB, which are 14.75 mm, 15.95 mm and 17.05 mm from the OW for the first, second and third row respectively. The mesh is finer at the nonslip fluid-solid boundary at the cochlear partition to resolve the Stokes boundary layer and also at the edge of the cochlear partition (which is indicated by the pink bar at the lower left bottom).

DISCUSSION

Already in the 1920's, Bekesy observed the steep slope at which the cochlear travelling wave decays beyond its peak[4]. Later measurements with more sensitive techniques confirmed this steep slope in form of spectral tuning curves with high-frequency slopes exceeding 100 dB/octave[5,6], even in post-mortem in human temporal bones[7,8]. In contrast to the tuning peak, the steepness of this slope is level-invariant, and also stable in its spectral and spatial position. Because normal hearing people have no difficulties understanding loud speech, although the narrow spectral tuning of the active cochlea diminishes quickly with rising sound level, we believe that the sharp tuning at low sound intensities is possibly just an epiphenomenon of preserving sound energy to achieve high displacement amplitude at the characteristic place. Instead, cochlear frequency analysis might be based on the steep and level-independend apical edge of the mechanical excitation pattern – a so far untested proposition.

This FEM study shows that only viscous dissipation can achieve these steep experimentally-observed slopes because its effect increases drastically with the decreasing wavelength past the TW peak. The reason behind this wavelength-dependence is that the spatial extent of the fluid viscosity field reduces with the shortening of the TW and is increasingly submerged into the Stokes boundary layer at the cochlear partition where power is rapidly dissipated. It was interesting to observe that the steepness of the TW decay is fairly independent of the fluid's viscosity value. In other words, mechanisms that could potentially reduce the viscous effect by thickening the Stoke boundary layer, and thus, reducing shear velocity[9] (like the compliant cell layers underneath the basilar membrane), might be able to shift the onset of power dissipation to more apical locations without reducing the final slope of the TW decay.

ACKNOWLEDGMENTS

Xinyu Zhou would like to thank the RNID for his PhD studentship award (S61).

REFERENCES

1. D. D. Greenwood, J. Acoust. Soc. Am. **87**, 2592–2605 (1990).
2. Y. Wang et al., Sci Rep **6**, 19475 (2016).
3. J. Lighthill, J. Fluid Mech. **239**, 551–606 (1992).
4. G. von Békésy, Experiments in Hearing (McGraw-Hill Book Co, New York, 1960).
5. W. S. Rhode, J. Acoust. Soc. Am. **49**, 1218–1231 (1971).
6. M. A. Ruggero et al., J. Acoust. Soc. Am. **101(4)**, 2151–2163 (1997).
7. T. Gundersen et al., Acta Otolaryngol. **86**, 225–232 (1978).
8. S. Stenfelt et al., Hear. Res. **181**, 131-143 (2003)
9. M. van der Heijden and A. Vavakou, Hear Res. Oct 2:108367 (2021).

Energetic Depletion and IsK Mutations Destabilize Potassium Resupply to the Endolymph

Julia Lasater[1, a)] and Robert M. Raphael[2, b)]

[1)] *Department of Biomedical Engineering, University of Pennsylvania, Philadelphia, PA*

[2)] *Department of Bioengineering, Rice University, Houston, Texas*

a) *jlasaster@seas.upenn.edu*

b) *Corresponding author: rraphael@rice.edu*

Abstract.

In the process of mechanotransduction, auditory and vestibular hair cells drain potassium ions (K+) from the endolymph. The resupply of K+ to the endolymph requires significant energy expenditure and is accomplished by a layer of specialized epithelial cells in the cochlea (marginal cells) and vestibular system (dark cells). We constructed a biophysical model of ion transport across these epithelial layers by implementing mathematical expressions that describe the activity of known ion channels and transporters expressed in the marginal/dark cells. This network constitutes a dynamical system and the net transepithelial potassium current (i_K^{te}) can be studied as a function of parameters such as external potassium concentration and ATP levels which influence ion transporter activity and affect K^+ homeostasis in the endolymph. The dependency of i_K^{te} on external potassium concentrations is nonlinear and can be used to study the stability of the K^+ cycle and gain insight into how potassium resupply is altered as a result of energetic depletion and genetic mutations. A particularly interesting result is that increasing the *Isk* conductance expressed on the apical surface of marginal cells and dark cells causes the development of hypersensitive regions in the transport phase diagram that provide insight into the mechanism of-induced hearing loss.

INTRODUCTION

In the inner ear, the apical surfaces of hair cells are bathed by endolymph, a specialized fluid within the cochlea and the vestibular system. The endolymph has a large potassium ion concentration (150 mM) that serves as the driving force for transduction by hair cells; in the cochlea the endolymph is also at an elevated electrical potential and the driving force is greater. Deflection of hair cell stereocilia by sound (cochlea) or head motion (vestibular system) opens mechanosensitive channels (MET). Potassium ions (K^+) flow through the MET channels into the hair cells and cause hair cell depolarization. The two types of hair cells in the cochlea respond differently to membrane depolarization - inner hair cells release neurotransmitters which signal the auditory nerve, while outer hair cells undergo mechanical deformation which amplifies vibrations of the organ of Corti. In the vestibular system, type I and type II hair cells release neurotransmitters to signal the vestibular nerve; type I cells additionally support non quantal transmission.

Hair cell mechanotransduction process drains K^+ from the endolymph. Sustaining a potassium current into the endolymph to replenish the lost K^+ is an energy intensive process that takes place in the stria vascularis (cochlea) and the dark cell layer (vestibular system). The stria vascularis is a highly vascular multilaminate epithelium composed of two functional cellular layers. The layer facing the scala media is composed of marginal cells that have extensive basolateral infoldings rich in mitochondria and a high density of the Na^+-K^+-ATPase and NKCC1. The apical ends of the cells contain an I_{sk} current composed of subunits of KCNQ1/KCNE1 that returns potassium to the endolymph. A similar arrangement occurs in the dark cells and is shown in Figure 1a. Due to the complexity of the cochlea and vestibule, the biophysics of marginal cell and dark cell ion transport processes are not fully understood but are essential for understanding function and dysfunction in these systems.

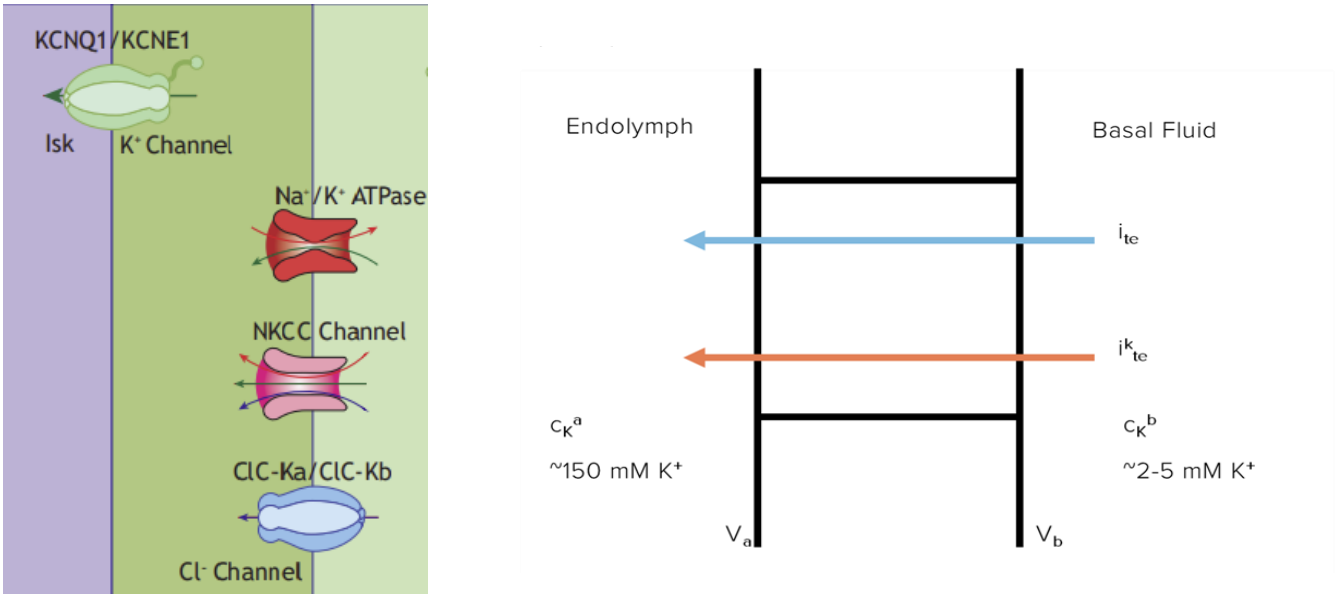


FIGURE 1. a) Major Ion Channel Distribution in Cochlear Marginal Cells and Vestibular Dark Cells. b) Definition of the variables used in the model

Mutations in ion channels expressed in the inner ear can cause hearing loss and vestibular dysfunction. One example is a mutation of the KCNE1 gene which encodes a protein that is a subunit of voltage activated potassium channels expressed in the apical membrane of the marginal/dark cells. This mutation is associated with a greater susceptibility to noise-induced hearing loss [1]. Interestingly, this mutation increases the Isk conductance, but it is unknown how this predisposes individuals to hearing loss.

We have expanded a previously developed computational model of ion transport across the marginal and dark cells [2] to improve accuracy and deliver better insight into the function and dysfunction of the marginal cells and dark cells. The upgraded model is able to simulate the influence of extracellular $[K^+]$ on the transepithelial K^+ current and investigate how genetic mutations in ion channels give rise to deafness and noise-induced hearing loss.

GENERAL DESCRIPTION OF THE MODEL

We present a steady-state model of ion transport across a marginal cell/vestibular dark cell exposed to a constant ionic environment on the apical and basolateral sides. The apical fluid represents the endolymph whose ion concentrations have been measured. In the cochlea, the basal fluid occupies the intrastrial space, in the vestibular system the basal fluid occupies the intercellular spaces surrounding the dark cells. Ion concentrations are difficult to study in these fluids due to their small volumes. The model enables the study of how perturbations in ion concentrations influence ion movement from and to these small spaces. For example, Quraishi and Raphael predicted how the potassium concentration in the basolateral fluid affected the generation of the endocochlear potential [3].

To develop the model, flux expressions that represent the activity of the major channels were utilized. The parameters in the model are taken from the literature as previously described [2]. In the original model, a simplified model of the Na/K ATPase was utilized that did not include ATP dependency. We modified the model of ion transport developed by Quraishi and Raphael [2] to include ATP as variable, implementing Oka, Cha, and Noma's [4] 15-state

post-Albers reaction cycle model for the Na^+-K^+-ATPase. We used parameters for the *α1 β2* isoform of the pump expressed in the marginal cells. Using these formulas in an extended computational model, we were able to derive a value for the rate of the Na^+-K^+-ATPase pump dependent on the concentration of ATP, which was then integrated into the full model. The system of ordinary differential equations was solved in Matlab. The output of the model are the concentrations of Na^+, K^+ and Cl^- inside the marginal cell, the steady state membrane potentials across the basal and apical surfaces of the cell and the net transepithelial currents across the cell.

RESULTS

Validation of the Model. We confirmed that the expanded model reproduced all the previous results obtained in the original model by Quraishi and Raphael [2]. This included predictions for how the density of Isk, Na/K ATPase, and NKCC affect the open and short circuit currents.

Open and Short Circuit Conditions. The predictions from the extended Na^+-K^{+-}ATPase model can be used to study the effect of various parameters on cell function. A standard approach used in epithelial physiology is to evaluate the functional capacity of the system in open/short circuit conditions. The open circuit generates maximal voltage while the short circuit generates maximal current. The physiological behavior of the system will depend on the resistance of the pathway that returns K^+ to the perilymph bathing the basal sides of the marginal/dark cells. Considering the importance of rapid potassium recycling in the cochlea it is likely that the short circuit condition is more representative of the native physiological system but we also show predictions for open circuit voltages.

Energetic Depletion. The expanded model permits an examination of how ATP levels affect the response of the system. The results indicate that the current carried by the Na^+-K^+-ATPase, the K^+ carried by the Na^+-K^+-Cl^- cotransporter (NKCC), the net K^+ current across the epithelium (i_K^{te}), and the net ionic current across the cell (i_{te}) all begin to decline when the ATP concentration in the cell falls below 1 mM (Figure 2). Of particular physiological significance is that the model predicts that i_K^{te} can reverse direction - meaning that potassium will be transported out of the endolymph into the cell. These results have implications for the cochlea and vestibular system to function and survive during episodes of energetic depletion which may occur in response to loud noises that generate reactive oxygen species.

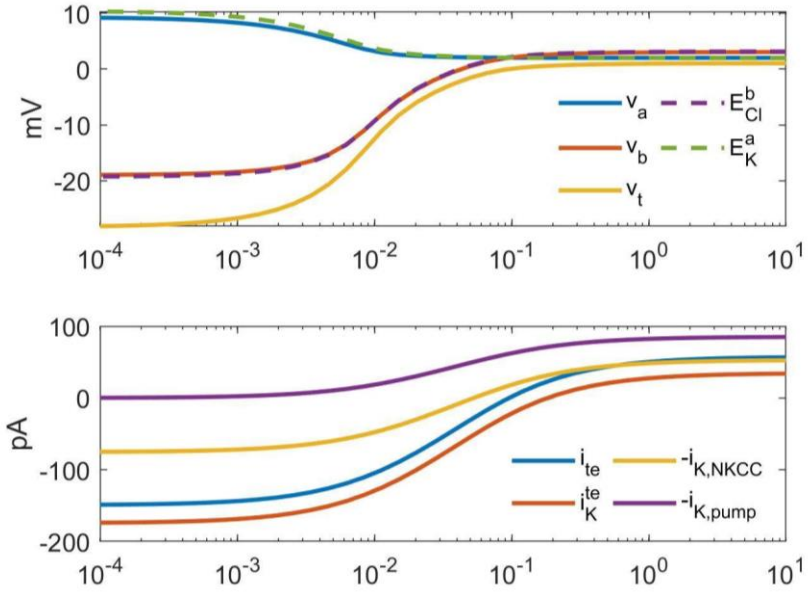


FIGURE 2. Effect of ATP on Potassium Transport Across the Marginal Cell. Legend: $\mathbf{v_a}$ = apical voltage; $\mathbf{v_b}$ basolateral voltage; $\mathbf{v_t}$ = transepithelial voltage difference; $\mathbf{E_{cl}^b}$ = basal Nernst potential for Cl^-; $\mathbf{E_k^a}$ = apical Nernst potential for K^+; $\mathbf{i_{te}}$ = total transepithelial current; $\mathbf{i_K^{te}}$ = transepithelial K^+ current; $\mathbf{-i_{K,NKCC}}$ = K^+ current across NKCC cotransporter; $\mathbf{-i_{K,pump}}$ = K^+ current across Na^+-K^+-ATPase pump. Note that currents are picoamps per cell.

Contour Graphs of Net Potassium Resupply Current (i_K^{te}). The model is able to study any desired parameter as a function of other parameters. For example, given that apical and basal potassium concentrations are major driving forces for potassium transport, a diagram can be made that predicts the current across the marginal cell as a function of the apical and basal potassium concentrations. This can be thought of as a phase diagram that shows how changes in potassium alter the steady state currents across the marginal cell. An example of this plot is shown in Figure 3a. In mathematical terms, the contours in the diagram reflect level curves with various i_K^{te} values. The values shown in the contours are pA/cell. The contour labeled zero (0) reflects the situation in which there is no net current across the epithelia. Below the zero contour, we see that currents are positive, meaning that potassium is transported into the endolymph and above the zero contour, we see that currents are negative, meaning that potassium is transported out of the endolymph. The negative currents are caused by high levels of endolymphatic potassium.

Looking at the predictions in Fig. 3a carefully we see that the dependency of i_K^{te} on external potassium concentrations is nonlinear. Under normal conditions, i_K^{te} increases monotonically for basolateral potassium concentrations up to ~ 7 mM, and then approaches saturation. The system exhibits stable behavior below values of endolymph potassium of ~160 mM for basal potassium concentrations up to 15 mM. However, a surprising finding of this model is that, above 160 mM, additional stable states can appear.

Isk (KCNE1) Mutations. At resting conditions, Isk is poised at approximately 50% of its maximal activity [2, see Fig 4]. This is expected as the marginal/dark cells will have to respond to a range of loads on the stereocilia. Mutations in the genes responsible for Isk conductance have been shown to cause hearing and vestibular dysfunction. For example, a mutation in a potassium subunit (KCNE1) responsible for I_{sk} is associated with susceptibility to noise-induced hearing loss [1]. This mutation increases the conductance of I_{sk}, and it is unclear why this leads to hearing loss. To simulate the effects of this mutation, we increase the conductance by 5x in the model. The results on the phase state of the system are shown in Fig. 3b. As shown, making the apical potassium channel (KCNE1) overactive has little effect on the system under normal conditions of ~ 4mM basolateral K^+. However, increasing I_{sk} conductance results in the appearance of a second steady state and a region of hypersensitivity at high basolateral K^+ concentrations (greater than 10 mM). These effects are not unexpected in dynamical systems where steady state behavior often displays sensitivity to parameter values. The appearance of unstable transport regimes is expected to lead to cochlear dysfunction as they will lead to imbalances in endolymphatic potassium that drives MET channels in hair cells.

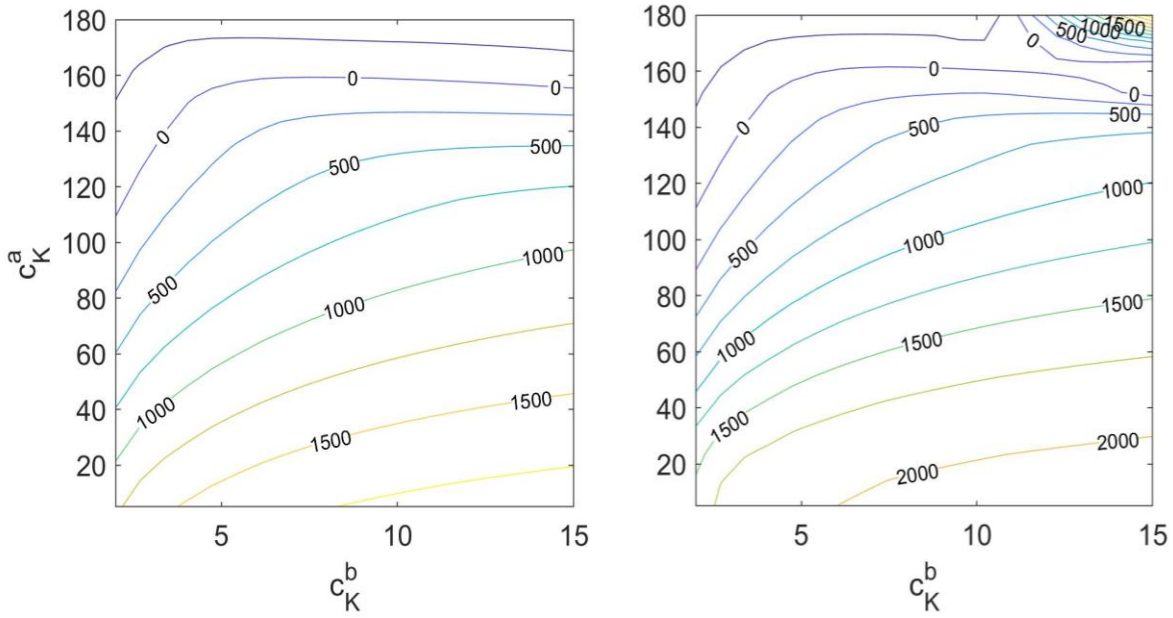


FIGURE 3. Increasing I_{sk} Conductance Causes Unstable Transport Regimes to Develop. The contours on the graphs below show the predicted transepithelial potassium current (i_K^{te}) in pA/cell as a function of the K^+ concentration on the apical and basolateral sides of the marginal/dark cells at a) normal I_{sk} conductances and b) elevated I_{sk} conductances (x5). As shown, making the apical potassium channel (KCNE1) overactive has little effect on the system under normal conditions of ~ 4mm basolateral K+. However, increasing I_{sk} conductance causes a second steady state and region of hypersensitivity to appear at higher basolateral K^+ concentrations.

Effect of ATP Reduction on Potassium Resupply Current. We can also study the effect of reduced ATP on the resupply current i_K^{te} . Energetic depletion was modeled by decreasing the ATP available for the Na+/K+ pump. To visualize the effects of reduced ATP on i_K^{te}, we chose a single contour of the phase diagram that represented i_K^{te} = 50 pA and studied how this contour changed at different ATP levels. ATP levels below 1 mM caused marked shifts in normal behavior of i_K^{te} , lowering the steady state i_K^{te} contours and causing the i_K^{te} current to reverse at normal values of endolymphatic potassium (Figure 4).

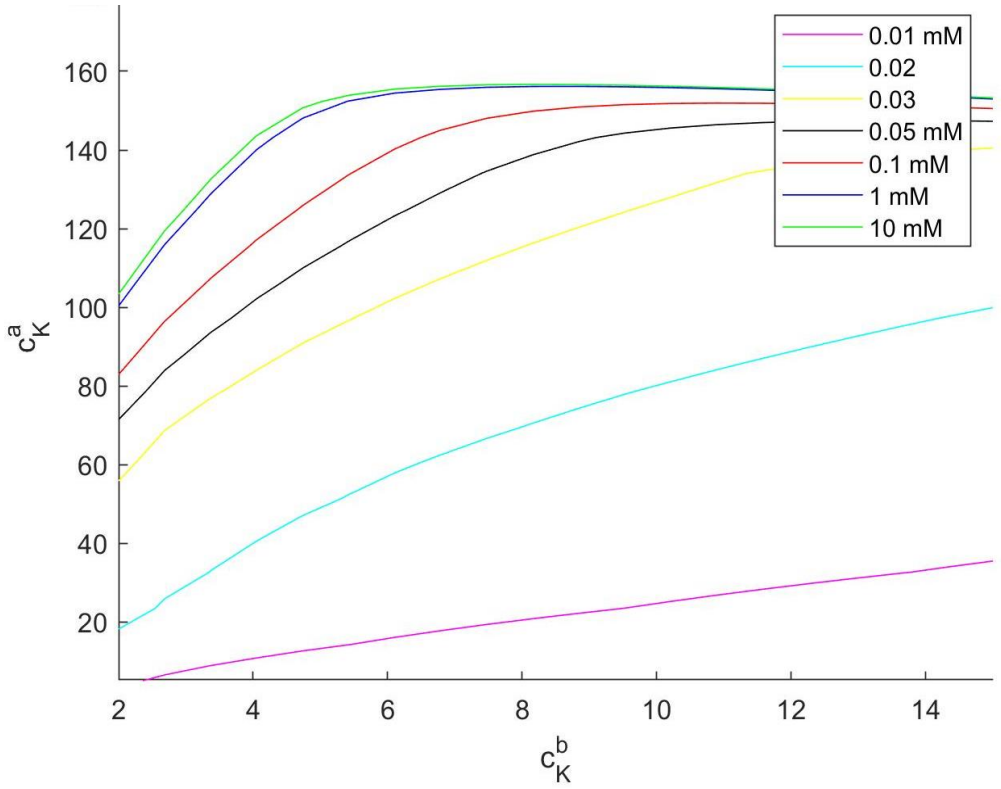


Figure 4. Effects of reducing ATP from control for a transeptihelial current of 50 pA/cell. Below 1 mM ATP, steady state will be established at lower values of endolymphatic K^+. System is most sensitive to ATP depletion at normal value of basolateral K^+ (3-4 mM)

DISCUSSION

This is, to our knowledge, the first model to simulate the effect of ATP reductions on ion transport across the marginal/dark cell layers. This is also the first model to predict the effect of overactivity in the Isk conductance on potassium resupply.

The results indicate that the current carried by the Na^+-K^+-ATPase, the K^+ carried by the Na^+-K^+-Cl^- cotransporter (NKCC) and the net K^+ current across the epithelium (i_K^{te}) all begin to decline when the ATP concentration in the cell falls. Of particular physiological significance is that the model predicts that i_K^{te} can reverses direction, meaning that potassium will be transported out of the endolymph. These results have implications for cochlear and vestibular function and survival during episodes of energetic depletion which may occur in response to loud noises that generate reactive oxygen species.

The model also predicts that even modest increases in Isk (by 5x) can introduce instabilities in K^+ transport when basolateral K^+ is elevated from its normal value of ~2-4 mM. These predictions provide a potential explanation for the genetic susceptibility to noise-induced loss in individuals with the KCNE1 mutation that increases Isk by 5-16x [1] - such individuals will experience dysregulation and instability in ion homeostasis if they experience elevated basolateral potassium.

This extension of the Qurashi and Raphael model [2] lays a foundation for future investigations of how energy depletion and genetic mutations affect the stability of potassium resupply to the endolymph and thereby ionic

homeostasis of the endolymph. The model made the surprising prediction that the marginal/dark cell ion transport system can exhibit multiple steady states and hypersensitivity. This type of behavior is inherent to dynamical systems and has been studied extensively in neuronal systems [5] but to our knowledge has not been reported in epithelial transport systems. In future work the dynamical behavior of the potassium resupply system needs to be examined in greater detail by determining whether the additional equilibrium states are stable or unstable and whether the system exhibits bifurcations or limit cycle oscillations.

The model has been solved at steady state but can be adapted to solve the system when perturbations occur. The model has also been solved in isolation from the other cells and fluids in the tissue, but can be integrated into models of potassium cycling in the entire cochlea [6,7].

CONCLUSION

The incorporation of the 15 state post-Albers model of the Na^+-K^+-ATPase to our previous ion transport model yields a more detailed and biologically accurate model of ion transport in the inner ear. Explicitly including ATP as a variable enables us to simulate the effects of energetic depletion, which is implicated in several inner ear disorders. Our results indicate that mutations affecting channel conductance can lead to instabilities in ion transport processes within marginal cells and vestibular dark cells and leave the system in an irrecoverable, pathologic, transport regime. This insight highlights the utility of computational modeling in analyzing the causes of hearing loss and vestibular dysfunction.

ACKNOWLEDGEMENTS

This work was supported by an NSF IGERT 1250104 and NIH R01 DC012347-08S17. We thank Aravind Govindaraju for assistance.

REFERENCES

1. L Van Laer, PI Carlsson, N Ottschytsch, ML Bondeson, A Konings, A Vandevelde, N Dieltjens, E Fransen, D Snyders, E Borg, A Raes, G Van Camp, The contribution of genes involved in potassium-recycling in the inner ear to noise-induced hearing loss. Hum Mutat. 2006 Aug;27(8):786-95. doi: 10.1002/humu.20360. PMID: 16823764.
2. IH Quraishi and RM Raphael, Computational model of vectorial potassium transport by cochlear marginal cells and vestibular dark cells. Am J Physiol Cell Physiol. 2007 Jan;292(1):C591-602. doi: 10.1152/ajpcell.00560.2005.
3. IH Quraishi and RM Raphael, Generation of the endocochlear potential: a biophysical model. Biophys J. 2008 Apr 15;94(8):L64-6. doi: 10.1529/biophysj.107.128082.
4. C Oka, CY Cha, A Noma, Characterization of the cardiac Na+/K+ pump by development of a comprehensive and mechanistic model. J Theor Biol. 2010 Jul 7;265(1):68-77. doi: 10.1016/j.jtbi.2010.04.028.
5. Eugene Izhikevich. *Dynamical Systems in Neuroscience*. (MIT Press, Boston, 2010).
6. IH Quraishi, "Computational Models of Potassium Transport in the Inner Ear," PhD thesis, Rice University, 2008
7. F Nin , H Hibino, S Murakami, T Suzuki, Y Hisa, Y Kurachi, Computational model of a circulation current that controls electrochemical properties in the mammalian cochlea. Proc Natl Acad Sci U S A. 2012 Jun 5;109(23):9191-6. doi: 10.1073/pnas.1120067109.

Energetic Depletion and IsK Mutations Destabilize Potassium Resupply to the Endolymph

Julia Lasater[1, a)] and Robert M. Raphael[2, b)]

[1)] *Department of Biomedical Engineering, University of Pennsylvania, Philadelphia, PA*

[2)] *Department of Bioengineering, Rice University, Houston, Texas*

a) *jlasaster@seas.upenn.edu*

b) *Corresponding author: rraphael@rice.edu*

Abstract.

In the process of mechanotransduction, auditory and vestibular hair cells drain potassium ions (K+) from the endolymph. The resupply of K+ to the endolymph requires significant energy expenditure and is accomplished by a layer of specialized epithelial cells in the cochlea (marginal cells) and vestibular system (dark cells). We constructed a biophysical model of ion transport across these epithelial layers by implementing mathematical expressions that describe the activity of known ion channels and transporters expressed in the marginal/dark cells. This network constitutes a dynamical system and the net transepithelial potassium current (i_K^{te}) can be studied as a function of parameters such as external potassium concentration and ATP levels which influence ion transporter activity and affect K^+ homeostasis in the endolymph. The dependency of i_K^{te} on external potassium concentrations is nonlinear and can be used to study the stability of the K^+ cycle and gain insight into how potassium resupply is altered as a result of energetic depletion and genetic mutations. A particularly interesting result is that increasing the *Isk* conductance expressed on the apical surface of marginal cells and dark cells causes the development of hypersensitive regions in the transport phase diagram that provide insight into the mechanism of-induced hearing loss.

INTRODUCTION

In the inner ear, the apical surfaces of hair cells are bathed by endolymph, a specialized fluid within the cochlea and the vestibular system. The endolymph has a large potassium ion concentration (150 mM) that serves as the driving force for transduction by hair cells; in the cochlea the endolymph is also at an elevated electrical potential and the driving force is greater. Deflection of hair cell stereocilia by sound (cochlea) or head motion (vestibular system) opens mechanosensitive channels (MET). Potassium ions (K^+) flow through the MET channels into the hair cells and cause hair cell depolarization. The two types of hair cells in the cochlea respond differently to membrane depolarization - inner hair cells release neurotransmitters which signal the auditory nerve, while outer hair cells undergo mechanical deformation which amplifies vibrations of the organ of Corti. In the vestibular system, type I and type II hair cells release neurotransmitters to signal the vestibular nerve; type I cells additionally support non quantal transmission.

Hair cell mechanotransduction process drains K^+ from the endolymph. Sustaining a potassium current into the endolymph to replenish the lost K^+ is an energy intensive process that takes place in the stria vascularis (cochlea) and the dark cell layer (vestibular system). The stria vascularis is a highly vascular multilaminate epithelium composed of two functional cellular layers. The layer facing the scala media is composed of marginal cells that have extensive basolateral infoldings rich in mitochondria and a high density of the Na^+-K^+-ATPase and NKCC1. The apical ends of the cells contain an I_{sk} current composed of subunits of KCNQ1/KCNE1 that returns potassium to the endolymph. A similar arrangement occurs in the dark cells and is shown in Figure 1a. Due to the complexity of the cochlea and vestibule, the biophysics of marginal cell and dark cell ion transport processes are not fully understood but are essential for understanding function and dysfunction in these systems.

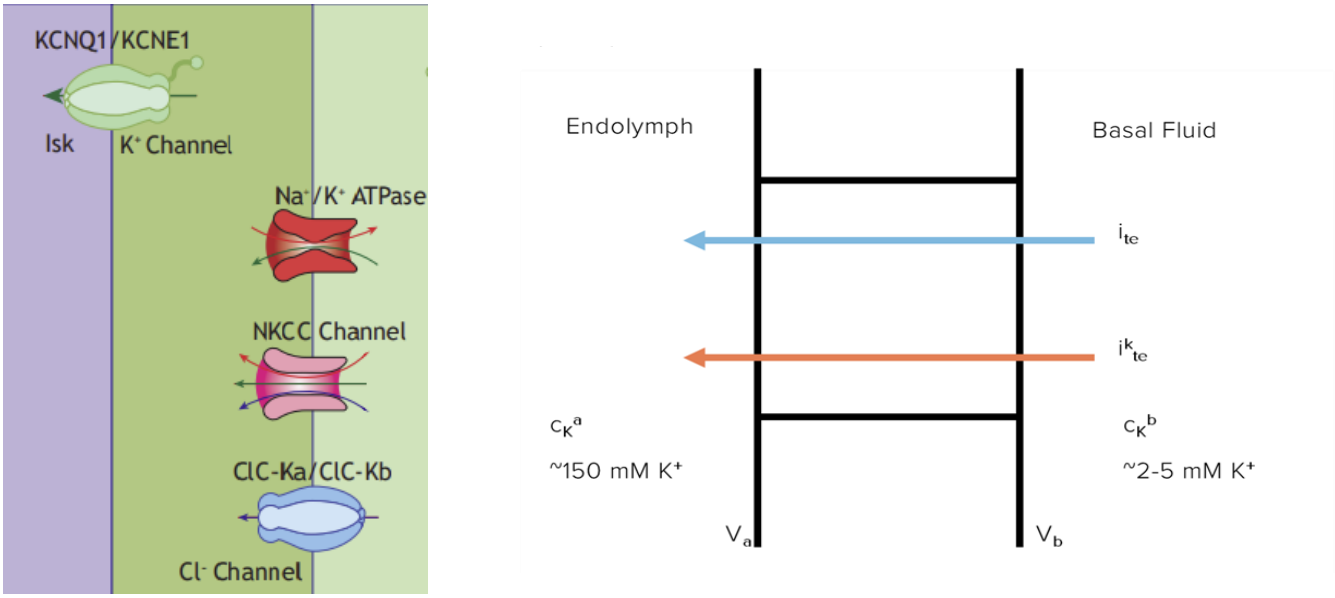


FIGURE 1. a) Major Ion Channel Distribution in Cochlear Marginal Cells and Vestibular Dark Cells. b) Definition of the variables used in the model

Mutations in ion channels expressed in the inner ear can cause hearing loss and vestibular dysfunction. One example is a mutation of the KCNE1 gene which encodes a protein that is a subunit of voltage activated potassium channels expressed in the apical membrane of the marginal/dark cells. This mutation is associated with a greater susceptibility to noise-induced hearing loss [1]. Interestingly, this mutation increases the Isk conductance, but it is unknown how this predisposes individuals to hearing loss.

We have expanded a previously developed computational model of ion transport across the marginal and dark cells [2] to improve accuracy and deliver better insight into the function and dysfunction of the marginal cells and dark cells. The upgraded model is able to simulate the influence of extracellular $[K^+]$ on the transepithelial K^+ current and investigate how genetic mutations in ion channels give rise to deafness and noise-induced hearing loss.

GENERAL DESCRIPTION OF THE MODEL

We present a steady-state model of ion transport across a marginal cell/vestibular dark cell exposed to a constant ionic environment on the apical and basolateral sides. The apical fluid represents the endolymph whose ion concentrations have been measured. In the cochlea, the basal fluid occupies the intrastrial space, in the vestibular system the basal fluid occupies the intercellular spaces surrounding the dark cells. Ion concentrations are difficult to study in these fluids due to their small volumes. The model enables the study of how perturbations in ion concentrations influence ion movement from and to these small spaces. For example, Quraishi and Raphael predicted how the potassium concentration in the basolateral fluid affected the generation of the endocochlear potential [3].

To develop the model, flux expressions that represent the activity of the major channels were utilized. The parameters in the model are taken from the literature as previously described [2]. In the original model, a simplified model of the Na/K ATPase was utilized that did not include ATP dependency. We modified the model of ion transport developed by Quraishi and Raphael [2] to include ATP as variable, implementing Oka, Cha, and Noma's [4] 15-state

post-Albers reaction cycle model for the Na^+-K^+-ATPase. We used parameters for the *α1 β2* isoform of the pump expressed in the marginal cells. Using these formulas in an extended computational model, we were able to derive a value for the rate of the Na^+-K^+-ATPase pump dependent on the concentration of ATP, which was then integrated into the full model. The system of ordinary differential equations was solved in Matlab. The output of the model are the concentrations of Na^+, K^+ and Cl^- inside the marginal cell, the steady state membrane potentials across the basal and apical surfaces of the cell and the net transepithelial currents across the cell.

RESULTS

Validation of the Model. We confirmed that the expanded model reproduced all the previous results obtained in the original model by Quraishi and Raphael [2]. This included predictions for how the density of Isk, Na/K ATPase, and NKCC affect the open and short circuit currents.

Open and Short Circuit Conditions. The predictions from the extended Na^+-K^{+-}ATPase model can be used to study the effect of various parameters on cell function. A standard approach used in epithelial physiology is to evaluate the functional capacity of the system in open/short circuit conditions. The open circuit generates maximal voltage while the short circuit generates maximal current. The physiological behavior of the system will depend on the resistance of the pathway that returns K^+ to the perilymph bathing the basal sides of the marginal/dark cells. Considering the importance of rapid potassium recycling in the cochlea it is likely that the short circuit condition is more representative of the native physiological system but we also show predictions for open circuit voltages.

Energetic Depletion. The expanded model permits an examination of how ATP levels affect the response of the system. The results indicate that the current carried by the Na^+-K^+-ATPase, the K^+ carried by the Na^+-K^+-Cl^- cotransporter (NKCC), the net K^+ current across the epithelium (i_K^{te}), and the net ionic current across the cell (i_{te}) all begin to decline when the ATP concentration in the cell falls below 1 mM (Figure 2). Of particular physiological significance is that the model predicts that i_K^{te} can reverse direction - meaning that potassium will be transported out of the endolymph into the cell. These results have implications for the cochlea and vestibular system to function and survive during episodes of energetic depletion which may occur in response to loud noises that generate reactive oxygen species.

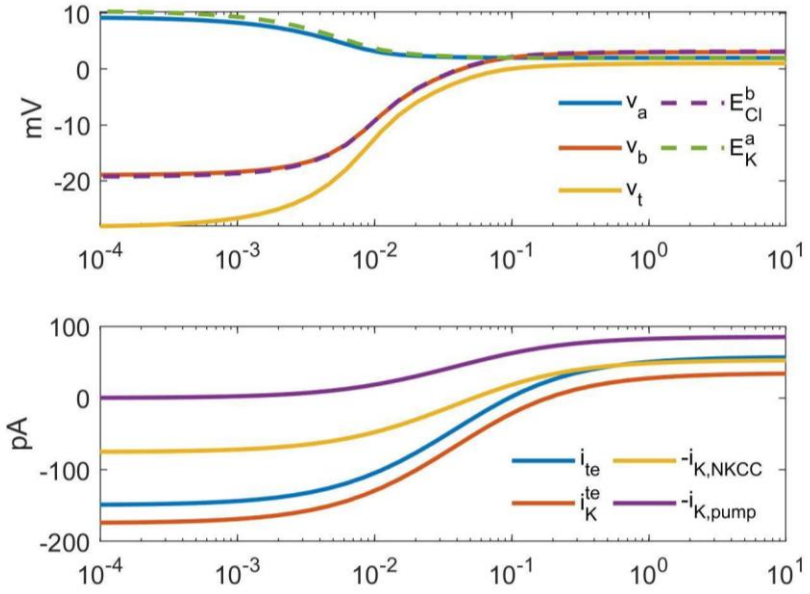


FIGURE 2. Effect of ATP on Potassium Transport Across the Marginal Cell. Legend: $\mathbf{v_a}$ = apical voltage; $\mathbf{v_b}$ basolateral voltage; $\mathbf{v_t}$ = transepithelial voltage difference; $\mathbf{E_{cl}^b}$ = basal Nernst potential for Cl^-; $\mathbf{E_k^a}$ = apical Nernst potential for K^+; $\mathbf{i_{te}}$ = total transepithelial current; $\mathbf{i_K^{te}}$ = transepithelial K^+ current; $\mathbf{-i_{K,NKCC}}$ = K^+ current across NKCC cotransporter; $\mathbf{-i_{K,pump}}$ = K^+ current across Na^+-K^+-ATPase pump. Note that currents are picoamps per cell.

Contour Graphs of Net Potassium Resupply Current (i_K^{te}). The model is able to study any desired parameter as a function of other parameters. For example, given that apical and basal potassium concentrations are major driving forces for potassium transport, a diagram can be made that predicts the current across the marginal cell as a function of the apical and basal potassium concentrations. This can be thought of as a phase diagram that shows how changes in potassium alter the steady state currents across the marginal cell. An example of this plot is shown in Figure 3a. In mathematical terms, the contours in the diagram reflect level curves with various i_K^{te} values. The values shown in the contours are pA/cell. The contour labeled zero (0) reflects the situation in which there is no net current across the epithelia. Below the zero contour, we see that currents are positive, meaning that potassium is transported into the endolymph and above the zero contour, we see that currents are negative, meaning that potassium is transported out of the endolymph. The negative currents are caused by high levels of endolymphatic potassium.

Looking at the predictions in Fig. 3a carefully we see that the dependency of i_K^{te} on external potassium concentrations is nonlinear. Under normal conditions, i_K^{te} increases monotonically for basolateral potassium concentrations up to ~ 7 mM, and then approaches saturation. The system exhibits stable behavior below values of endolymph potassium of ~160 mM for basal potassium concentrations up to 15 mM. However, a surprising finding of this model is that, above 160 mM, additional stable states can appear.

Isk (KCNE1) Mutations. At resting conditions, Isk is poised at approximately 50% of its maximal activity [2, see Fig 4]. This is expected as the marginal/dark cells will have to respond to a range of loads on the stereocilia. Mutations in the genes responsible for Isk conductance have been shown to cause hearing and vestibular dysfunction. For example, a mutation in a potassium subunit (KCNE1) responsible for I_{sk} is associated with susceptibility to noise-induced hearing loss [1]. This mutation increases the conductance of I_{sk}, and it is unclear why this leads to hearing loss. To simulate the effects of this mutation, we increase the conductance by 5x in the model. The results on the phase state of the system are shown in Fig. 3b. As shown, making the apical potassium channel (KCNE1) overactive has little effect on the system under normal conditions of ~ 4mM basolateral K^+. However, increasing I_{sk} conductance results in the appearance of a second steady state and a region of hypersensitivity at high basolateral K^+ concentrations (greater than 10 mM). These effects are not unexpected in dynamical systems where steady state behavior often displays sensitivity to parameter values. The appearance of unstable transport regimes is expected to lead to cochlear dysfunction as they will lead to imbalances in endolymphatic potassium that drives MET channels in hair cells.

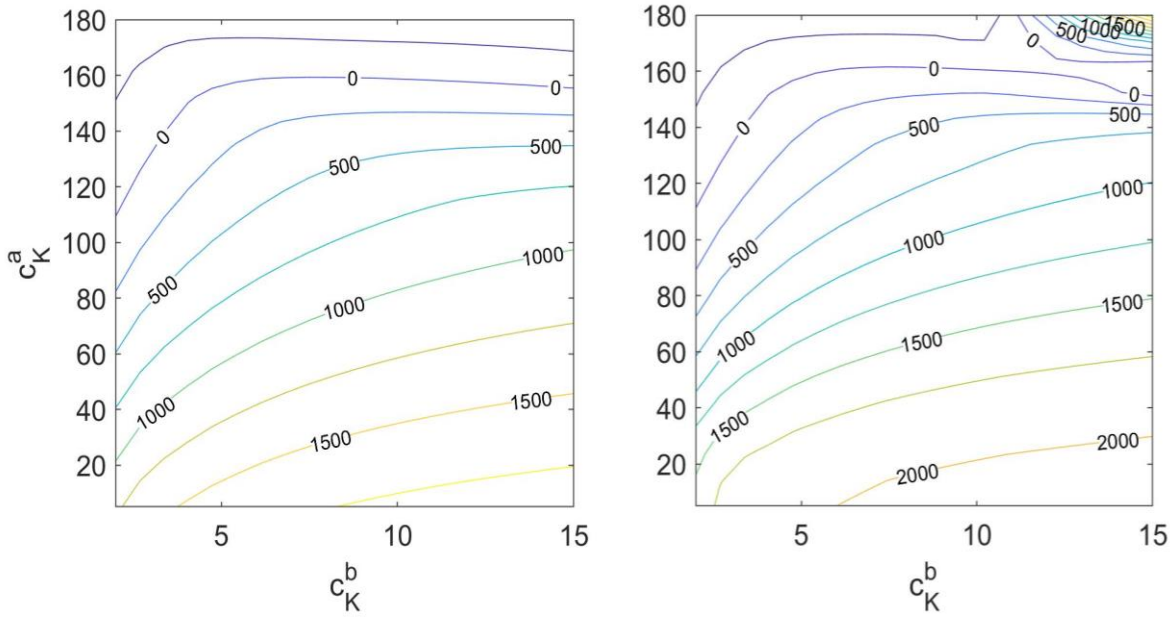


FIGURE 3. Increasing I_{sk} Conductance Causes Unstable Transport Regimes to Develop. The contours on the graphs below show the predicted transepithelial potassium current (i_K^{te}) in pA/cell as a function of the K^+ concentration on the apical and basolateral sides of the marginal/dark cells at a) normal I_{sk} conductances and b) elevated I_{sk} conductances (x5). As shown, making the apical potassium channel (KCNE1) overactive has little effect on the system under normal conditions of ~ 4mm basolateral K+. However, increasing I_{sk} conductance causes a second steady state and region of hypersensitivity to appear at higher basolateral K^+ concentrations.

Effect of ATP Reduction on Potassium Resupply Current. We can also study the effect of reduced ATP on the resupply current i_K^{te} . Energetic depletion was modeled by decreasing the ATP available for the Na+/K+ pump. To visualize the effects of reduced ATP on i_K^{te}, we chose a single contour of the phase diagram that represented i_K^{te} = 50 pA and studied how this contour changed at different ATP levels. ATP levels below 1 mM caused marked shifts in normal behavior of i_K^{te} , lowering the steady state i_K^{te} contours and causing the i_K^{te} current to reverse at normal values of endolymphatic potassium (Figure 4).

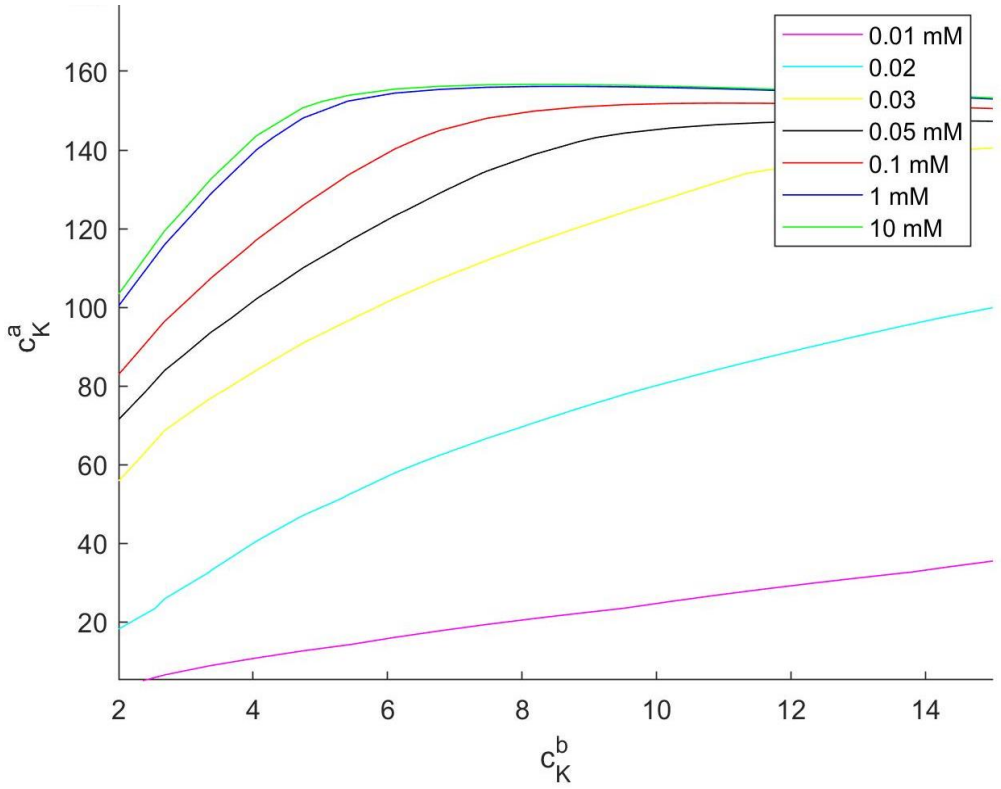


Figure 4. Effects of reducing ATP from control for a transeptihelial current of 50 pA/cell. Below 1 mM ATP, steady state will be established at lower values of endolymphatic K^+. System is most sensitive to ATP depletion at normal value of basolateral K^+ (3-4 mM)

DISCUSSION

This is, to our knowledge, the first model to simulate the effect of ATP reductions on ion transport across the marginal/dark cell layers. This is also the first model to predict the effect of overactivity in the Isk conductance on potassium resupply.

The results indicate that the current carried by the Na^+-K^+-ATPase, the K^+ carried by the Na^+-K^+-Cl^- cotransporter (NKCC) and the net K^+ current across the epithelium (i_K^{te}) all begin to decline when the ATP concentration in the cell falls. Of particular physiological significance is that the model predicts that i_K^{te} can reverses direction, meaning that potassium will be transported out of the endolymph. These results have implications for cochlear and vestibular function and survival during episodes of energetic depletion which may occur in response to loud noises that generate reactive oxygen species.

The model also predicts that even modest increases in Isk (by 5x) can introduce instabilities in K^+ transport when basolateral K^+ is elevated from its normal value of ~2-4 mM. These predictions provide a potential explanation for the genetic susceptibility to noise-induced loss in individuals with the KCNE1 mutation that increases Isk by 5-16x [1] - such individuals will experience dysregulation and instability in ion homeostasis if they experience elevated basolateral potassium.

This extension of the Qurashi and Raphael model [2] lays a foundation for future investigations of how energy depletion and genetic mutations affect the stability of potassium resupply to the endolymph and thereby ionic

homeostasis of the endolymph. The model made the surprising prediction that the marginal/dark cell ion transport system can exhibit multiple steady states and hypersensitivity. This type of behavior is inherent to dynamical systems and has been studied extensively in neuronal systems [5] but to our knowledge has not been reported in epithelial transport systems. In future work the dynamical behavior of the potassium resupply system needs to be examined in greater detail by determining whether the additional equilibrium states are stable or unstable and whether the system exhibits bifurcations or limit cycle oscillations.

The model has been solved at steady state but can be adapted to solve the system when perturbations occur. The model has also been solved in isolation from the other cells and fluids in the tissue, but can be integrated into models of potassium cycling in the entire cochlea [6,7].

CONCLUSION

The incorporation of the 15 state post-Albers model of the Na^+-K^+-ATPase to our previous ion transport model yields a more detailed and biologically accurate model of ion transport in the inner ear. Explicitly including ATP as a variable enables us to simulate the effects of energetic depletion, which is implicated in several inner ear disorders. Our results indicate that mutations affecting channel conductance can lead to instabilities in ion transport processes within marginal cells and vestibular dark cells and leave the system in an irrecoverable, pathologic, transport regime. This insight highlights the utility of computational modeling in analyzing the causes of hearing loss and vestibular dysfunction.

ACKNOWLEDGEMENTS

This work was supported by an NSF IGERT 1250104 and NIH R01 DC012347-08S17. We thank Aravind Govindaraju for assistance.

REFERENCES

1. L Van Laer, PI Carlsson, N Ottschytsch, ML Bondeson, A Konings, A Vandevelde, N Dieltjens, E Fransen, D Snyders, E Borg, A Raes, G Van Camp, The contribution of genes involved in potassium-recycling in the inner ear to noise-induced hearing loss. Hum Mutat. 2006 Aug;27(8):786-95. doi: 10.1002/humu.20360. PMID: 16823764.
2. IH Quraishi and RM Raphael, Computational model of vectorial potassium transport by cochlear marginal cells and vestibular dark cells. Am J Physiol Cell Physiol. 2007 Jan;292(1):C591-602. doi: 10.1152/ajpcell.00560.2005.
3. IH Quraishi and RM Raphael, Generation of the endocochlear potential: a biophysical model. Biophys J. 2008 Apr 15;94(8):L64-6. doi: 10.1529/biophysj.107.128082.
4. C Oka, CY Cha, A Noma, Characterization of the cardiac Na+/K+ pump by development of a comprehensive and mechanistic model. J Theor Biol. 2010 Jul 7;265(1):68-77. doi: 10.1016/j.jtbi.2010.04.028.
5. Eugene Izhikevich. *Dynamical Systems in Neuroscience*. (MIT Press, Boston, 2010).
6. IH Quraishi, "Computational Models of Potassium Transport in the Inner Ear," PhD thesis, Rice University, 2008
7. F Nin , H Hibino, S Murakami, T Suzuki, Y Hisa, Y Kurachi, Computational model of a circulation current that controls electrochemical properties in the mammalian cochlea. Proc Natl Acad Sci U S A. 2012 Jun 5;109(23):9191-6. doi: 10.1073/pnas.1120067109.

The Inner Spiral Sulcus and Outer Tunnel Alter Inner-Hair-Cell Bundle Motion in a Finite-Element Model of the Mouse Cochlea

Yanli Wang[1, 2] and Sunil Puria1[1, 2, 3, a)]

[1]*Dept. of Otolaryngology, Harvard Medical School, Boston, MA.*
[2]*Eaton-Peabody Laboratories, Massachusetts Eye & Ear, Boston, MA.*
[3]*Graduate Program in Speech and Hearing Bioscience and Technology, Harvard University, Cambridge, MA.*

[a)] *sunil_puria@meei.harvard.edu*

Abstract. The inner hair cells (IHCs) situated within the organ of Corti (OoC) inside the cochlea translate mechanical sound vibrations into electrical potentials via mechanoelectrical transduction (MET) channels within the IHC stereocilia bundles. Exactly how the IHC bundles are stimulated by their environment within the OoC remains unclear. The current work utilizes a 3D finite-element (FE) model of a slice from the middle-turn 20-kHz region of a mouse cochlea to study the passive modes of vibration within the OoC at different frequencies and how they influence IHC-bundle stimulation. The results indicate that the OoC exbibits different modes of motion at different frequencies, and that the IHC-bundle stimulation magnitude peaks at the best frequency (BF). The results also indicate that the inner spiral sulcus and outer tunnel of Corti can both have a significant effect on the IHC-bundle stimulation at BF.

INTRODUCTION

The intricate cellular structure within the organ of Corti (OoC) provides the structural basis for how the stereocilia bundles of the inner hair cells (IHCs) are stimulated, which triggers the mechanoelectrical transduction (MET) channels of individual IHC stereocilia. The stimulation mechanisms of the IHC bundles play a crucial role in how the mechanical motions of the traveling wave and the OoC get translated into electrical signals in the hair cells via MET channels. However, the mechanisms by which the IHC bundles are stimulated by their surroundings and the fluid spaces within the OoC remain to be determined[1]. Recent experimental data measured on the individual stereocilia within *in situ* IHC bundles reveal that the intra-bundle stereocilia do not move together as a cohesive whole[2]. This finding provides a motivation for employing modeling studies to better understand the stimulation mechanisms for the IHC stereocilia in terms of the surrounding anatomy and fluid environment.

The present work studies the cytoarchitectural motion of the OoC and the resulting stimulation of the IHC bundles via a 3D finite-element (FE) model of the mouse OoC in the middle turn. Based on recent data, the two key hypotheses tested by the model are: 1) whether the OoC exhibits different modes of concerted motion at different frequencies that lead to different stimulation modes for the IHC bundles, and 2) whether the inner spiral sulcus (ISS) and outer tunnel (OT) of the OoC fluid spaces[3] (Fig. 1A) have a significant influence on IHC-bundle stimulation.

An FE model of a cross section of a passive mouse cochlea from the middle-turn 20-kHz best-frequency (BF) region has been constructed with details relevant to the mechanical and fluid responses pertaining to OoC motion. The modes of OoC motion at different frequencies are studied, and the IHC-bundle motions are compared against experimental data. Furthermore, model versions with geometric variations to the ISS and OT have also been created for testing the structural significance of the two fluid spaces to OoC motion.

METHODS

Geometry

The realistic representation of the detailed cellular structure in the FE model is crucial for studying IHC-bundle stimulation. The model was constructed in FreeCAD based on two-photon[4] and bright-field[2] microscopy measurements on fresh tissue, as well as scanning electron microscopy[5] and histological slices on fixed tissue (EPL otopathology). The model captures the structurally important features, including the mosaiclike pattern of the reticular lamina (RL; Figure 1B), and realistic representation of the fluid spaces of the ISS and cortilymph, which consists of the inner tunnel of Corti (IT), the space of Nuel (SN), and the outer tunnel (OT; Fig. 1A). It also captures the longitudinal inclinations of the outer hair cells (OHCs), Deiters' cells (DCs), phalangeal processes (PhPs), and outer pillar cells (OPCs; Fig. 1C and D), with the RL consisting of a mosaic of the tops of the OHCs, PhPs, and OPCs (Fig. 1B). The lengths, diameters, and longitudinal and radial inclinations of the OHCs, DCs, inner and outer pillar cells, and inner and outer hair bundles are taken from Soons et al. (2015)[4]. The dimensions of the peripheral structures such as the tectorial membrane (TM), Hensen's cells, stria vascularis, inner spiral lamina, spiral ligament, and inner supporting cells are taken from histological slices. The bundle heights are estimated according to multiple resources[2, 4, 6, 7, 8]. The tips of the OHC bundles are embeded into the TM. Despite an ongoing debate[9], the IHC bundles are modelled as not touching the TM, with a fluid gap connecting the sub-tectorial space (Fig. 1E) and ISS. For the aims of the current study, the stereocilia bundles are simplified as not consisting of individual stereocilia, but instead as a V shape and curvy wall for the OHC and IHC bundles, respectively (Fig. 1B–D). To mimic the thinner diameter at the root of the stereocilia, the bottoms of the bundle shapes are assigned much softer material properties to allow bending there. Variations in the sizes of the ISS and OT are made to study the roles of these fluid spaces in OoC cytoarchitectural and IHC-bundle motions.

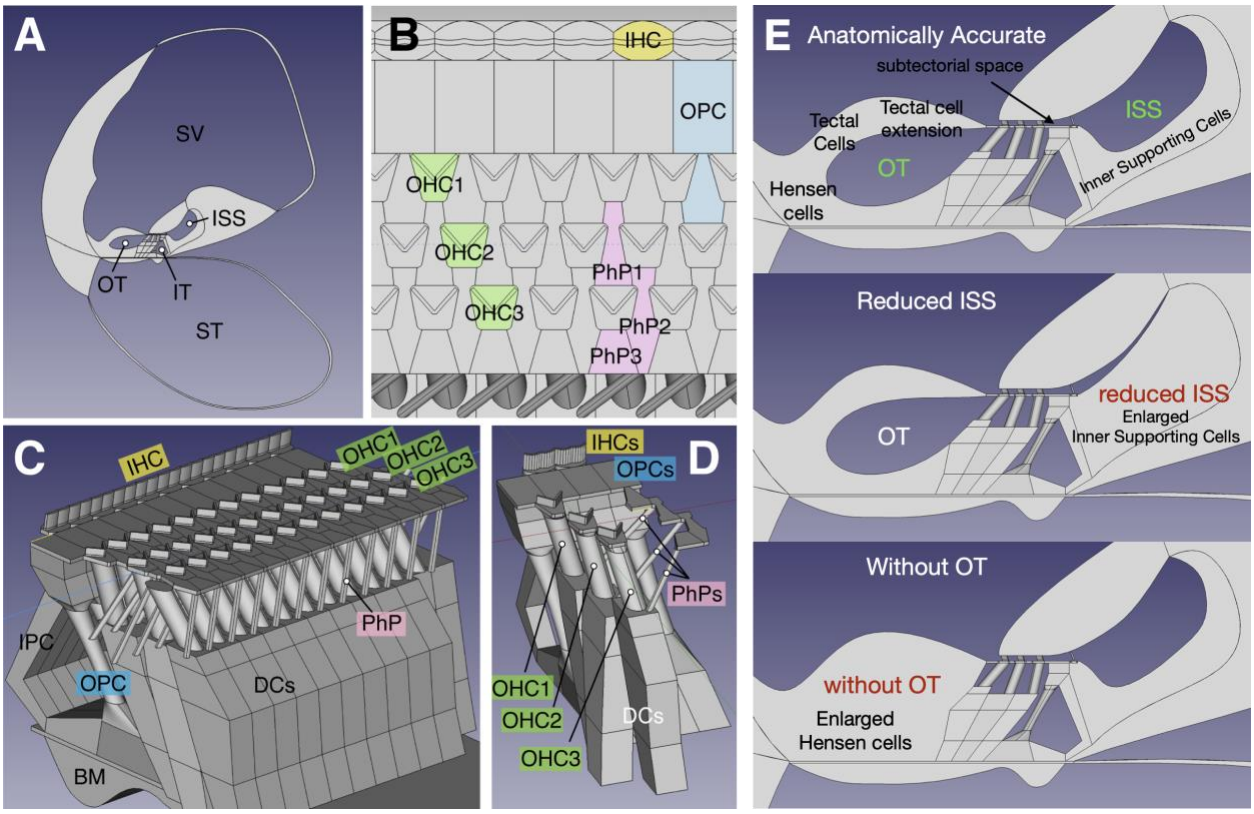


FIGURE 1. The geometry and components of the orgain of Corti (OoC) finite-element (FE) model. (A) The fluid spaces around the OoC. (B) The mosaic pattern forming the reticular-lamina (RL) surface, composed of the upper surfaces of the outer hair cells (OHCs), phalangeal processes (PhPs), and outer pillar cells (OPCs). This view displays a 46.2-µm-wide longitudinal section

(horizontal in this view). Note that the presently reported model calculations are based on a smaller model that is 23.1-µm wide. (C) Another view of the model, showing how the OHCs, DCs, PhPs, OPCs, and inner pillar cells (IPCs) form the core of the OoC as built upon the basilar membrane (BM). This view displays an approximately 92-µm-wide longitudinal section. (D) One isolated set of OHCs, DCs, and PhPs, from rows 1–3, showing the longitudinal inclinations of these cells. (E) Side views of the OoC for the anatomically accurate model (top), for the inaccurate model with a reduced inner spiral sulcus (ISS; middle), and for the inaccurate model without an outer tunnel of Corti (OT; bottom).

Finite-Element Modeling

The geometry was constructed in FreeCAD and then imported into COMSOL Multiphysics (ver 5.3a), which is used for the FE modeling. The slice used for the current model calculations is 23.1 µm wide, consisting of 3 sets of IHCs and OHCs, etc. (as a reference, Fig. 1B shows a section that is twice as wide as the slice used for the current results). Uniform pressure is applied to the bottom of the basilar membrane (BM), while continuity boundary conditions are applied to the two longitudinal-facing sides of the slice. The uniform-pressure and continuity boundary conditions mimic the condition in which the wavelength is much longer than the width of the slice, which would be the case when the 20-kHz BF location is within the long-wave region during low-frequency stimulation (<10 kHz). It is also the condition used for the experimental data in Wang et al. (2021)[2], where two holes were made in the cochlea to allow IHC-bundle imaging at the 20-kHz BF location during 2- and 3-kHz stimulations. The stimulation frequencies applied in the simulations are 2, 3, 5, 10, 20, 30, and 40 kHz. Note that in our experimental setup, only data for 2- and 3-kHz stimulations were measured due to experimental limitations[2]. Moreover, although the results with 20-, 30-, and 40-kHz stimulations are still illuminating in the current study, the present continuity boundary conditions are not well-suited to represent conditions with a shorter-wavelength traveling wave, which may be the case for these frequencies, such that Floquet boundary conditions should be used instead[10]. Please see the DISCUSSIONS section for more details.

The material properties of the various structures are directly taken from previous FE models of mouse and gerbil organs of Corti[10, 11]. Notably, both the Hensen's cells and inner supporting cells that respectively surround the OT and ISS are composed of soft ground substance (60 kPa), as compared to the stiffer BM (200 MPa).

RESULTS

Different modes of OoC motion are apparent when the stimulation frequency varies. Each of these modes has a different influence on IHC-bundle stimulation. In this section, different modes are shown (Fig. 2). The effects of these different modes on IHC-bundle stimulation are then studied by looking at the IHC-bundle deflections normalized to the cuticular-plate motion of the IHC (Table 1). The magnitude of the cuticular-plate motion and the bundle deflections are shown as well (Tables 2 and 3). These results are also compared to the results from the anatomically inaccurate models with a reduced ISS and omitted OT (Tables 2 and 3).

OoC Motion Modes at Different Stimulation Frequencies

To show the modes of OoC motion at different stimulation frequencies, color is mapped to the displacement amplitudes across the OoC in Figure 2, with 2, 20, and 40 kHz shown in the left, middle, and right columns, respectively. The transverse and radial amplitudes are shown in the first and second rows, respectively. By comparing the patterns of coloring, it can be easily seen that different modes of motion are displayed for the different stimulation frequencies. To show the motions more clearly, the scales of the different colormaps are not the same across the panels. All of the results shown are for a driving pressure of 1 Pa applied to the bottom of the BM. As mentioned in the METHODS section, these conditions are not intended to mimic the conditions with the presence of a traveling wave for two reasons: 1) in a passive or active cochlea with a traveling wave, the pressure at a given location will change as the stimulation frequency changes, 2) at high frequencies and frequencies close to the 20-kHz BF, Floquet boundary conditions are required for capturing traveling waves with a wavelength close to the width of the slice. Nevertheless, these simulations indicate that different modes of OoC motion are to be expected with different stimulation frequencies.

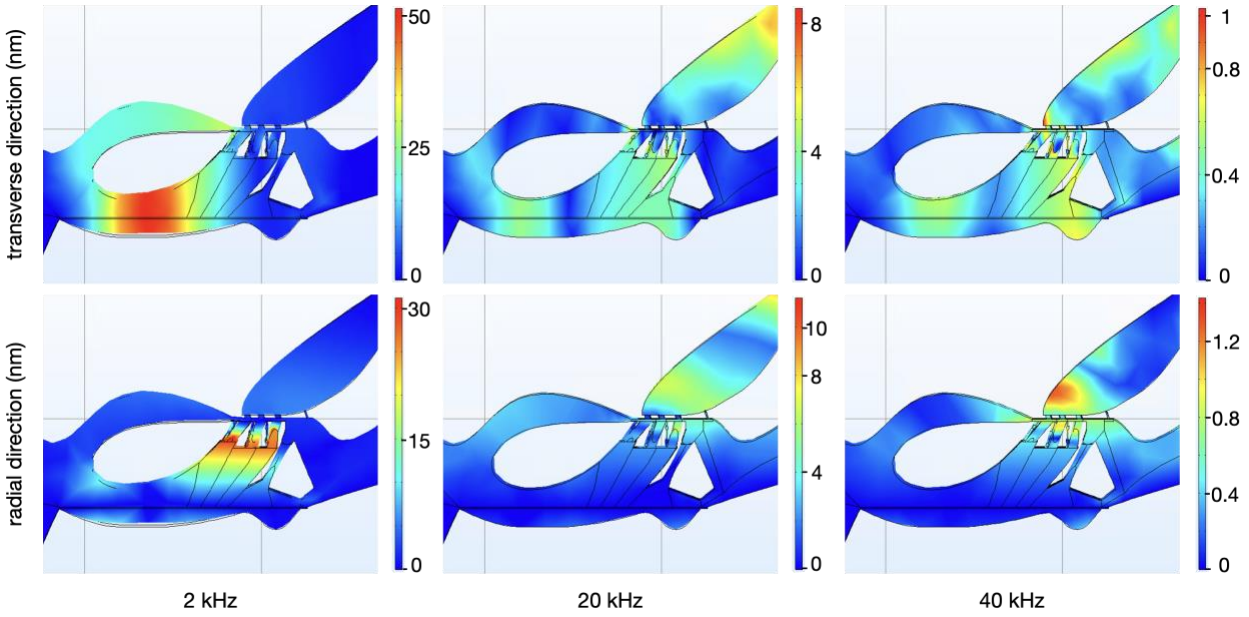


FIGURE 2. Displacement amplitudes (μm) mapped to color, showing different modes of OoC motion at different stimulation frequencies. (A–C) Displacement amplitudes in the transverse (vertical) direction. (D–F) Displacement amplitudes in the radial (horizontal) direction. (A, D), (B, E), and (C, F) show results for 2-, 20-, and 40-kHz stimulations, respectively.

Comparisons of IHC-Bundle Deflections Normalized by the Magnitude of the IHC Cuticular-Plate Motion, from Experiments and Anatomically Accurate and Inaccurate Models

The IHC-bundle deflection is calculated as the differential radial-direction motion between the top of the bundle and the bottom of the bundle (the cuticular-plate motion). This differential motion is also called the relative bundle motion re. the cuticular plate in previous work[2]. For this section, the deflection magnitude is normalized by the magnitude of the cuticular-plate motion to compare against the experimental data, and to indicate the efficacy of bundle stimulation for a given cuticular-plate motion magnitude. The unnormalized deflections and cuticular-plate motions are reported in the following section.

Table 1 lists and compares the results from experimental data and those from anatomically accurate and inaccurate models. The experimental data are only available for 2- and 3-kHz stimulations, whereas the model simulations went up to 40 kHz for the 20-kHz BF location. For the anatomically accurate model, the results at 2 and 3 kHz are close to the ranges of expected results based on experimental data, but lie on the high ends for both 2 and 3 kHz. There is also a peak of stimulation efficacy at 20 kHz for the anatomically accurate model, but not for the inaccurate model with a reduced ISS. The peak is still present in the model without an OT, but is reduced.

TABLE 1. Magnitudes of the IHC-bundle deflection (relative motion) normalized by the magnitude of the IHC cuticular-plate motion.

	Experiment	**Model**		
Frequency	**Wang et al. 2021[2] Fig. 4**	**Accurate**	**reduced ISS**	**without OT**
2 kHz	0.63±0.16	0.90	0.44	0.09
3 kHz	0.48±0.12	0.59	0.34	0.20
5 kHz	N/A	0.63	0.28	0.45
10 kHz	N/A	0.85	1.36	1.08
20 kHz (BF)	**N/A**	**2.44**	**0.98**	**1.64**
30 kHz	N/A	1.07	1.28	0.80
40 kHz	N/A	1.4	4.84	8.28

Unnormalized Magnitudes of the Radial IHC Cuticular-Plate Motions and IHC-Bundle Deflections, for Anatomically Accurate and Inaccurate Models

Table 2 lists and compares the unnormalized radial-motion magnitudes of the IHC cuticular plate, for the anatomically accurate and inaccurate models. The listed magnitudes are for a 20-Pa pressure load on the BM, which is equivalent to 120 dB SPL at the BM. This corresponds roughly to 114 dB SPL at the ear canal for frequencies below 5 kHz, and 90 dB SPL at the ear canal for frequencies above 5 kHz, based on an estimated 6 dB of middle-ear gain below 5 kHz and 30 dB of middle-ear gain above 5 kHz[12]. These estimated SPLs at the ear canal are on the high side, but are considered reasonable for detecting motions of the BM in a passive cochlea. Due to our lack of knowledge about how the pressure acted on the BM at the experimental location during experiments, we are not able to compare these results against experiments. Nevertheless, the trend of the results is illuminating. For the anatomically accurate model, there is a peak at BF. In contrast, for the model with a reduced ISS, the response drops sharply with frequencies above 5 kHz. For the model without an OT, the response drops sharply with frequencies above 20 kHz.

TABLE 2. Magnitudes of the radial IHC cuticular-plate motion of the models (20-Pa pressure load on the BM).

Frequency	**Anatomically Accurate**	**reduced ISS**	**without OT**
2 kHz	44.9 nm	29.4 nm	26.8 nm
3 kHz	63.3 nm	35.9 nm	27.0 nm
5 kHz	52.8 nm	32.2 nm	27.4 nm
10 kHz	4.7 nm	6.0 nm	35.0 nm
20 kHz (BF)	**91.2 nm**	**4.2 nm**	**26.6 nm**
30 kHz	11.1 nm	13.2 nm	2.16 nm
40 kHz	14.2 nm	3.9 nm	0.41 nm

Table 3 lists and compares the unnormalized radial-deflection magnitudes of the IHC bundles, for the anatomically accurate and inaccurate models. Again, the listed magnitudes are for a 20-Pa pressure load on the BM. For the anatomically accurate model, there is a significant peak at the BF location as a compounded effect of the higher cuticular-plate motion at the BF (Table 2) and the higher stimulation efficacy at BF (Table 1). There is no peak at the BF location for the model with a reduced ISS, and a lesser peak for the model without an OT.

TABLE 3. Magnitudes of the radial IHC-bundle deflections of the models (20-Pa pressure load on the BM).

Frequency	**Anatomically Accurate**	**reduced ISS**	**without OT**
2 kHz	39.6 nm	15.7 nm	2.30 nm
3 kHz	37.6 nm	11.0 nm	5.36 nm
5 kHz	33.4 nm	7.63 nm	12.4 nm
10 kHz	3.98 nm	8.16 nm	37.7 nm
20 kHz (BF)	**222.8 nm**	**4.13 nm**	**43.5 nm**
30 kHz	11.8 nm	16.9 nm	1.73 nm
40 kHz	20.0 nm	18.7 nm	3.43 nm

DISCUSSIONS

An anatomically detailed FE model of a slice of the mouse cochlea from the middle turn (20-kHz BF location) has been constructed. The model captures the mechanical and fluid structures relevant to deflections of the IHC and OHC stereocilia bundles. The material properties were taken directly from previous studies[10, 11]. The model results indicate different modes of OoC motion at different stimulation frequencies, and show a significant increase in the IHC-bundle motion magnitude when stimulated at BF. The variations to the geometry of the fluid spaces within the OoC also indicate the importance of the ISS and OT in the frequency responses, due to changes in the stimulation modes within the OoC and therefore the mechanisms of hair-bundle stimulation.

The results from the two anatomically inaccurate models cannot be explained by a simple increase in the mass or stiffness of the system due to the corresponding increased volume of the inner supporting cells and Hensen's cells. With the traditional view that the radial motion of the OoC is directly linked to the transverse motion of BM, the

stiffness in the system for the transverse direction is dominated by the stiffness of the BM as well as that of the DCs, OHCs, OPCs and IPCs. Both the inner supporting cells and Hensen's cells are modeled as soft ground substance having a stiffness more than 3 orders of magnitude softer than that of the BM, and 2–3 orders of magnitude softer than that of the DCs, OHCs, IPCs, and OPCs. Thus, the stiffness increase due to the increased volume of the inner supporting cells and Hensen's cells would have a negligible influence on the BM transverse motion. The present results offer a new perspective on the different modes of vibration of the intricate structures within the OoC, and suggest that the radial motions of the various parts of the OoC, including the manner by which the IHC bundles are stimulated, are predominantly determined by these vibration modes.

The preliminary results from altered models with a reduced ISS show that the presence of the ISS space is highly responsible for the peak in the IHC-bundle stimulation efficacy at the BF (20 kHz). Without an OT, this peak is also reduced at the BF. One possible explanation is that the IHC bundles are influenced by fluid motion from the ISS, and that this fluid motion is altered in the model with reduced ISS volume. More studies can be done by observing the flow of fluid in the different model versions, and also by varying the size of the ISS to see whether that causes the peak of stimulation efficacy for the IHC bundles to shift in frequency.

The absence of an OT generally causes a significant reduction of the peak and a broadening of the peak in the radial motion of the IHC cuticular plate. We hypothesize that these reductions could also be seen in the OHC region of the RL, and that the fluid motion in the OT, together with the motion of the tectal-cell extension (Fig. 1E top panel), i.e., the thin piece of tissue connecting the Hensen's cells (Fig. 1E top panel) to the RL, induce more radial motion of the RL. The tectal cells and nearby Hensen's cells are sometimes called the "lateral compartment", which has been shown to exhibit significant motion relative to the BM motion in both active and passive cochleae[13], which is consistent with the present simulations.

We have chosen to model the 20-kHz region because of the availability of experimental data for IHC cuticular-plate and hair-bundle motions from this location. We are now developing a similar model for the apical 10-kHz BF region, as this will allow validation of the model against 2D experimental data available for this region[13, 14]. There are significant geometrical differences in the OoC between the apical and middle-turn regions. These differences include the inclination angles of OHCs, DCs, OPCs, and IPCs, and different shapes and sizes of the ISS and OT. The comparison of cytoarchitectural motions between apical and middle-turn OoC locations will help to determine the structural implications and importance of the various geometrical parameters in the OoC.

In summary, our approach to FE modeling of a section of the mouse cochlea offers a unique opportunity to study the detailed concerted motion within the OoC at different frequencies, and the structural importance of various components to the stimulation mechanisms for the IHC bundles. Preliminary results find a significant peak in the stimulation magnitude of the IHC bundles for a 20-kHz stimulation, which corresponds to the BF for the slice location. Results from anatomically inaccurate models indicate that the ISS is responsible for this peak, and that the presence of the OT generally increases the radial motion of the RL at the IHC location.

ACKNOWLEDGMENTS

This work is supported by grant R01 DC007910 from the NIDCD of the NIH. We also thank the generous help offered by Andrew Tubelli on working with FreeCAD and COMSOL Multiphysics, as well as by Hamid Motallebazadeh for his expertise and experience in working with COMSOL Multiphysics.

REFERENCES

1. J. J. Guinan, Jr., Hear Res **292** (1-2), 35-50 (2012).
2. Y. Wang, C. R. Steele, S. Puria and A. J. Ricci, Commun Biol **4** (1), 958 (2021).
3. N. H. Cho, H. Wang and S. Puria, J Assoc Res Otolaryngol (2022).
4. J. A. Soons, A. J. Ricci, C. R. Steele and S. Puria, J Assoc Res Otolaryngol **16** (1), 47-66 (2015).
5. R. R. Taylor, D. J. Jagger and A. Forge, PLoS One **7** (1), e30577 (2012).

6. V. Zampini, L. Ruttiger, S. L. Johnson, C. Franz, D. N. Furness, J. Waldhaus, H. Xiong, C. M. Hackney, M. C. Holley, N. Offenhauser, P. P. Di Fiore, M. Knipper, S. Masetto and W. Marcotti, PLoS Biol **9** (4), e1001048 (2011).
7. X. Wu, M. V. Ivanchenko, H. Al Jandal, M. Cicconet, A. A. Indzhykulian and D. P. Corey, Nat Commun **10** (1), 3801 (2019).
8. K. K. Miller, P. Atkinson, K. R. Mendoza, O. M. D and N. Grillet, Front Cell Dev Biol **9**, 742529 (2021).
9. P. Hakizimana and A. Fridberger, Nat Commun **12** (1), 2604 (2021).
10. A. Tubelli, H. Motallebzadeh, J. J. Guinan, Jr. and S. Puria, Bio Phys Journal **(under review)** (2022).
11. H. Motallebzadeh, J. A. M. Soons and S. Puria, Proc Natl Acad Sci U S A **115** (22), 5762-5767 (2018).
12. H. Motallebzadeh and S. Puria, J Acoust Soc Am **(in revision)** (2022).
13. S. S. Gao, R. Wang, P. D. Raphael, Y. Moayedi, A. K. Groves, J. Zuo, B. E. Applegate and J. S. Oghalai, J Neurophysiol **112** (5), 1192-1204 (2014).
14. H. Y. Lee, P. D. Raphael, A. Xia, J. Kim, N. Grillet, B. E. Applegate, A. K. Ellerbee Bowden and J. S. Oghalai, J Neurosci **36** (31), 8160-8173 (2016).

The Drive to Inner and Outer Hair-Cell Bundles in a Slice Model of the Gerbil Cochlea

Andrew Tubelli[1], Hamid Motallebzadeh[1,2], Michelle Qin[3], John Guinan[1,2], Sunil Puria[1,2,a]

[1] Eaton-Peabody Laboratories, Mass. Eye and Ear, 243 Charles St., Boston MA 02114, USA.
[2] Harvard Medical School, Dept. of Otolaryngology Head and Neck Surgery, Boston MA, USA.
[3] Harvard University, Boston MA, USA.

[a)] Corresponding author: sunil_puria@meei.harvard.edu

Abstract. Despite the importance of inner hair cells (IHCs) and outer hair cells (OHCs) to cochlear physiology, we have an incomplete understanding of the motions that stimulate these hair cells. Among the unanswered questions are: What is the phase relationship between stereocilia deflections of the three rows of OHCs and the IHCs? How does basilar-membrane (BM) transverse motion translate into radial drive to OHC and IHC stereocilia? Gross motions of the organ of Corti (OoC) in the gerbil apex have been experimentally shown to have low-pass characteristics, yet neural-fiber responses are bandpass. Does the mechanical drive to the IHC stereocilia differ from low-pass and can it explain the difference between mechanical and neural responses? To begin to address these questions, we developed a finite-element model of a 20-μm slice of the passive gerbil cochlea at the middle turn. The model used realistic anatomy of the organ of Corti with stereocilia bundles represented by the series of stereocilia in the tallest row. OHC stereocilia spanned the subtectorial gap, but IHC stereocilia attached only at the reticular lamina. The model included viscous fluid-structure interaction. Material properties of the model elements that were not described in literature were estimated, aided by parameter tuning to emulate key characteristics of experimental data. The cochlear traveling wave was simulated with a Floquet boundary condition that constrained the phase differences between the slice edges to be consistent with the local frequency-dependent wavelength of the traveling wave. Input pressure was applied to the BM on the scala tympani side. The model showed that OHC stereocilia deflection was low-pass, similar to experimentally measured gross OoC motion, but IHC stereocilia deflection was band-pass. At low frequencies, the RL stretched and caused the bases of OHC stereocilia to move predominantly in the radial direction, with more stretching and more stereocilia deflection around the pillar cells and OHC row 1. At frequencies approaching the 2.5 kHz best frequency and higher, the motion was more complex with substantial longitudinal components of motion. OHC stereocilia deflections were generally in phase with each other, and at high frequencies were in phase with IHC stereocilia deflections, but at low frequencies OHC and IHC deflections were out of phase.

INTRODUCTION

The cochlea is a complex structure with many parts whose roles are poorly understood. To make models that are mathematically tractable it has been necessary to simplify, and sometimes use lumped-parameter representations of parts of the organ of Corti (OoC) that are, at best, crude approximations of the mechanical properties of the anatomy. To get a realistic assessment of the important OoC mechanical properties, a model must be realistic. Only with realistic models can we hope to answer such questions as how the mechanical drive to the outer-hair-cell (OHC) stereocilia achieves the required phase to produce cochlear amplification at frequencies within ~1/2 octave below the best frequency (Dong and Olson, 2013).

It is essential to choose a modeling method that is not built upon pre-conceived views of how the OoC functions. With this consideration in mind, we chose a finite-element (FE) modeling approach that used reasonably accurate mechanical anatomy, although we have been limited by a lack of detailed knowledge about many of the structurally important cochlear elements. FE models are computationally expensive, so in this model we included only a 20-μm longitudinal extent of the cochlea, i.e., a cochlear slice. To model the effects of the slice being driven by a traveling

wave, we used a Floquet boundary condition on the longitudinally facing sides of the slice, and a pressure drive on the basilar membrane (BM) that mimicked the pressure difference across the cochlear partition of the traveling wave. We chose the gerbil as the model species because gerbils are a popular animal model for cochlear studies; they are conveniently small but have a frequency range of hearing that extends to the low frequencies that are important for human speech.

METHODS

We based the model on the Edge et al. (1998) image of the middle turn of the gerbil hemicochlea with a characteristic frequency (CF) of ~2.5 kHz (Muller, 1996). The model attempted to accurately portray the gerbil BM and had a fully arched pectinate zone (PZ) and a smaller arch in the arcuate zone (AZ) (Fig. 1a). In the PZ, the upper collagen-fiber layer was a flat shell, while the lower shell layer formed an arch, and there was ground substance between the shells. The AZ was similar with the two fiber layers closer together. Since the slice thickness was 20 µm, the attachment of each Deiters-cell (DC) phalangeal process (PhP) was just apical and lateral to its OHC (a more-accurate attachment at three-OHCs apical would require a thicker slice and more computing time). We included all of the OoC cells thought to be structurally important, and used dimensions based on anatomy. The hair-cell cuticular plates, the tops of the pillar cells and the PhP tops formed the reticular lamina (RL) which was made to be a 0.8-µm-thick homogeneous region (Fig. 1b). The Hensen's and Claudius' cells, extending to the spiral ligament, were treated as a single homogeneous region (Fig. 1a).

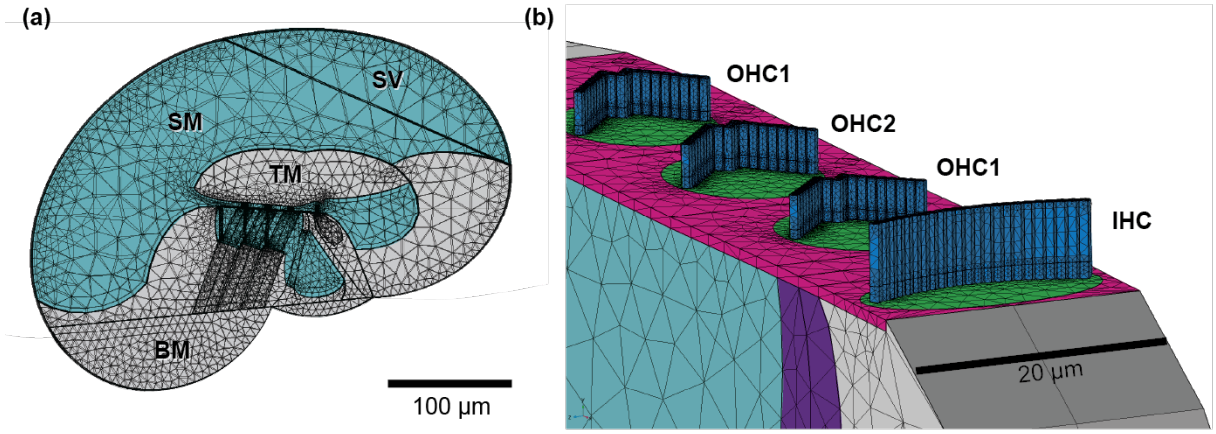


FIGURE 1. Model geometry: (a): The geometry and finite-element meshing for a slice location in the gerbil middle turn with a characteristic frequency (CF) of 2.5 kHz. (b): An enlarged view that highlights anatomical details. BM=basilar membrane; IHC=inner hair cell; OHC=outer hair cell; SM=scala media; SV scala vestibuli; TM=tectorial membrane.

The stereocilia were solid elements, and like the rest of the model, had fluid-structure interactions. The three "rows" of stereocilia in a bundle were modeled just by the tallest stereocilium row, which was formed by a single row of abutting cylinders (Fig. 1b) with appropriate stiffness but no nonlinearity. The OHC stereocilia spanned the subtectorial gap but did not span the entire 20-µm slice thickness (Fig. 1b). The inner-hair-cell (IHC) stereocilia were attached to the cuticular plate but were freestanding in fluid at the top and did not span the entire 20-µm slice thickness (Fig. 1b). To compensate for the lack of stereocilia tapering and rootlets, the base (lower ~20%) of each stereocilium was a separate material with a 0.1 times lower Young's modulus to allow greater bending near the rootlet. The dimensions of hair cells and stereocilia were taken from anatomical images.

Scala media and scala vestibuli were combined to a single fluid domain that was continuous with the fluid in the reticular-lamina (RL) tectorial-membrane (TM) gap and the fluid in the inner sulcus. Inside the OoC was the cortilymph space, which was composed of the outer tunnel, the fluid surrounding the OHCs, the space of Nuel and (connected via gaps between the outer pillar cells) the inner tunnel of Corti. The fluid domains and soft-tissue regions (i.e., everything except the DCs and hair cells) extended throughout the 20-µm slice thickness.

Cellular structure (*e.g.*, the presence and orientation of collagen fibers) and stiffness estimates from the literature were used to set initial values of material properties and to set expected limits of values for the Young's moduli (Slepecky, 1996; Tolomeo et al., 1996; Tolomeo and Holley, 1997; Sellon et al., 2015). These values were tuned (i.e.,

iteratively varied) so that the resulting model had a radial profile of BM motion that was largest near the AZ-PZ junction, as expected from experimental data (Cooper, 2000; Rhode and Recio, 2000; Homer *et al*., 2004; Cooper and van der Heijden, 2018).

Structural components were treated as elastic and isotropic, except the BM which was orthotropic. The BM orthotropy used a local normal-tangential coordinate system parallel to the outer shell of the BM with the tangential direction closest to radial as the stiffest and the other directions an order of magnitude less stiff. The ground-substance gel between the two collagen layers was solid. Thicknesses of the fiber layers were taken from the ~2.5-kHz place (Müller, 1996; Schweitzer *et al*., 1996). We used a Young's modulus of 200 MPa for the collagen layers and 100 kPa for the BM ground substance. The shear modulus for the BM was one third of the respective Young's modulus in each direction. The densities of soft tissues were 1,100 kg/m^3, the Poisson's ratios of all solid materials were 0.485, and the structural damping was 10% (Motallebzadeh et al., 2018). Fluid was approximated as water (ten Kate and Kuiper, 1970). Electrical and/or active processes were not modeled.

Physics and boundary conditions

The model was implemented in COMSOL Multiphysics version 5.3a (COMSOL AB, Stockholm, Sweden) with solid, shell and fluid domains. OoC structures, except the BM outer layer, were modeled with solid mechanics. The BM was modeled using linear-elastic shell elements with three translational and three rotational displacements. The outermost radial edges of the BM, the medial face of the sulcus, and the lateral face of the Hensen cell region were all clamped, i.e., set to zero displacement and no rotation. The fluids were modeled with the thermoviscous-acoustics capabilities of the COMSOL Acoustics module, which uses linearized Navier-Stokes equations and includes fluid viscosity. We neglected heat transmission. Rigid, no-slip boundary conditions were applied at the fluid-region walls. Velocity continuity was enforced at the interface between the viscoacoustic domain and the OoC solid elements.

Due to the irregularity of most of the geometry, the solid and fluid domains were meshed with coarse tetrahedral elements, except the viscous layer at the fluid-domain edges in which we generated a finer, more structured mesh. The thickness of the viscous boundary layer decreased with rising angular frequency as in water at room temperature. The fluid mesh surrounding the OoC and in the vicinity of the bony scalae surfaces had at least 3 nodes within the viscous boundary layer. Quadratic shape functions were used for the fluid, solid, and shell elements.

The mechanical drive to the model was an oscillatory pressure of 0.2 Pa applied to the scala-tympani side of the BM. There are no measurements of the scala-tympani or scala-media sound pressures near the BM in the middle region of the gerbil cochlea. We obtained the wavelength (and its reciprocal, the wavenumber "**k**") as a function frequency (**k** vs. frequency is termed the "wavenumber vs. frequency relationship", or "WFR") from the measured OoC motion at a second-turn place in the gerbil (Dong et al., 2018). The WFR changes with frequency can be thought of as expressing the effects of the rest of the cochlea on the traveling wave as seen at the slice.

To allow realistic transverse motion in the model, we used a computationally efficient method: a "Floquet" boundary condition. This considers the slice as a part of an infinite array of abutted identical slices driven identically. A wave traveling along the infinite array is realized by compelling the phase change between the basal and apical slice edges to match the phase change of the wavelength being mimicked.

The model was linear so motion scaled linearly with input pressure. The model was solved for frequencies between 100 Hz and 5 kHz in 1/6-octave steps.

RESULTS

The overall OoC motion profile is presented in Tubelli et al (2022). Here we concentrate on the motion drives to OHC and IHC stereocilia. The maximum deflections in any direction of OHC and IHC stereocilia are shown in Fig. 2. The OHC deflections were greatest in the OHC first row (OHC1), were smaller in the OHC second row (OHC2) and were least in the OHC third row (OHC3). Surprisingly, at low frequencies the amplitude of OHC3 was an order of magnitude less than that of OHC1. OHC1 and OHC2 had simple low-pass amplitude-vs.-frequency characteristics but OHC3 showed a small increase in amplitude at a frequency near the 2.5 kHz CF of the slice location. The deflection of the IHCs was more bandpass than lowpass with the largest amplitude approximately an octave below the location CF.

The deflections of OHC and IHC stereocilia bundles in each direction are shown in Fig. 3 relative to BM motion (this normalization removes the low-pass characteristics). The direction of the largest stereocilia deflection was generally in the radial direction (Fig. 3, left column), which is the main excitatory direction of both IHC and OHC stereocilia bundles.

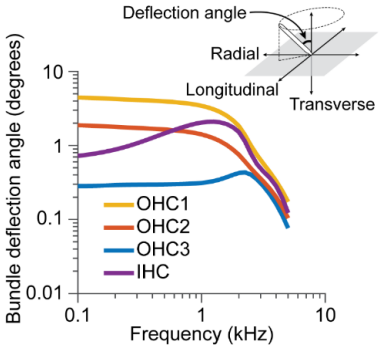


FIGURE 2. Bundle deflection angles for each of the four stereocilia rows as functions of frequency. Bundle deflection was calculated as the average maximum deflection angle in the radial-longitudinal plane of all stereocilia within each row. Deflections are for a sound that produced a 1 kHz BM motion of ~100 nm.

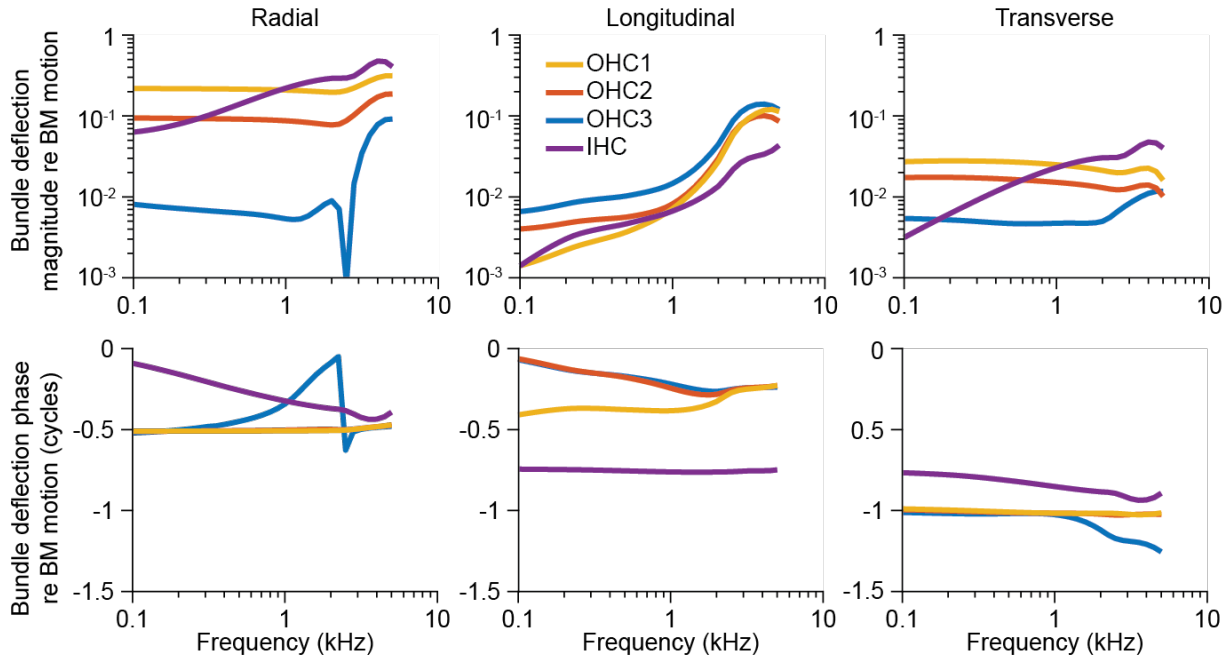


FIGURE 3. Bundle deflections in three directions for the four stereocilia rows as functions of frequency. For each direction, bundle deflection was calculated as stereocilia tip motion minus base motion averaged across the stereocilia of a hair cell. For averaging, the amplitude and phase of each motion were expressed as a complex number re. basilar membrane transverse motion at the arcuate-pectinate junction.

To understand why there were large differences in the drives across the three OHC rows, for two frequencies, we separately show the radial motions at the tops and bottoms of the OHC stereocilia (Fig. 4). For the three OHC rows, the motions varied relatively little at the tips (open circles, i.e., across the TM) but varied more at the bottoms (solid circles, i.e., across the RL). Thus, the change in OHC bundle deflection across the three rows was mostly due to

variation in the RL motion. At OHC3, the radial RL motion was almost the same as the motion of the TM; this small difference explains why the deflection of OHC3 was so much less than the deflection of OHC1.

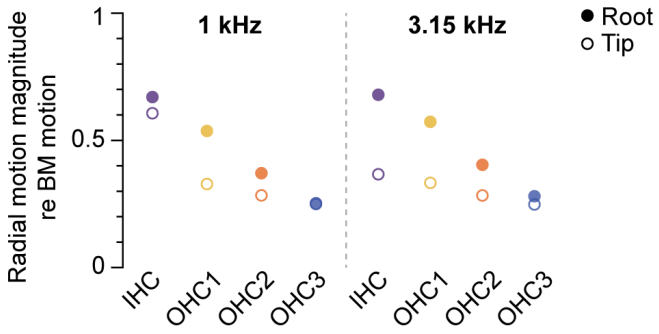


FIGURE 4. The radial motions at the tips (open circles) and the roots (solid circles) of the four stereocilia bundles at 1 kHz and 3.15 kHz. The difference in these motions produces the stereocilia bundle deflections in the radial direction. For each bundle, the stereocilia motions were averaged with the amplitude and phase of each motion expressed as complex numbers re. basilar membrane transverse motion at the arcuate-pectinate junction.

DISCUSSION

In the cochlear base, auditory-nerve (AN) fiber tuning curves (TCs) and mechanical tuning of the BM are similar in shape (Robles and Ruggero, 2001) but this is not the case in cochlear apical regions with CFs below ~4 kHz (Recio-Spinoso and Oghalai, 2017). In low-CF regions, AN TCs are bandpass, almost V-shaped, while gross mechanical structures show low-pass-like tuning (Recio-Spinoso and Oghalai, 2017; Dong et al., 2018). The low-pass characteristic of gross mechanical responses is reflected in the low-pass shapes of the OHC drives of Fig. 2. In contrast, the model IHC drive is bandpass (Fig. 2). It should be noted that the IHC-drive low-frequency slope is not enough to account for the low-frequency slopes of AN TCs, and that other low-pass factors must also be also involved.

Published calculations for the expected IHC stereocilia deflection due to shear between the RL and the TM indicate that the IHC stereocilia deflection versus frequency should follow structure velocity at the lowest frequencies, i.e., show a 6 dB per octave positive slope (Freeman and Weiss, 1990). The IHC deflection does show a positive slope at the lowest frequency but somewhat less than 6 dB per octave. The predicted 6 dB slope does not take into account a contribution to the IHC drive from fluid flow due to changes in the RL-TM gap (Guinan, 2012). We are now working to determine the fractions of the IHC drive accounted for by RL-TM shear versus RL-TM gap-changes.

It is also interesting that the phase of the IHC deflections almost reversed over the frequencies considered (Fig. 3, bottom-left column). This is not consistent with the IHC drive having a pure velocity dependance on the RL-TM shear. The model prediction that IHC drive phase can reverse fits with the observation that that the phase at which AN fibers are excited by low-frequency tones becomes reversed in the middle of the cochlea in guinea pigs, gerbils, chinchillas, and cats (Oshima and Strelioff, 1983; Ruggero and Rich, 1983; Temchin *et al.*, 2012; Nam and Guinan, 2016).

The drive to the OHCs has generally been thought to be similar across the three rows of OHCs. Previous models have not been good enough to shed light on this. Our model is the first with adequate detail to make a credible case that the drive can be different across the OHC rows. Model motion within the OoC indicates that as the BM moves up, the pillar cells rotate with their tops moving toward the modiolus (the classic direction) which produces RL radial motion. The RL radial motion is less at OHC1 than at the heads of the pillar cells because the RL stretches. RL stretching extends out to OHC3 so that RL motion at RL3 is substantially less than at RL1 (Fig. 4). One can ask what is holding back the RL radial motion near RL3 so that it doesn't follow the radial motion at RL1? We don't fully understand this, but some things can be suggested: (1) the OHC3 stereocilia provide a bridge between the RL and the TM that ties their motions together, (2) radial movement toward the modiolus of OHC3 requires stretching of the structures lateral of OHC3 (the Hensen cell region) and/or (3) the squeezing of the OoC produced by the BM moving up (in the model the BM moves up more than the RL moves up – Tubelli et al., 2022) produces an OoC internal pressure that expands the lateral part of the OoC outward, which pulls the RL outward. Empirical, in-vivo evidence that the RL is not rigid comes from the transverse motions at the tops of the three OHC rows, which show substantially more transverse motion of the RL over OHC3 than over OHC1 (Cho and Puria, 2022). Our model indicates that even in a passive cochlea there is enough elasticity in the RL that RL radial motion is different across the three OHC rows.

ACKNOWLEDGEMENTS

Supported in part by NIH NIDCD R01 DC007910 (to Sunil Puria).

REFERENCES

1. W. Dong, E. S. Olson, Detection of cochlear amplification and its activation. *Biophys J* **105**, 1067-1078 (2013).
2. R. Edge *et al.*, Morphology of the unfixed cochlea. *Hearing Res.* **124**, 1-16 (1998).
3. M. Müller, The cochlear place-frequency map of the adult and developing Mongolian gerbil. *Hearing research* **94**, 148-156 (1996).
4. N. B. Slepecky, "Structure of the Mammalian Cochlea" in The Cochlea, P. J. Dallos, A. N. Popper, R. R. Fay, Eds. (Springer-Verlag, New York, 1996), pp. 44-129.
5. J. A. Tolomeo, C. R. Steele, M. C. Holley, Mechanical properties of the lateral cortex of mammalian auditory outer hair cells. *Biophys J* **71**, 421-429 (1996).
6. J. A. Tolomeo, M. C. Holley, Mechanics of microtubule bundles in pillar cells from the inner ear. *Biophys J* **73**, 2241-2247 (1997).
7. J. B. Sellon, S. Farrahi, R. Ghaffari, D. M. Freeman, Longitudinal spread of mechanical excitation through tectorial membrane traveling waves. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 12968-12973 (2015).
8. N. P. Cooper, "Radial variation in the vibrations of the cochlear partition" in Recent Developments in Auditory Mechanics, H. Wada, T. Takasaka, K. Ikeda, K. Ohyama, Eds. (World Scientific, Singapore, NJ., 2000), pp. 109-115.
9. W. S. Rhode, A. Recio, Study of mechanical motions in the basal region of the chinchilla cochlea. *J. Acoust. Soc. Am.* **107**, 3317-3332 (2000).
10. M. Homer, A. Champneys, G. Hunt, N. P. Cooper, Mathematical modeling of the radial profile of basilar membrane vibrations in the inner ear. *J Acoust Soc Am* **116**, 1025-1034 (2004).
11. N. Cooper, M. Van der Heijden, "Spatial Profiles of sound-evoked vibration in the gerbil cochlea" in To the ear and back again: Advances in auditory biophysics. Proc. 13th Mechanics of Hearing Workshop, C. Bergevin, S. Puria, Eds. (AIP Conference Proceedings, Melvile, NY, 2018), vol. 1965, pp. 080001.
12. M. Müller, The cochlear place-frequency map of the adult and developing Mongolian gerbil. *Hearing research* **94**, 148-156 (1996).
13. L. Schweitzer, C. Lutz, M. Hobbs, S. P. Weaver, Anatomical correlates of the passive properties underlying the developmental shift in the frequency map of the mammalian cochlea. *Hear Res* **97**, 84-94 (1996).
14. H. Motallebzadeh, J. A. M. Soons, S. Puria, Cochlear amplification and tuning depend on the cellular arrangement within the organ of Corti. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 5762-5767 (2018).
15. J. H. ten Kate, J. W. Kuiper, The viscosity of the pike's endolymph. *J Exp Biol* **53**, 495-500 (1970).
16. W. Dong *et al.*, Organ of Corti vibration within the intact gerbil cochlea measured by volumetric optical coherence tomography and vibrometry. *Journal of neurophysiology* **120**, 2847-2857 (2018).
17. A. A. Tubelli, H. Motallebzadeh, J. J. Guinan, Jr., S. Puria, A gerbil cochlear-slice model is tuned by the traveling wave and not by a local frequency resonance. Biophysical J. (in review) (2022)
18. L. Robles, M. A. Ruggero, Mechanics of the mammalian cochlea. *Physiological Reviews* **81**, 1305-1352 (2001).
19. A. Recio-Spinoso, J. S. Oghalai, Mechanical tuning and amplification within the apex of the guinea pig cochlea. *The Journal of physiology* **595**, 4549-4561 (2017).
20. D. M. Freeman, T. F. Weiss, Hydrodynamic forces on hair bundles at low frequencies. *Hearing Res.* **48**, 17-30 (1990).
21. J. J. Guinan, Jr., How are inner hair cells stimulated? Evidence for multiple mechanical drives. *Hear Res* **292**, 35-50 (2012).
22. W. Oshima, D. Strelioff, Responses of gerbil and guinea pig auditory nerve fibers to low- frequency sinusoids. *Hear Res* **12**, 167-184 (1983).
23. M. A. Ruggero, N. C. Rich, Chinchilla auditory-nerve responses to low-frequency tones. *J. Acoust. Soc. Am.* **73**, 2096-2018 (1983).

24. A. N. Temchin, A. Recio-Spinoso, H. Cai, M. A. Ruggero, Traveling waves on the organ of corti of the chinchilla cochlea: spatial trajectories of inner hair cell depolarization inferred from responses of auditory-nerve fibers. *J Neurosci* **32**, 10522-10529 (2012).
25. H. Nam, J. J. Guinan, Jr., Low-frequency bias tone suppression of auditory-nerve responses to low-level clicks and tones. *Hear Res* **341**, 66-78 (2016).
26. N.-H. Cho, S. Puria, Motion of the cochlear reticular lamina varies radially across the outer-hair-cell rows. *Biorvix* submitted (2022).

The Transepithelial Potential Can Control Gating Compliance of the Hair-Cell Bundle

A. JOLIOT, P. MARTIN [a)]

Laboratoire Physico-Chimie Curie, Institut Curie, Université PSL, Sorbonne Université, CNRS UMR168, F-75248 Paris, France

[a)] Corresponding author: pascal.martin@curie.fr

Abstract. Taking advantage of a two-chamber excised preparation of the frog saccule, we study the effects of imposing a static transepithelial current, and in turn a transepithelial potential, on the force-displacement relation of single hair-cell bundles. Our observations suggest that the transepithelial potential may serve as a control parameter of gating forces of the transduction channels and in turn of hair-bundle mechanosensitivity.

SUMMARY

Hearing is initiated by vibrations of the hair bundle, which operates as a mechanosensitive antenna for the sensory hair cells of the inner ear [1]. Mechanosensitivity is thought to result from the direct mechanical activation of ion channels by tension change in tip links interconnecting the stereocilia of the hair bundle, resulting in an electrical transduction current [2,3]. Within this framework, mechanoelectrical transduction is reciprocal: gating of the transduction channels affects tip-link tension, which effectively reduces the stiffness of the hair bundle — a phenomenon known as 'gating compliance' [4].

This defining mechanical signature of hair-cell mechanosensitivity has been reported with variable magnitudes across animal species and ex-vivo preparations of sensory epithelia [4-10]. Despite its importance, the regulation of gating compliance, and more generally of hair-cell mechanosensitivity, by physiological parameters remains poorly understood. Here, using a two-chamber excised preparation of the frog saccule [9], we studied the effects on hair-bundle mechanics of an electric potential difference applied across the sensory epithelium, mimicking the endocochlear potential observed in the mammalian cochlea.

By imposing a static transepithelial current of ± 7 µA, the potential into the artificial endolymph that bathes the hair bundles, with respect to the perilymph to which the hair-cell somas are exposed, was varied within a range of ± 80 mV. In addition to the expected adaptive shift of the force-displacement relations, we found that the magnitude of the whole-bundle change in tip-link tension upon channel gating—the global gating force—displayed a negative correlation with the transepithelial potential. Correspondingly, upon application of a negative potential in endolymph, the gating force could increase from an average value of 18 ± 8 pN ($n = 12$) by as much as by 30 pN or about 100%. These effects were graded and reversible. Although a negative correlation was also observed between the hair-bundle stiffness and the transepithelial potential, stiffness

changes were too small to account for the observed variations of the gating force. We conclude that the transepithelial potential may serve as a control parameter of gating compliance and in turn of hair-bundle mechanosensitivity.

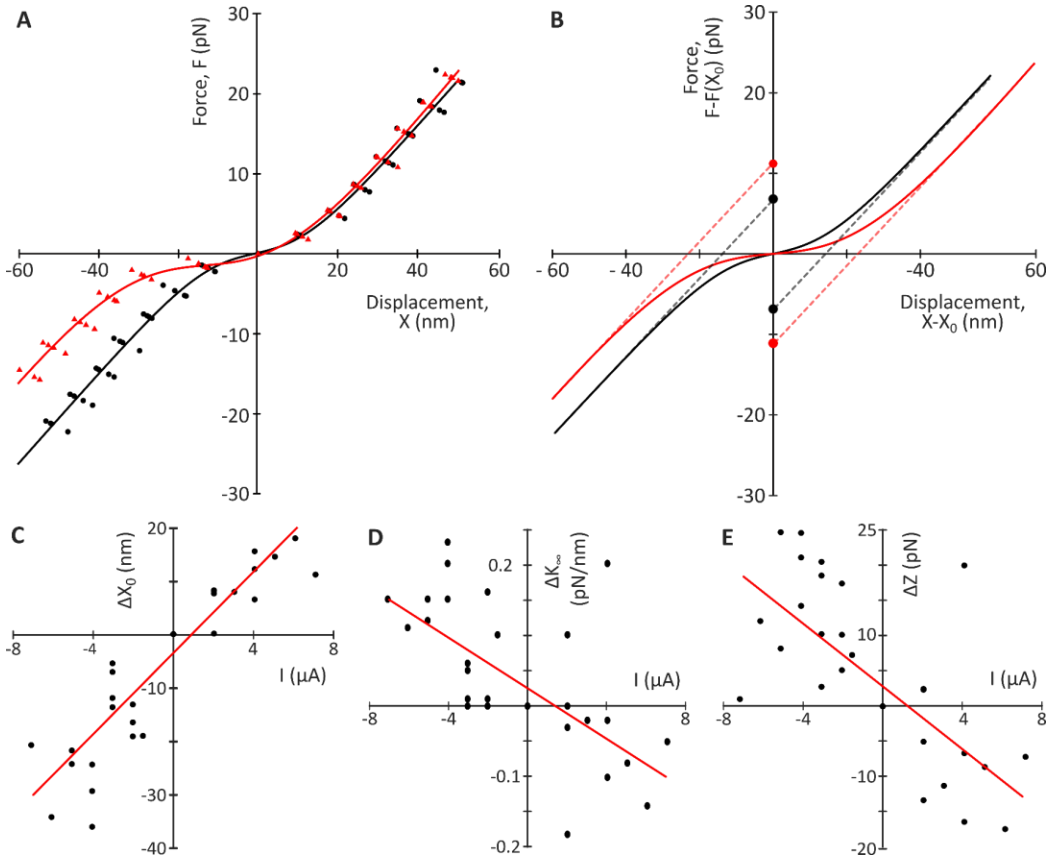


FIGURE 1: Effects of a transepithelial current on hair-bundle mechanics. (A) Force-displacement relation under control conditions (transepithelial current $I = 0$ μA, black) and in the presence of a fixed transepithelial current $I = -3$ μA (red). The hair-bundle mechanics was probed alternatively 4 times under each condition, demonstrating reversibility and repeatability of the effects. The solid lines correspond to a global fit of the 4 measurements under each condition by the equation $F(X) = K_\infty X - Z/[1 + \exp(-(X - X_0)/\delta)] + F_0$. Under control conditions ($I = 0$ μA, black line), we found the asymptotic hair-bundle stiffness $K_\infty = 0.56$ pN.nm^{-1}, the whole-bundle gating force $Z = 14$ pN, the set-point deflection $X_0 = -1$ nm, the offset force $F_0 =$ 7.5 pN, and the characteristic distance $\delta =$ 7.9 nm, whereas in the presence of a fixed transepithelial current ($I = 3$ μA, red line) we found $K_\infty = 0.61$ pN.nm^{-1}, $Z = 29$ pN, $X_0 = -13$ nm, $F_0 = 21$ pN, and $\delta =$ 13 nm. (B) Fits to the force-displacement relations shown in A are here plotted with respect to the origin $(X_0, F(X_0))$. Under each condition, the magnitude of the whole-bundle gating force Z is given by the vertical distance between the filled disks on the force axis, showing the increase in gating-force upon application of the negative transepithelial current. (C-E) Change ΔX_0 of the set-point deflection (C), change ΔK_∞ of the asymptotic hair-bundle stiffness (D), and change ΔZ of the gating force (E) as a function of the transepithelial current I with respect to the control condition ($I = 0$ μA). In each panel, linear fit to the data are shown as thin lines. In (C), the adaptive shift ΔX_0 is positively correlated to the transepithelial current I (Pearson test: $r =$ 0.9, p $-$ value $= 7 \cdot 10^{-15}$), with a linear fit (red line) of slope 3.8 ± 0.7 nm $\cdot$ μA^{-1} and intercept of -3 ± 2 nm. In (D), the change ΔK_∞ is negatively correlated to the transepithelial current I (Pearson test: $r = -0.65$, p $-$ value $= 1 \cdot 10^{-5}$), with a linear fit (red line) of slope -0.02 ± 0.01 pN $\cdot$ nm$^{-1} \cdot \mu A^{-1}$. and intercept of $0.03 \pm 0.01\, pN \cdot nm^{-1}$. In (E), the change in gating force Z is negatively correlated to the transepithelial current I (Pearson test: $r = -0.70$, p $-$ value $= 1 \cdot 10^{-6}$), with a linear fit (red line) of slope -2.2 ± 0.8 pN $\cdot$ μA^{-1} and intercept of 3 ± 2 pN. Under control conditions, over an ensemble of 12 hair cells, the average values of these parameters were $X_0 = -4 \pm 9$ nm, $K_\infty = 0.9 \pm 0.3$ pN $\cdot$ nm^{-1}, and $Z = 18 \pm 8$ pN. Note that applying a negative transepithelial current could increase the gating force by nearly 100%, whereas the asymptotic hair-bundle stiffness increased by at most 25%.

ACKNOWLEDGMENTS

This research was supported by the Fondation pour l'Audition (project n° FPA RD 2020-7).

REFERENCES

1. Hudspeth, A. J. 1989. How the ear's works work. Nature. 341:397–404.
2. Corey, D. P., and A. J. Hudspeth. 1983. Kinetics of the receptor current in bullfrog saccular hair cells. J. Neurosci. 3:962–976.
3. Martin, P., and A.J. Hudspeth. 2021. Mechanical Frequency Tuning by Sensory Hair Cells, the Receptors and Amplifiers of the Inner Ear. Annu. Rev. Condens. Matter Phys. 12:29–49.
4. Howard, J., and A.J. Hudspeth. 1988. Compliance of the hair bundle associated with gating of mechanoelectrical transduction channels in the bullfrog's saccular hair cell. Neuron, 1, 189-199.
5. Russel, I. J., M. Kossl, and G. P. Richardson. 1992. Nonlinear mechanical responses of mouse cochlear hair bundles. Pric. R. Soc. Lond. B 250: 217-227
6. Martin, P., A. D. Mehta, and A. J. Hudspeth. 2000. Negative hair-bundle stiffness betrays a mechanism for mechanical amplification by the hair cell. Proc. Natl. Acad. Sci. USA. 97:12026–12031.
7. Ricci, A. J., A. C. Crawford, and R. Fettiplace. 2000. Active hair bundle motion linked to fast transducer adaptation in auditory hair cells. J. Neurosci. 20:7131–7142.
8. Kennedy, H. J., A. C. Crawford, and R. Fettiplace. 2005. Force generation by mammalian hair bundles supports a role in cochlear amplification. Nature. 433:880–883.
9. Tinevez, J. Y., F. Jülicher, and P. Martin. 2007. Unifying the various incarnations of active hair-bundle motility by the vertebrate hair cell. Biophys. J. 93:4053–4067.
10. Tobin, M., A. Chaiyasitdhi, V. Michel, N. Michalski, and P. Martin. 2019. Stiffness and tension gradients of the hair cell's tip-link complex in the mammalian cochlea. eLife, 8, e43473.

Hair Bundle Micromechanics Including Stereocilia Kinematics and the Interaction of Stimulus and Bundle Rate Constants

Varun Goyal[1, a)] and Karl Grosh[1, 2, 3, b)]

[1)]*Department of Mechanical Engineering, University of Michigan, G.G. Brown Building, 2350 Hayward St., Ann Arbor, MI 48109-2125, U.S.A.*

[2)]*Department of Biomedical Engineering, University of Michigan, 1107 Carl A. Gerstacker Building, 2200 Bonisteel, Blvd., Ann Arbor, MI 48109-2099, U.S.A.*

[3)]*Kresge Hearing Research Institute, 4605 Medical Science Unit II, Ann Arbor, MI 48109-5616, U.S.A.*

[a)]*Corresponding author: varungo@umich.edu*

[b)]*Electronic mail: grosh@umich.edu.*

Abstract. In this paper, the relation between the current and displacement responses due to an external mechanical stimulus on the outer hair cell (OHC) hair bundle (HB) is studied using a theoretical approach. We seek to understand the interplay between the time constants of the external loading and those intrinsic to the HB. To incorporate HB adaptation and channel gating, we used the model of an isolated HB, denoted as the TMJ model. We solved the nonlinear equations for the bundle dynamic response due to an externally applied force that consisted of an exponential temporal rise to a constant value. We determined the dependence of the bundle displacement over which the current continued to increase (the apparent operating range (OR)) on the rise time (τ_F) of the applied force. In addition, we developed a model linearized about the resting open probabilities for a given static, biasing load to provide closed-form approximations of the dependence on the stimulus rise time, τ_F, and bundle adaptation time constants. Finally, we wanted to determine if the inclusion of more precise kinematics of the tip link motion relative to the stereociliary rotation influenced model predictions of the channel opening and bundle stiffness. Hence, we developed geometrical relations between the two rows of stereocilia to establish coupled kinematic relations for inclusion in our HB kinetic model. We predict an OR of $30-50$ nm for small τ_F and an overestimation by a factor of 10 in the OR for τ_F higher than the slow adaptation time constant. Finally, with accurate bundle kinematics, lower HB displacement, current, and adaptation motor displacement were predicted in contrast to the TMJ model. We are exploring the implications of this model on nanoscale mechano-electrical transduction.

INTRODUCTION

The outer hair cells (OHCs) in the mammalian cochlea are primarily responsible for the active, nonlinear processes occurring under acoustic stimulations. They enhance the sensitivity and frequency-selectivity of the sound (nonlinear amplification/compression), possibly via two main mechanisms- change in OHC length (so-called somatic motility [1]) and electromotility of the hair bundle (HB) [2] situated on the apical surface of OHC. Anatomically, each HB consists of hair-like structures called stereocilia, typically arranged in three rows with a species- and spatially-dependent number of columns. The stereocilia are arranged in the increasing order of their heights, with the shortest radially closer to the modiolus (the central conical axis of the cochlea). Adjacent stereocilia are connected through lateral links, and within each row, the tip of the shorter stereocilium is linked to the lateral surface of its taller neighbor via tip links [3]. As all the tallest rows of stereocilia are embedded in the tectorial membrane (TM), an external force moving the TM pushes/pulls these stereocilia uniformly. Within each row, for an excitatory force, the tallest row pulls the two shorter stereocilia, thereby activating the mechano-electrical transduction (MET) channels located at each of the lower tip link densities (LTLD- attachment of the tip link to the shorter stereocilium) [4]. This activation leads to the influx of cations like Ca^{2+} and K^{+} into the stereocilia, where they depolarize the cell and may interact with the myosin molecules to adapt the HB to sustained stimuli [1, 5].

Researchers have performed numerous experimental procedures on different mammalian cochleae to gauge the micromechanics involved and comprehend how the ear functions. However, the cochlea is small and located inside the temporal bone, making it difficult to access, and there is also a risk of permanently damaging the ear. Hence, the focus has been on building mathematical models and conducting *in-vitro* experiments on the excised HBs to understand this mechano-electrical machinery. Howard *et al.* [6] developed a two-state (open/closed) gating-spring model where transduction channels opened when HB deflection tensed the tip links. The gate open probability density was averaged over the HB and defined as a first-order Boltzmann function of bundle displacement. However, they assumed the actin-myosin interaction as a passive component in their gating-spring model. Choe *et al.* [7] proposed a six-state model where ions bound to the transduction channels [8, 9] to promote channel closure. Their model provided a mechanism for the generation of mechanical energy over a cycle, putatively from the binding energy of ions. Beurg *et al.* [5] suggested that adaptation may have two components- fast adaptation near the gate and slow adaptation resulting from

myosin motor action. Tinevez *et al.* [1] extended the gating-spring model [6], and added a degree-of-freedom for myosin motion (called adaptation motor). Their probability function depended on relative motion between HB and adaptation motor. However, in their simulations, a force was applied to the HB instantaneously, which is not possible experimentally. Nam *et al.* [10] highlighted the effects of the rise time of the external stimulus on the experimental estimates of the operating range (OR) (10% - 90% width of current-displacement relation). We wanted to discern whether a simpler model could replicate this prediction.

In the present study, our first goal is to understand the interplay between force rise times and adaptation time constants based on the two-degree-of-freedom adaptation model developed by Tinevez *et al.* [1], called the TMJ model. We achieve this by linearizing the system of equations and comparing the nonlinear and linear bundle displacements, currents, and dependence of the OR on the aforementioned time constants. This study's second and ongoing goal is to carefully model the stereocilia geometry for two rows and use precise kinematic relations to estimate kinetic bundle displacements and tip link motion with stereociliary rotation.

MODEL AND METHODS

Rise-Time Analysis

The TMJ model [1] is an adaptation model for an isolated HB with two-degrees-of-freedom, one for the bundle motion X and the other for the adaptation motor motion X_a, measured from a reference where gating springs bear no tension. Below are the equations of motion:

$$\lambda \frac{dX}{dt} = -K_{GS}\left(X - X_a - DP_o\right) - K_{SP}X + F_{ext} + \eta, \tag{1}$$

$$\lambda_a \frac{dX_a}{dt} = K_{GS}\left(X - X_a - DP_o\right) - K_{ES}\left(X_a - X_{ES}\right) - F_{max}\left(1 - SP_o\right) + \eta_a. \tag{2}$$

The variables are defined in Table I. η and η_a are noise terms taken zero in this study. A two-state Boltzmann function for the channel open probability (P_o) is defined as a function of relative motion between the bundle displacement and the adaptation motor displacement in this model, as shown in Eq. (3) with $A = \exp([\Delta G + (K_{GS}D^2)/(2N)]/k_BT)$ [11].

$$P_o = \frac{1}{1 + A\exp\left(-\frac{K_{GS}D}{Nk_BT}\left(X - X_a\right)\right)} \tag{3}$$

A schematic of this model is shown in Fig. 1 a). We solved this nonlinear model using the 4^{th}-order Runge-Kutta

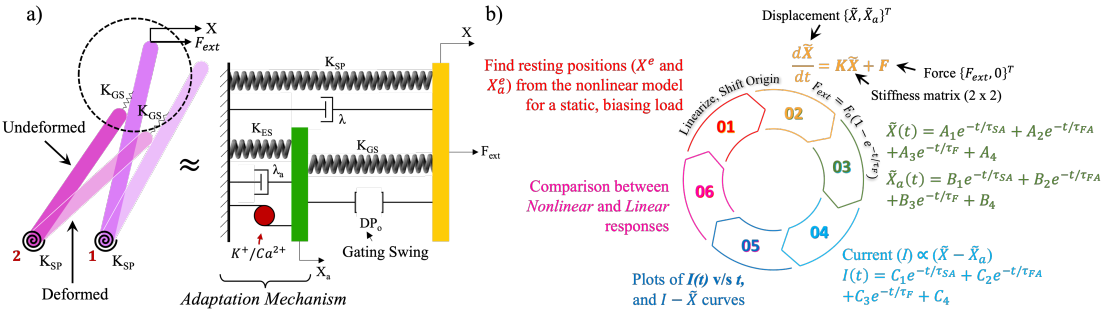


FIGURE 1. a) A mechanical system representation of the two-row TMJ model [1] with the undeformed configuration (dark purple) and deformed configuration (light purple) is shown on the left and the mechanistic model on the right (yellow- a represntation of the tallest stereocilium, green- a representation of the adaptation motor). b) Linearization procedure to study the dependence of bundle displacement and current on τ_F and adaptation time constants (τ_{FA} and τ_{SA}). In step 1, a static force brings the system to a resting state (X^e, X_a^e) shown in yellow. The nonlinear system is then linearized about the evaluated resting state to obtain the Jacobian (stiffness matrix K) and shifted coordinates $(\tilde{X}, \tilde{X}_a)$. In step 2, the linearized system is solved for an input force (F_{ext}) with varying rise times (τ_F) and amplitude (F_o) to obtain the closed-form solutions for linear $\tilde{X}$ and $\tilde{X}_a$ shown in green. Steps $4-6$ highlight the post-processing of the linearized solutions obtained in step 3.

TABLE I. Definitions of the parameters in the TMJ model [1] and the model developed in this paper with higher fidelity geometric considerations. The subscript '1' indicates the parameters used in the model developed in this paper.

Variable	Definition and Value	Variable	Definition and Value
λ	Viscous coefficient of the hair bundle and the stimulus fiber; $\lambda = 1\ \mu Ns/m$	K_{SP}	Combined pivot stiffness and stimulus fiber's stiffness; $K_{SP} = 3.1\ mN/m$
λ_1, λ_2	Rotational viscous damping in the vicinity of the taller (λ_1) and the shorter (λ_2) stereocilia; $\lambda_1 = \lambda_2 = 9.05\ aNsm$ (atto Newton)	K_{SP_1}, K_{SP_2}	Pivot torsional stiffness for the taller (K_{SP_1}) and the shorter (K_{SP_2}) stereocilium (or row 2); $K_{SP_1} = K_{SP_2} = 27.2\ fNm$
D, d_1	Gating swing in the x direction (D) and along the inclined gating spring (d_1); $D = 45\ nm$, $d_1 = 11.3\ nm$	K_{ES}, K_{ES_1}	Stiffness of the extent springs regulating the adaptation motor motion; $K_{ES} = K_{ES_1} = 1\ mN/m$
F_{max}, F_{max_1}	Stall force for the motor when $P_o = 0$; $F_{max} = F_{max_1} = 140\ pN$	K_{GS}, K_{GS_1}	Stiffness of the gating spring; $K_{GS} = 8\ mN/m$, $K_{GS_1} = 64.4\ mN/m$
N, N_1	Number of transduction channels; $N = 120$, $N_1 = 24$	S, S_1	Strength of the Ca^{2+} on the motor force; $S = S_1 = 6$
λ_a, λ_{a1}	Force-velocity relation's slope for the adaptation motor; $\lambda_a = \lambda_{a1} = 0.5\ \mu Ns/m$	X_{ES}	Position of the adaptation motor when the tension in the extent springs is zero

method in MATLAB for a time-varying force $F_{ext} = F_o(1-\exp(-t/\tau_F))$ with rise time τ_F and magnitude F_o.

Along with the complete nonlinear analysis, we linearized the system about different resting conditions. A static, biasing load was applied on the tallest row and the equations for small motion were obtained and solved as described in Fig. 1 b). This analysis allowed us to obtain closed-form estimates of the eigenvalues related to the slow (τ_{SA}) and fast (τ_{FA}) response time constants. Finally, we varied τ_F and predicted the changes in the responses.

Including Accurate Kinematics and Geometry in a Kinetic HB Model

We denote the tall and short stereocilia as rows 1 and 2, respectively. In the geometry, we assumed that both the rows have round tips and that row 2 made sliding contact with row 1. This point of contact slips as the bundle rotates. A gating spring with stiffness K_{GS_1} represents the tip link. We define two generalized coordinates, the rotation of the tallest stereocilium ϕ and the motion of the adaptation motor along the axis of the tallest stereocilium δ_{a1}. Thereby, we determined the tip link length (l_1), location of the point of contact from row 1 pivot point (point A), and associated angles (α_2 and α_3) portrayed in Fig. 2 a) as functions of the two generalized coordinates. At $\alpha_1 = 90^o$, we assume that the length of the tip link insertion point in row 1 from point A (l_1^{gs}) is known and that it varies as ($l_1^{gs} - \delta_{a1}$) along its axis as the bundle rotates. A static force rotates row 1 by ϕ^o, and the rotation ϕ is measured from this resting state. We evaluated the geometries of the two polygons $ABHB'D$ and $ABGJIB'D$, as depicted in Fig. 2 a) to find 1^{st}-order Taylor series approximations of varying lengths (r_{BA} and l_1) and angles (α_1, α_2, α_3) as functions of ϕ and δ_{a1}. The dimensions are estimated from the model developed by Beurg *et al.* [5]. The linearized expressions for the above mentioned variables are:

$$l_1 = l_1^o + \varepsilon_4\phi + \varepsilon_5\delta_{a1}, \tag{4}$$

$$r_{BA} = r_{BA}^o + \varepsilon_2\phi, \tag{5}$$

$$\alpha_1 = \alpha_1^o + \phi, \tag{6}$$

$$\alpha_2 = \alpha_2^o + \varepsilon_1\phi, \tag{7}$$

$$\alpha_3 = \alpha_3^o + \varepsilon_3\phi + \varepsilon_6\delta_{a1}, \tag{8}$$

where ε_1, ε_2, ε_3, ε_4, ε_5, and ε_6 are coefficients obtained from the Taylor series expansion, and l_1^o, r_{BA}^o, α_1^o, α_2^o, and α_3^o are the values at the resting state. Inspired by the gating-spring model developed by Howard *et al.* [6], we estimated the energies of channel open (ξ_o) and closed (ξ_c) states. Their difference ($\xi_o - \xi_c = K_{GS_1}d_1(l_1 - l_1^o)$) gives the channel opening probability density as:

$$P_{o1} = \frac{1}{1 + A_1\exp\left(-\frac{K_{GS_1}d_1}{N_1k_BT}(\varepsilon_4\phi + \varepsilon_5\delta_{a1})\right)}, \tag{9}$$

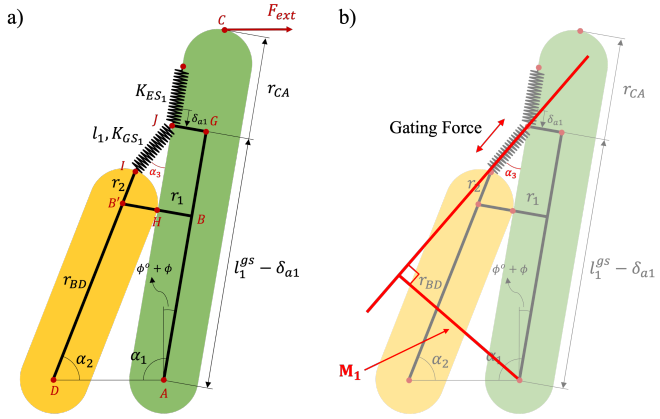


FIGURE 2. a) A higher precision kinematic and geometric model for two rows of stereocilia showing various dimensions, model parameters, and generalized coordinates ϕ and δ_{a1}. $r_1 = 113$ *nm* and $r_2 = 118$ *nm* are the radii of rows 1 and 2, respectively. $r_{CA} = 4.25 \mu m$ is the length of row 1 from the pivot point A. $r_{BD} = 2.07$ μm is the length of row 2 minus its radius (r_2) from the pivot point D. l_1 is the length of the gating spring/tip link. $l_1^{gs} = 2.36$ μm is where the tip link contacts row 1 lateral surface when row 1 is vertical. Finally, K_{GS_1} and K_{ES_1} are the gating spring and extent spring stiffness, respectively. The separation between the pivot points A and D is $b_{12} = AD = 609$ *nm*. The dimensions are estimated from the model developed by Beurg *et al.* [5]. b) The time-dependent moment arm (M_1) for the gating force. $M_1 = (l_1^{gs} - \delta_{a1})\sin(\alpha_3) + r_1 \cos(\alpha_3)$ is a function of both ϕ and δ_{a1}.

where $\varepsilon_4 \phi + \varepsilon_5 \delta_{a1} = l_1 - l_1^o = \delta l_1$ is the tip link extension from Eq. 4 and A_1 sets the open probability at rest. Note that the tip link extension and the angle it makes with row 1 are functions of both stereociliary rotation and adaptation motor displacement. Using the geometrical relations obtained in Eq. 4 - Eq. 8 and applying Lagrange's energy method [12], we obtain two equations of motion for perturbations from the equilibrium (ϕ^o, δ_{a1}^o):

$$(\lambda_1 + \varepsilon_1^2 \lambda_2)\frac{d\phi}{dt} = -K_{GS_1}\left(\varepsilon_4^2 \phi + \varepsilon_4 \varepsilon_5 \delta_{a1} - d_1 P_{o1} M_1\right) - \left(K_{SP_1} + \varepsilon_1^2 K_{SP_2}\right)\phi + F_{ext} r_{CA} \cos(\phi^o + \phi), \tag{10}$$

$$\lambda_{a1}\frac{d\delta_{a1}}{dt} = K_{GS_1}\left(-\varepsilon_4 \varepsilon_5 \phi - \varepsilon_5^2 \delta_{a1} - d_1 P_{o1} \cos(\alpha_3)\right) - K_{ES_1}\delta_{a1} - F_{max_1}(1 - S_1 P_{o1}), \tag{11}$$

where $M_1 = (l_1^{gs} - \delta_{a1})\sin(\alpha_3) + r_1 \cos(\alpha_3)$ is the moment arm for the nonlinear gating force. Table I lists the other parameters' definitions. Eq. 10 and Eq. 11 are similar to Eq. 1 and Eq. 2, respectively, in their forms with newly added nonlinear, time-dependent moment arm (Eq. 10) and component (Eq. 11) of the gating force in our model.

RESULTS

Modified TMJ Model

We biased the system using multiple static forces in the nonlinear model to vary the resting channel open probability from 5% to 40%. A static force of 800 pN and 1300 pN corresponds to 5% and 40% resting channel open probability, respectively. From the equilibrium position, we simulated the model by varying F_o and τ_F. In Fig. 3, we plot the bundle displacement and inward current (proportional to the channel open probability (P_o)) as a function of time for a variety of rise times. We observed that the saturation of the bundle displacement is delayed as τ_F increased (as shown in Fig. 3 a)), and the inward current peak diminished while shifting rightwards until saturation with increasing τ_F as illustrated in Fig. 3 b).

The current displacement (I-X) relations to predict the OR and its dependence on τ_F are shown in Fig. 4 a). Fig. 4 b) shows how the OR increased with τ_F and saturated as τ_F approached 0.5 ms. From the linear model, we found closed-form solutions for HB and adaptation motor displacement shown in Fig. 1 b) as sums of three exponentials-fast adaptation time constant (τ_{FA}), slow adaptation time constant (τ_{SA}), and force rise-time (τ_F). We also saw a high temporal correspondence between the nonlinear and linear responses. We then evaluated the current proportional to the relative motion between the bundle and motor displacements $X - X_a$. We observed that the current peaked due

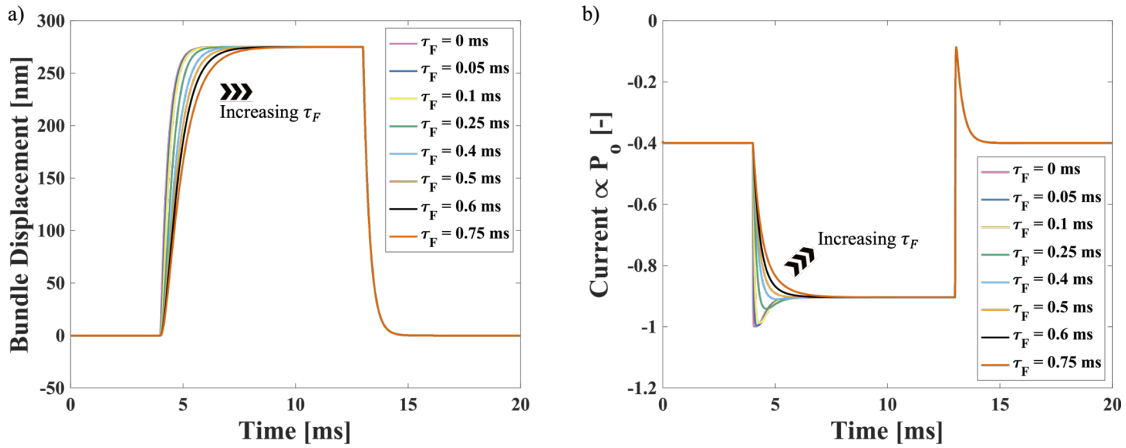


FIGURE 3. a) Bundle displacement from the nonlinear model for $F_o = 700$ pN and $P_o^e = 40\%$. As the force rise-time increases, the bundle displacement saturates more slowly. b) Inward current (proportional to channel open probability (P_o)) responses from the nonlinear model for $F_o = 700$ pN and $P_o^e = 40\%$. The current peak decreases in magnitude and shifts to the right in time with increasing τ_F until there is no distinguishable peak for τ_F values approaching 0.5 ms.

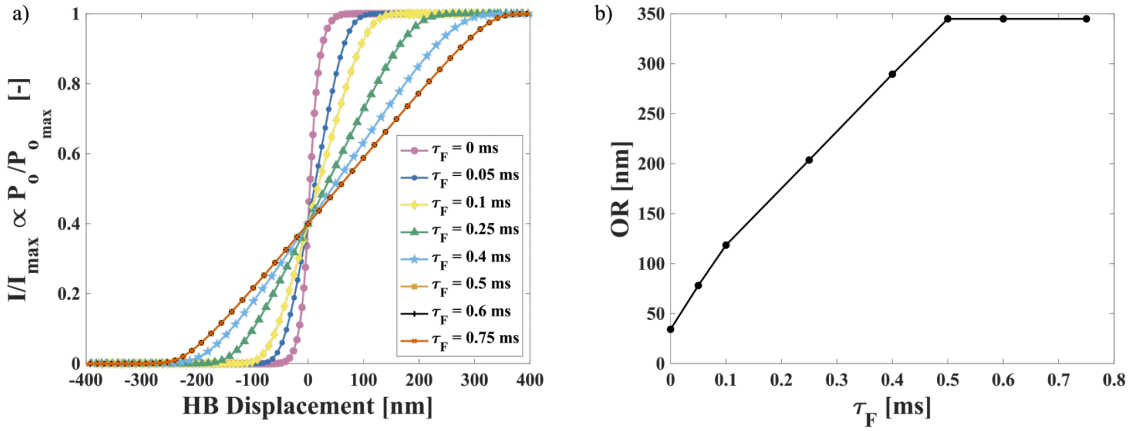


FIGURE 4. a) I-X curves obtained by simulating the TMJ model with varying F_o and τ_F for $P_o^e = 40\%$. The current continues to increase until saturation for a wider range of HB displacements as τ_F increases. As $\tau_F > 0.5$ ms, the broadening stops. b) The OR as a function of τ_F leading to a constant value. As force application speed slows down, the OR ceases to change once $\tau_F > 0.5$ ms.

to the fast adaptation mechanism when $\tau_F < \tau_{FA}$. When τ_F is increased from τ_{FA} to τ_{SA}, it is the external force that controlled the current peak, likely causing the peak reduction and rightward shift in the nonlinear model (in Fig. 3 b)). Finally, when τ_F exceeded τ_{SA}, there is no peak seen as the force application is too slow for adaptation to influence it. For the case of 1300 pN static force, the linear model estimated $\tau_{FA} \approx 4\mu$s and $\tau_{SA} \approx 0.3$ ms. Correspondingly, the OR from the nonlinear model also reached its highest value when τ_F approached to the slow adaptation time constant from the linear model.

Predictions from the New Kinematic and Geometric HB Model

We simulated this model with the same parameters used in the case of the modified TMJ model [1] and added kinematic constraints. For individual stereociliary pivot stiffness (K_{SP_1} and K_{SP_2}), we assigned half the stiffness value in the TMJ model for K_{SP} times the square of r_{CA} (length of row 1) to obtain the torsional spring stiffness. Similarly, for the fluid damping, we allocated half the value of λ multiplied by the squared length of row 1 (r_{CA}) to both λ_1 and λ_2

(torsional damping coefficients). The upright position ($\alpha_1 = 90^o$) is the default state of the HB. Since our model is in the development phase, we set the values of A_1 and N_1 to match the resting open probabilities for a static force of 800 pN in the TMJ model and the new kinematic and geometric model.

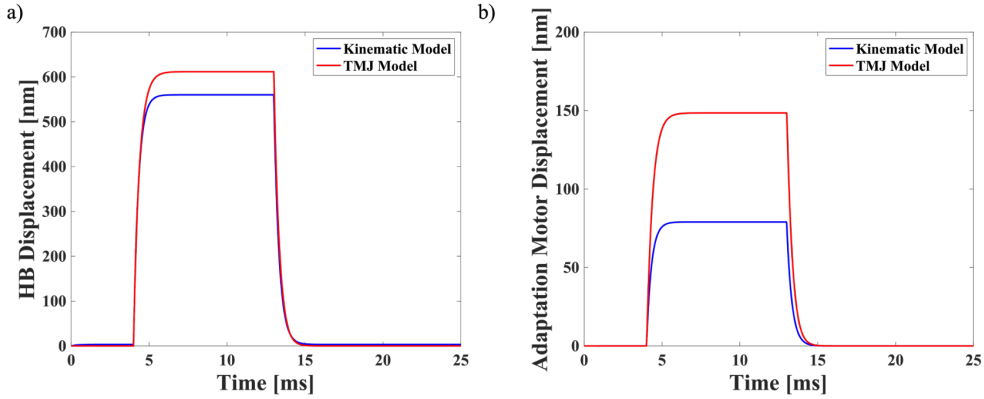


FIGURE 5. A comparison between the responses obtained for a) the HB displacement, and b) the adaptation motor motion from the TMJ model ($X, \gamma X_a$) and the new kinematic and geometric HB model ($r_{CA}\phi, \delta_{a1}$). In both the models, static force $F_{st} = 800$ pN, $F_o = 1700$ pN, and $\tau_F = 0$ ms. The bundle displacement from our model is marginally lower than the modified TMJ model, therefore, both models predict a biologically valid displacement. The motion of the adaptation motor in a stereocilium is experimentally unknown. Consequently, we cannot conclude whether the displacements (γX_a and δ_{a1}) predicted are realistic or not yet.

First, we observed similar trends (not shown in this paper) from our model as in Fig. 3 from the TMJ model for varying τ_F. We then compared the responses with the TMJ model by setting $\tau_F = 0$ ms and found that the HB displacement is lower for the same set of parameters (except $A_1 \neq A$ and $N_1 \neq N$ in some cases). For instance, for a static force of 800 pN ($P_o^e \approx P_{o1}^e \approx 5\%$), $F_o = 1700$ pN, and $\tau_F = 0$ ms, the steady-state displacement from the TMJ model along the x-axis is 611.6 nm, whereas that from our model is 560.1 nm (an 8% reduction). There is a significant difference in the saturation values for the adaptation motor motion (δ_{a1} and γX_a) along row 1 (γ is the geometrical gain [6]). For the same static and external force conditions, the TMJ model predicted the value (for adaptation motor displacement) as 149 nm, while our model estimated this value as 79 nm (a 47% lower value). It is unknown at this moment whether these predictions (from both the models) are realistic or not as there is no experimental data available for validation. We are currently working on parameter (A_1, N_1, K_{GS_1}, and d_1) optimization to correspond the currents and bundle displacements predicted from the model to those obtained experimentally [8].

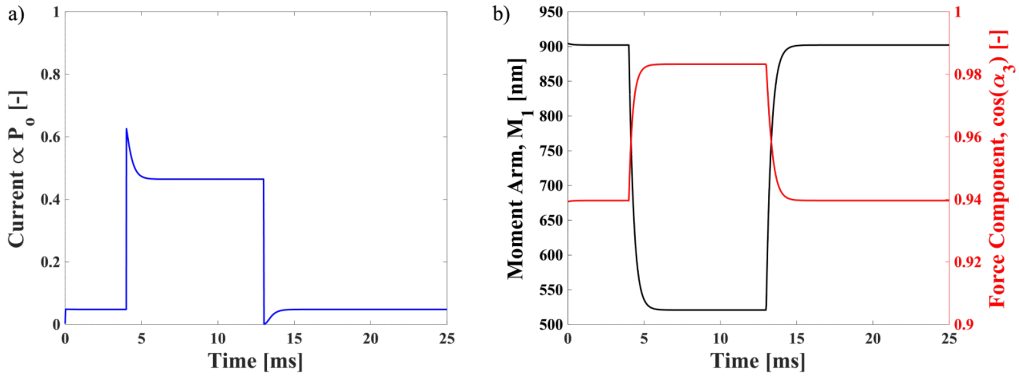


FIGURE 6. A description of the nonlinearities in our model. The nonlinear behavior of a) the two-state Boltzmann function for defining the channel open probability/influent current, and b) the moment arm $M_1 = (l_1^{gs} - \delta_{a1})\sin(\alpha_3) + r_1\cos(\alpha_3)$ and force component $\cos(\alpha_3)$. The results correspond to static force $F_{st} = 800$ pN, $F_o = 1700$ pN, and $\tau_F = 0$ ms. The moment arm nonlinearity decreases by $\approx 42\%$ at saturation from its resting value, whereas the force component increases by $\approx 4.6\%$ at saturation from its resting value. Therefore, the change in these factors seem substantial and they may affect the prediction of the responses.

Our model introduces dual nonlinearity in the system, one from the two-state Boltzmann probability function (Eq. 9) and second from the nonlinear moment arm (M_1) and force component ($\cos(\alpha_3)$) of the gating force in Eq. 10 and Eq. 11, respectively, as opposed to just the Boltzmann probability (Eq. 3) in the TMJ model. It is evident from Fig. 6 b) that the length of the moment arm decreases as the HB rotates in the excitatory direction. Fig. 6 b) also shows how $\cos(\alpha_3)$ increases due to a decrease in α_3 with stereociliary rotation. The nonlinearity of a force follower (moment arm here) is well-known in dynamic systems, therefore, its inclusion in our model might be critical to bundle displacement, current responses, and the operating range. However, it is yet to be determined precisely.

CONCLUSION

We simulated the model developed by Tinevez *et al.* [1] with a time-varying exponential force for numerous static forces. For resting channel open probabilities between 5% and 40%, we found that HB motion saturated to steady-state with a delay, the current peak declined and shifted rightwards, and the apparent OR broadened as τ_F was increased. We found, using the TMJ model, that the force rise time is important and alters the predicted OR, as shown by Nam *et al.* [10]. There is a close qualitative correspondence in the response trends (HB displacement, current, and the OR) observed by Nam *et al.* [10] in their model and experiments to those predicted by a simpler TMJ model. Consequently, we were motivated to develop a kinetic model with more precise kinematics of two stereocilia rows of the HB.

Including more precise kinematic and geometric constraints in the model predicted lower bundle displacement ($\approx 8\%$) and adaptation motor displacement ($\approx 47\%$) than the TMJ model for the same applied force. With a more precise geometric and kinematic model, we are now working towards optimal estimation of the model parameters. Our focus is also on assessing how important the geometric nonlinearities (M_1 and $\cos(\alpha_3)$) in equations of motion are compared to the constitutive nonlinearity (Eq. 9). Finally, we are developing and validating our model with available experimental data in the literature to gain better insights into the responses and the operating range with mathematically assessed model parameters.

ACKNOWLEDGMENTS

We acknowledge the support of the National Institutes of Health grant- NIH-NIDCD R01 04084.

REFERENCES

1. J.-Y. Tinevez, F. Jülicher, and P. Martin, "Unifying the various incarnations of active hair-bundle motility by the vertebrate hair cell," Biophysical journal **93**, 4053–4067 (2007).
2. H. Kennedy, A. Crawford, and R. Fettiplace, "Force generation by mammalian hair bundles supports a role in cochlear amplification," Nature **433**, 880–883 (2005).
3. C. D. Geisler, *From sound to synapse: physiology of the mammalian ear* (Oxford University Press, USA, 1998).
4. W. Denk, J. R. Holt, G. M. Shepherd, and D. P. Corey, "Calcium imaging of single stereocilia in hair cells: localization of transduction channels at both ends of tip links," Neuron **15**, 1311–1321 (1995).
5. M. Beurg, J.-H. Nam, A. Crawford, and R. Fettiplace, "The actions of calcium on hair bundle mechanics in mammalian cochlear hair cells," Biophysical journal **94**, 2639–2653 (2008).
6. J. Howard and A. Hudspeth, "Compliance of the hair bundle associated with gating of mechanoelectrical transduction channels in the bullfrog's saccular hair cell," Neuron **1**, 189–199 (1988).
7. Y. Choe, M. O. Magnasco, and A. Hudspeth, "A model for amplification of hair-bundle motion by cyclical binding of ca2+ to mechanoelectrical-transduction channels," Proceedings of the National Academy of Sciences **95**, 15321–15326 (1998).
8. R. Fettiplace and K. X. Kim, "The physiology of mechanoelectrical transduction channels in hearing," Physiological reviews **94**, 951–986 (2014).
9. C. L. Cunningham, X. Qiu, Z. Wu, B. Zhao, G. Peng, Y.-H. Kim, A. Lauer, and U. Müller, "Tmie defines pore and gating properties of the mechanotransduction channel of mammalian cochlear hair cells," Neuron **107**, 126–143 (2020).
10. J.-H. Nam, A. W. Peng, and A. J. Ricci, "Underestimated sensitivity of mammalian cochlear hair cells due to splay between stereociliary columns," Biophysical journal **108**, 2633–2647 (2015).
11. B. Nadrowski, P. Martin, and F. Jülicher, "Active hair-bundle motility harnesses noise to operate near an optimum of mechanosensitivity," Proceedings of the National Academy of Sciences **101**, 12195–12200 (2004).
12. D. Mead, "Elements of vibration analysis, leonard meirovitch," McGraw-Hill Book Company, New York. Journal of Sound and Vibration **117**, 603–604 (1987).

A nonlinear mechano-electro-acoustic model of the human cochlea

Naman Agarwal[*b], Sripriya Ramamoorthy[*a]

[*]*Department of Mechanical Engineering, Indian Institute of Technology Bombay, Mumbai, Maharashtra, India*
[a)]Corresponding author: ramamoor@iitb.ac.in
[b)] 154100025@iitb.ac.in

Abstract. While direct measurements of the basilar membrane (BM) and auditory nerve fiber tuning are possible in lab animals [1] but no psychoacoustic measurements (with rare exceptions [2]), only otoacoustic emission and psychophysical tuning curve measurements [2,3,4] are possible in humans. In an attempt to partly address this missing link, the three-dimensional physiology-based nonlinear mechanical-electrical-acoustic (MEA) model of the human cochlea is developed in this study to enable prediction of both otoacoustic emissions as well as BM and organ of Corti vibration tuning. The nonlinear MEA model of the human cochlea adapts our prior linear MEA model of the guinea pig cochlea [5,6] to humans and includes the nonlinearity due to mechanoelectrical transduction in hair bundles. The transmission line human middle ear model, inspired by O'Connor and Puria [7], is integrated with the nonlinear MEA model. The passive BM displacement predicted by the model is compared with the *in vivo* physiological measurements for the BM displacement in human cadavers [8], and the predicted active BM tuning curves are compared with human psychophysical tuning curves (PTC) from Moore (1978) [9]. The distortion product otoacoustic emissions (DPOAE), predicted by the model at equal stimulus levels and at different stimulus frequencies, have a maximum at around f2/f1 = 1.2, agreeing with measured data in the literature [10,11].

INTRODUCTION

Nonlinear compressive growth is a characteristic of healthy mammalian cochleae [12], which enables the stimulation of distortion products (DPs). The mammalian cochlea also demonstrates high sensitivity, sharp tuning, and tonotopic frequency separation of the incoming sounds along the basilar membrane (BM) [12]. Consequently, each sensory hair cell responds to sound within the limited frequency range due to this frequency separation or filtering. Owing to its role as a frequency analyser, the cochlea has been represented as filter banks [13], also referred to as cochlear filters [12]. The frequency tuning of these filters signifies our ability to distinguish sounds. The loss of cochlear tuning and deficits in the ability to extract the correct sound from background noise [3] are associated with hearing loss and broadening the cochlear filters. A physiological study based on otoacoustic emissions indicates that human cochlear filters may be sharper than other species [3]. However, this view has been challenged by Ruggero et al. (2005) [4] where it is argued that the sharpness of the human tuning curve is not exceptional, and that it is similar to other mammals. In Ruggero et al. (2005) [4], an equal tuning in humans is proved using a compound action potential tuning curve against auditory nerve fibers tuning curve. The sharpness of cochlear tuning in humans, as well as several other aspects of human cochlear responses are unclear yet. The primary challenge is that, so far, only otoacoustic emissions [12,14] and psychophysical tuning curve measurements [2,9] are possible in humans.

DPOAE is a robust phenomenon routinely used for clinical purposes. The intermodulation of two tones in the cochlea generates the acoustic distortion product (ADP), 2f1-f2, measured in the ear canal of animals and humans. DPOAEs of humans have been measured for a wide range of stimulus frequencies and levels [10]. DPOAE level had

a band-pass filtering characteristic is observed in humans and rodent [10], when the frequency of one primary tone was fixed and other was swept.

The goals of the present study are to construct a nonlinear human MEA model and to estimate the model's frequency domain response to single and multitoned stimuli, particularly for simulation of DPOAEs. DPOAEs for a wide range of primary levels and primary frequencies were simulated so results can be compared to data reported by Gaskill and Brown (1990) [10]. The nonlinear MEA model of the human cochlea adapts our prior linear MEA model of the guinea pig cochlea [5, 6] to humans and includes the nonlinearity due to mechanoelectrical transduction in hair bundles. Inspired by O'Connor and Puria [7], the transmission line human middle ear model is also integrated with the nonlinear MEA model.

METHOD

Nonlinear Mechano Electro Acoustical (MEA) Model: In this article, nonlinearities are incorporated into a previously developed quasi-linear mechano-electro acoustical (MEA) model of the cochlea (Fig 1) [5, 6]. This physiologically motivated finite element model consists of three physical domains (mechanical, electrical, and acoustical domain), integrated along with their coupling terms. The three-dimensional intracochlear fluid dynamics are coupled to a micromechanical model of the organ of Corti (OoC) and to the electrical potentials of outer hair cells (OHC) and the cochlear ducts. The model includes three structural degrees of freedom at each cross-section: BM, HB, and OHC, as shown in Fig 1. To incorporate cochlear nonlinearities into the model, the HB mechano-electrical transduction current (MET current) is modelled by a first-order Boltzmann nonlinear function. The MET current of HB can be written as

$$I_{hb} = G_a^{max}(P_0 - P_0^s)\Delta V_{hb}$$

Where G_a^{max} is the maximum saturating HB mechano-electric conductance, V_{hb} is the resting value of the difference between scala media potential and intercellular OHC potential,

$$P_0 = \frac{1}{1 + e^{\left(\frac{X_0 - u_{hb}}{\Delta X}\right)}}$$

P_0^s is the resting open probability of the MET channel, X_0 and ΔX are displacement constants, u_{hb} is the HB displacement in the time domain. The nonlinear governing equations of cochlear dynamics are discretized using Bubnov-Galerkin finite element method and shown in the following matrix form

$$\begin{bmatrix} K_f & Q_{fs} & 0 \\ Q_{sf} & K_s & Q_{se} \\ 0 & Q_{es} & K_e \end{bmatrix} \begin{bmatrix} p_m \\ u_m \\ \emptyset_m \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ NL_m(u) \end{bmatrix} = \begin{bmatrix} f_m \\ 0 \\ 0 \end{bmatrix}$$

The quantities K_f, K_s and K_e are the dynamic stiffness of the fluid, the micromechanical structure, and the electrical degrees of freedom, respectively. Q_{sf} and Q_{fs} are from fluid-structure coupling at the BM. Q_{se} and Q_{es} represent the electrical-structural coupling at the OHC. NL_m is the nonlinear force vector due to the MET current. f_m is the forcing term on the fluid due to stapes, which is attached to the middle ear model, adopted from O'Conner *et al.* (2008) [7] (shown in Fig 1). m is the index to map the frequency components eg. m=1,2 for primary frequencies f_1 and f_2, respectively and 3, 4 for DP frequency components ($2f_1$-f_2 and $2f_2$-f_1) and so on. To determine the nodal value of fluid pressure, p_m structure displacement u_m, and electrical potential $\emptyset_m$,

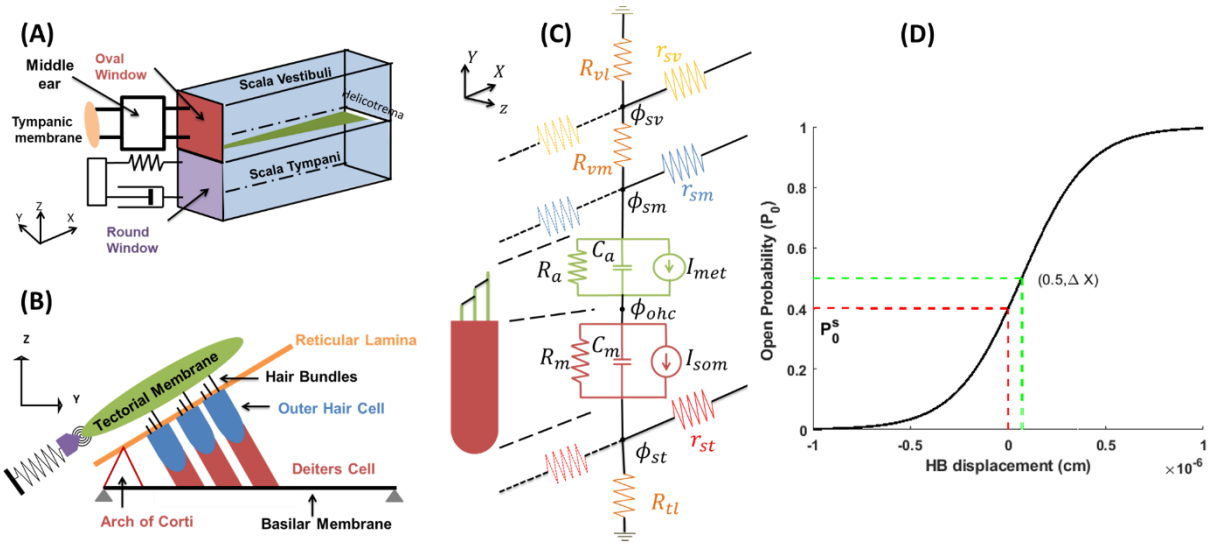


FIGURE 1 Nonlinearities incorporated into the mechano-electroacoustical (MEA) model of the cochlea (Ramamoorthy et al. (2007) [5], Agarwal and Ramamoorthy (2020) [6]). **(A)** Fluid model. SV: Scala Vestibuli ST: Scala Tympani, OW: Oval Window and RW: Round Window **(B)** Micromechanical model for the OoC structures. BM: Basilar Membrane, TM: Tectorial Membrane, RL: Reticular Lamina, OHC: Outer hair cell, DC: Deiter Cell. **(C)** Electrical network at given cross section of cochlea. The SV, SM and ST cables run longitudinally along the cochlea. SM: Scala Media. r_{SV} ,r_{Sm} and r_{St} are the resistance per unit length along the SV, SM and ST cables respectively. R_{V1} and R_{tl} are the resistance to current flowing to ground from SV and ST respectively. R_{Vm} is the resistance seen by current flowing from SV to SM. R_{a0} and C_a are the apical resistance and capacitance while R_m and C_m are the basolateral resistance and capacitance respectively of the OHCs. I_{met} and I_{som} are the current sources due to HB conductance and OHC electromotility respectively. **(D)** Open probability of MET channel P_0 as a function of HB displacement u_{hb} modeled as a first order Boltzmann function. the slope of open probability is given as $P_0(1- P_0)/\delta X$. Choses resting probability $P_0^s = 0.4$.

Human middle ear model: In this work, the transmission line human middle ear model, shown in Fig 2, inspired by O'Connor and Puria (2008) [7], is integrated with the nonlinear MEA model. The design of the middle ear cavity is adapted from Voss *et al* (2000) [15], and the cavity is thereby modeled as two acoustic volume compliances, C_{B1} and C_{B2}, separated by a narrow passage of acoustic inertia, L_B, and resistance, R_B. For forward transmission, the cochlear input impedance acts as the load which is computed using the predicted Scala Vestibule (SV) pressure and stapes volume velocity. In the reverse direction, rigid termination is assumed at the end of the ear canal, mimicking experimental conditions used for validation. The forward pressure transfer function (SV pressure per unit input pressure at the ear canal) and the reverse pressure transfer function (pressure measured at ear canal by a microphone per unit SV pressure) have been validated with available measured data for human, Puria (2003) [16]. DPOAE is predicted as the ear canal pressure resulting from the propagation of DP from the cochlea to the ear canal via the middle ear.

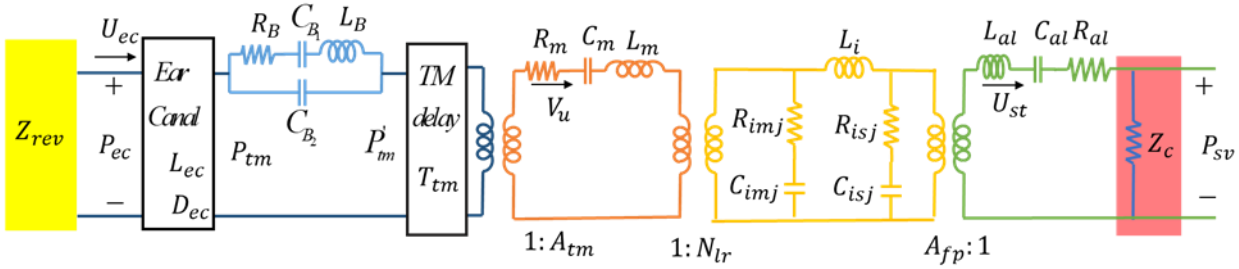


FIGURE 2 Transmission line human middle ear model, inspired by O'Connor and Puria (2008) [7]. In this model, the ear canal (EC) is assumed to be of cylindrical cross-section with A_{ec} the area of EC and l_{ec} the length of EC from the tip of the probe-tube microphone (used in DPOAE experiments) to the tympanic membrane. The tympanic membrane (TM) is modeled as a lumped uniform, lossless transmission line element, with characteristic impedance Z_{tm} and delay T_{tm}. A_{tm} is the effective area of the TM. The malleus is represented as mechanical resistance R_m, and compliance C_m. The mass of the malleus and incus are L_m and L_i

,respectively, with turns ratio N_{lr}. The incudo-malleus joint is modeled as mechanical resistance R_{imj} and compliance C_{imj}. The incudo-stapedial joint is modeled as mechanical resistance R_{isj} and compliance C_{isj}. A_{fp} is the effective area of the stapes footplate. The middle-ear cavities are modeled as two acoustic volume compliances, C_{B1} and C_{B2} , separated by a narrow passage of acoustic inertia, L_B, and resistance, R_B. R_{al}, C_{al}, and M_{al} are the equivalent resistance, compliance and mass of the complex formed, respectively, by the annular ligament, stapes, and cochlea. C_{al} is mostly due to the annular ligament and L_{al} is due to the mass of the stapes and some cochlear fluid proportions. Z_C is mainly due to the acoustic input resistance of the cochlea. Z_{rev} is reverse impedance at ear canal when sound is emitted from the cochlea to ear canal. P_{ec} and P_{sv} are the ear canal pressure and scala vestibule pressure, respectively.

RESULTS

The human nonlinear MEA model discussed in the METHODS section is used to predict the BM and organ of Corti displacements, SV and ST fluid pressures, cochlear microphonic potentials and OHC intracellular potential, and the middle ear responses, to sound input at the ear canal. Both pure tone as well as multi-tone stimuli are considered as discussed next.

Figure 3 (A) and (C) show the variation of BM response as a function of the stimulus frequency at 14.4 kHz best place for 108 dB sound pressure level with experimental response in human cadavers from Raufer *et al.* (2019) [8]. In the magnitude plot, the BM gain relative to ear canal pressure is similar as the experiment response around the CF. However, at low frequencies, the model response differs from the experiment. Similar match is seen in the phase responses too. This mismatch points to the need to improve the model parameters for the human cochlea. Indeed, the primary challenge faced is the unavailability of parameters for the human cochlea model. This challenge is partly addressed here by adapting the parameters of the guinea pig cochlea used in Agarwal and Ramamoorthy (2020) [6] to the human cochlea using the tonotopic scaling method discussed in [17]. Figure 3 (B) and (D) show the variation of normalized BM response at 5 kHz best place relative to stapes, as a function of the stimulus frequency for different sound pressure levels. At low sound pressure levels, the normalized BM response curves are sharply tuned with a steep high-frequency slope near the characteristic frequency (CF), while at high sound pressure levels, the responses are broadly tuned as the peak response shifts toward lower frequencies. BM phase responses to tones exhibit lagging as a function of increasing frequency. Phase curve (Fig 3 (D)) has a distinct shallow segment at low frequency and steep segment at a frequency around CF. At frequencies lower than CF, the model response (BM magnitude or phase) are insensitive to alteration of ear canal pressure, which is consistent with experimental observation of other animals [12]. At CF, the BM gain is higher at low intensity than at high intensity, indicative of a compressive nonlinearity.

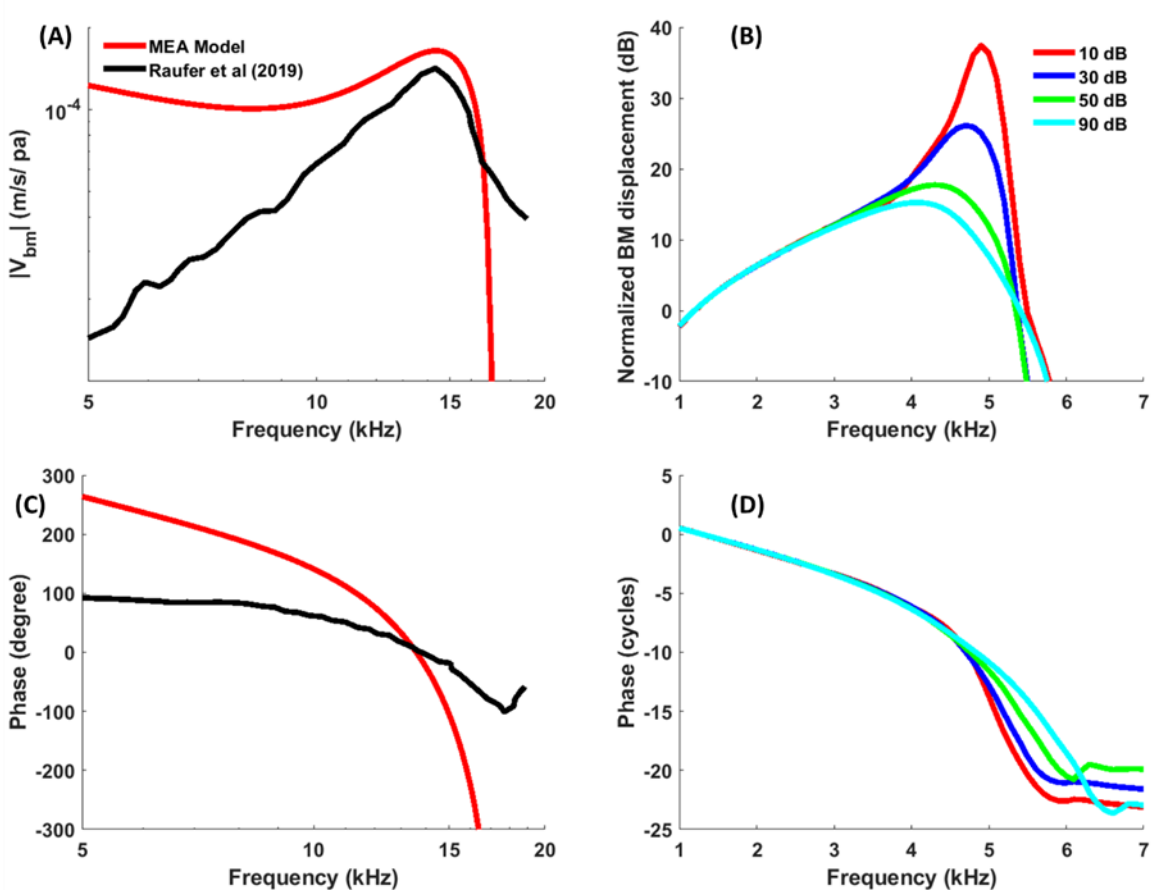


FIGURE 3 (A) Magnitude and **(C)** Phase of BM velocity per unit stimulus pressure in frequency domain at 14.4 kHz BP and compared with experiment results [8]. **(B)** Magnitude and **(D)** Phase of normalized BM with stapes displacement in frequency plot at 5 kHz BP with different stimulus level. The response decreases with increasing stimulus level, as assumed.

Whether human cochlear tuning is sharper than other mammals is being debated in the literature. Ruggero *et al* (2005) [4] suggest that the sharpness of human cochlear tuning is similar to other mammals based on measurements of compound action potential tuning. However, Shera *et al* (2002) [3] suggest human cochlear filters are sharper than other mammals using otoacoustic emission measurements. Since direct measurements of intracochlear vibrations in humans is not yet possible, the cochlear tuning sharpness is explored in this study using the nonlinear human MEA model. The iso-displacement tuning curve predicted by the model is compared with the psychoacoustic tuning curve (PTC) measured in human subjects [9] at three different frequencies, 2, 4, 8 kHz. The amplitude of iso-displacement at high frequency concurs with the PTC. However, there is again a disagreement at the low-frequencies, which could be due to the imprecise parameters used in the model.

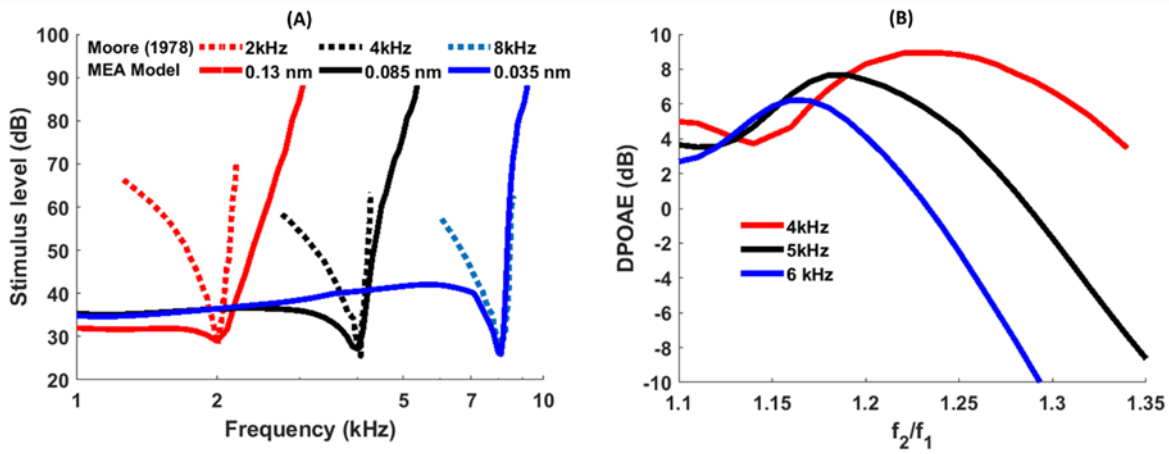


FIGURE 4 (A) Iso displacement curve estimated from nonlinear MEA model of human and compared with PTC, measured by Moore (1978) [9]. The iso intensity curve qualitatively similar to PTC. **(B)** Cubic DPOAE estimated from model for three different frequencies (f1 = 4 , 5 and 6 kHz) at 55 dB SPL (L1=L2). The DPOAE response peaks around 1.2 to 1.22 and decreases with increasing frequency ratio, as expected from Gaskill and Brown 1990 [10].

The level of $2f_1 - f_2$ cubic DPOAE at the ear canal is predicted as a function of the stimulus frequency separation (ratio = f2/f1) using the human nonlinear MEA model. In each simulation sequence with a constant f1 frequency and variable f2 frequency, the ratio was increased from 1.1 to 1.35 in 0.01 increments. The stapes displacement is predicted at the DP frequency and converted into ear canal pressure using the reverse middle ear path (refer to Fig 2). The effect of frequency separation is estimated at three different f1 frequencies at four different stimulus levels (L1=L2). Figure 4 (B) shows how the DPOAE response pattern changes with the frequency separation for a constant stimulation level of 55 dB SPL. The DPOAE response increases with frequency separation until an optimum f2/f1. Beyond that value, the response level declines with further increase in f2/f1. In this simulation, the primary frequencies f1 are taken to be 4, 5, and 6 kHz with variable f2/f1 for 55 dB SPL input at the ear canal. Figure 4 (B) shows that the optimum ratio lies between 1.15 to 1.2 for three f1 frequencies, similar to human DPOAE measurements shown in Gaskill and Brown 1990 [10]. At the optimum f2/f1, the predicted cubic DPOAE level is 50 dB below the level of the primaries, similar to experimental observations [10]. From the literature, it may be noted that the optimal ratio changes from subject to subject, as well as with age. For neonates, the ratio could vary from 1.2 to 1.23, while for adult subjects, it could vary from 1.18 to 1.2 [11]. The band-pass response with frequency separation is similar to different stimulus frequencies (shown in the Fig 4 **Error! Reference source not found.**(B)) [10] for the equal stimulus level (L1 = L2).

CONCLUSION

The primary objective of this study to design a physiology-based human cochlea model. The mechano-electro acoustical, previously developed by Ramamoorthy et al. (2007) [5] (also see [6]), is integrated with MET transduction nonlinearity and modified with human cochlea parameters to predict sound induced response within the hearing range (20 Hz to 20kHz). A transmission line middle ear model is combined with the nonlinear MEA model in order to predict DPOAE. The model is first benchmarked with passive intra-cochlear response measured in human cadavers [8]. The predicted BM response matches with measured data well around the CF but deviated at lower frequencies. This difference points to the need to improve the human nonlinear MEA model parameters. Except for a few parameters taken from literature for human cochlea, most of the model parameters are extrapolated from guinea pig model for the 20 Hz to 20 kHz frequency range using the tonotopic scaling principle. Except for the differences at low frequencies, the model predicts the expected basic characteristics of cochlea (such as amplification at low stimulus levels, frequency selectivity, compressive nonlinearity) similar to other mammals [12].

The predicted normalized BM response relative to stapes peaks at CF for low stimulus levels and the peak frequency decreases with increasing stimulus level. At frequencies away from the CF, the response is almost independent of stimulus level. The BM iso-displacement curve estimated by the model agrees qualitatively with PTC measurements in human subjects. With further improvement in the model parameters, and better agreement with PTC over a wider frequency range, this model could be used to help address the sharpness of human cochlear tuning.

In response to two tone sound stimulus, nonlinearities in the cochlea cause intermodulation distortion of the sound-induced responses within the OoC. These distortion products can be measured at the ear canal as DPOAE. The cubic DPOAE (or the 2f1 – f2 component) predicted for varying primary frequency ratio (f2/f1) demonstrates a band-pass shape with a maximum at f2/f1 = 1.2 for f1 = 5 kHz. The model shows that the f2/f1 value at which DPOAE is maximum varies with f1. This trend is qualitatively similar to experimental observations [10] but differs in the quantitative value of the primary frequency ratio.

REFERENCES

1. Ruggero, M. A., Narayan, S. S., Temchin, A. N., & Recio, A. (2000). Mechanical bases of frequency tuning and neural excitation at the base of the cochlea: comparison of basilar-membrane vibrations and auditory-nerve-fiber responses in chinchilla. *PNAS*, *97*(22), 11744-11750. https://doi.org/10.1073/pnas.1810766115.
2. Sumner, C. J., Wells, T. T., Bergevin, C., Sollini, J., Kreft, H. A., Palmer, A. R., ... & Shera, C. A. (2018). Mammalian behavior and physiology converge to confirm sharper cochlear tuning in humans. *PNAS*, *115*(44), 11322-11326. https://doi.org/10.1073/pnas.1810766115.
3. Shera, C. A., Guinan, J. J., & Oxenham, A. J. (2002). Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. *PNAS*, *99*(5), 3318-3323. https://doi.org/10.1073/pnas.032675099.
4. Ruggero, M. A., & Temchin, A. N. (2005). Unexceptional sharpness of frequency tuning in the human cochlea. *PNAS*, *102*(51), 18614-18619. https://doi.org/10.1073/pnas.0509323102.
5. Ramamoorthy, S., Deo, N. V., & Grosh, K. (2007). A mechano-electro-acoustical model for the cochlea: response to acoustic stimuli. *JASA*, *121*(5), 2758-2773. https://doi.org/10.1121/1.2713725.
6. Agarwal, N., & Ramamoorthy, S. (2020). Balance in the feedback loop components of the mammalian cochlear amplifier. *JAP*, *128*(3), 034701. https://doi.org/10.1063/5.0010802.
7. O'Connor, K. N., & Puria, S. (2008). Middle-ear circuit model parameters based on a population of human ears. *JASA*, *123*(1), 197-211. https://doi.org/10.1121/1.2817358.
8. Raufer, S., Guinan, J. J., & Nakajima, H. H. (2019). Cochlear partition anatomy and motion in humans differ from the classic view of mammals. *PNAS*, *116*(28), 13977-13982. https://doi.org/10.1073/pnas.1900787116
9. Moore, B. C. (1978). Psychophysical tuning curves measured in simultaneous and forward masking. *JASA*, *63*(2), 524-532. https://doi.org/10.1121/1.381752.
10. Gaskill, S. A., & Brown, A. M. (1990). The behavior of the acoustic distortion product, 2 f 1− f 2, from the human ear and its relation to auditory sensitivity. The Journal of the Acoustical Society of America, 88(2), 821-839. https://doi.org/10.1121/1.399732.
11. Abdala, C. (1996). Distortion product otoacoustic emission (2 f 1− f 2) amplitude as a function of f 2/f 1 frequency ratio and primary tone level separation in human adults and neonates. The Journal of the Acoustical Society of America, 100(6), 3726-3740. https://doi.org/10.1121/1.417234.
12. Robles, L., & Ruggero, M. A. (2001). Mechanics of the mammalian cochlea. Physiological reviews, 81(3), 1305-1352. https://doi.org/10.1152/physrev.2001.81.3.1305.
13. Lopez-Poveda, E. A., & Meddis, R. (2001). A human nonlinear cochlear filterbank. The Journal of the Acoustical Society of America, 110(6), 3107-3118. https://doi.org/10.1121/1.1416197 .
14. Kemp, D. T. (2002). Otoacoustic emissions, their origin in cochlear function, and use. *British medical bulletin*, *63*(1), 223-241. https://doi.org/10.1093/bmb/63.1.223
15. Voss, S. E., Rosowski, J. J., Merchant, S. N., & Peake, W. T. (2000). Acoustic responses of the human middle ear. Hearing research, 150(1-2), 43-69. https://doi.org/10.1016/S0378-5955(00)00177-5
16. Puria, S. (2003). Measurements of human middle ear forward and reverse acoustics: implications for otoacoustic emissions. *The Journal of the Acoustical Society of America*, *113*(5), 2773-2789. https://doi.org/10.1121/1.1564018
17. Greenwood, D. D. (1990). A cochlear frequency-position function for several species—29 years later. *The Journal of the Acoustical Society of America*, *87*(6), 2592-2605. https://doi.org/10.1121/1.399052.

Interpeak characterizations for spontaneous otoacoustic emissions

Christopher Bergevin,[1, a)] Rebecca Whiley,[2, b)] Hero Wit,[3, c)] and Pim van Dijk[3, d)]

[1)]*Department of Physics & Astronomy, York University, Toronto, ON Canada*
[2)]*Department of Biology, York University*
[3)]*University of Groningen, University Medical Center Groningen, Department of Otorhinolaryngology/Head and Neck Surgery, The Netherlands*

[a)]*Corresponding author: cberge@yorku.ca*
[b)]*Electronic mail: rwhiley@yorku.ca*
[c)]*Electronic mail: hero.wit@ziggo.nl*
[d)]*Electronic mail: p.van.dijk@umcg.nl*

Abstract. One manifestation of the "active ear" is the presence of spontaneous otoacoustic emission (SOAE), which also exhibit salient connections to perception such as threshold microstructure. Historically, SOAE modeling efforts initially focused upon a single limit-cycle oscillator. However, SOAE spectra from a given ear typically exhibit multiple peaks, and more current models consider a spatially-distributed tonotopic system with various types of coupling. SOAEs have nonstationary properties (e.g., amplitude and frequency modulations), which may be crucially tied to the coupling of active elements in the ear. Thus to better biophysically constrain models, this study seeks to improve characterization of general non-stationary features of SOAE peaks as well as interrelations of such between them. Given the ubiquitous nature of SOAE across the animal kingdom, we analyze SOAE waveforms from a variety of species exhibiting disparate inner ear morphologies (e.g., human, barn owls, Anolis lizards). This manuscript provides a preliminary account of our analyses and focuses on the *Anolis* lizard. Upon filtering in the spectral domain, we characterize temporal properties of individual peaks, including possible amplitude-modulation (AM) and frequency-modulation (FM). Further, we perform correlation analyses of such between peaks to determine types of interactions and how such might vary across time. Initial results are consistent with previous reports (e.g., [1, 2]) in that an SOAE interpeak correlations for a given ear are idiosyncratic: Sometimes peaks (adjacent or not) exhibit correlated (positive or negative) AM and/or FM fluctuations with delays up to the order of milliseconds (typically longer for humans, shorter for lizards), while sometimes no correlation is observed. We attempt to frame these results within the broader context of specific SOAE modeling approaches.

BACKGROUND

A common feature of the healthy ear across the animal kingdom is the generation of spontaneous otoacoustic emission (SOAE). This phenomenon is often described as a by-product of an underlying active mechanism that metabolically boosts the sensitivity and selectivity of the ear. While many theories have been proposed (e.g., [3, 4, 5, 6]), SOAE generation remains relatively poorly understood, especially when considering gross inner ear morphological differences across the animal kingdom [7].

Several key characteristics of SOAE are commonly observed. First, not all ears emit. A healthy individual can have normal hearing but exhibit no SOAE. Further, the presence of SOAE is common but not universal across the animal kingdom: While relatively robust in primates and numerous lizard species, SOAE activity is conspicuously absent in many animals commonly used in auditory neurophysiology. For example, SOAE in mice are exceedingly rare unless mutations to their tectorial membrane are present [8]. Second, SOAE activity is mostly confined to the most sensitive portion of the audiogram. Third is the general characteristic that SOAE commonly manifest as peaks of variable width in the spectrum of the measured microphone signal (hence "SOAEs" from an ear). This narrowband feature has provided a focal point for many facets of SOAE analysis and theory, despite the fact SOAE activity can also be present in a broadband fashion (e.g., the "baseline" activity in geckos [9] and skinks). Fourth, these peaks can (but not always) exhibit statistical properties consistent with self-sustained sinusoids [10, 11, 12]. Lastly, SOAE activity appears to readily interact with external acoustic stimuli, allowing for measurements such as "suppression tuning curves" that can exhibit selectivity similar to that of single auditory nerve fibers (e.g., [13]).

To help constrain theoretical models, the present study focuses on characterizing the distinctive set of SOAE peaks that an individual ear can exhibit. Assuming that hair cells are a key element for SOAE generation (i.e., they act as a primary unit of force generation), and that the inner ear is a collection of coupled hair cells, we seek to elucidate how

those elements interact by virtue of characterizing SOAE interpeak relationships (if any). We do this by examining both stationary and non-stationary properties (e.g., amplitude and/or instantaneous frequency fluctuations) for SOAEs from an individual ear. It is worth noting from the outset that animals with evident interaural coupling can exhibit SOAE correlations between both ears [14], such that SOAE-related interactions need not be confined to the inner ear alone.

While the goal of the present study is to ultimately characterize interpeak relationships across a broad range of terrestrial vertebrate ears (e.g., human, barn owl, lizard species with variable tectorial membrane morphology), this manuscript focuses on preliminary results from the *Anolis carolinensis* lizard. There are several advantages for such a choice of *Anolis*: • robust SOAE • short auditory papilla (0.5 mm) with $\sim$150 hair cells • most hair cells lack an overlying tectorial structure [15, 16]. These features will presumably make it easier to interpret the otoacoustics and how such relates to the biomechanics of the inner ear.

Approach

Our approach to examining SOAE interpeak relations is two-fold. First, we quantify stationary frequency relationships between peaks, as apparent in their (relatively stable) spectra. Such includes their relative frequency spacing (e.g., N_{SOAE}, as per [4, 12]) and ratios to examine if peaks correspond to distortions (e.g., cubic intermodulations, [17]) or harmonics of other peaks. This latter facet may be particularly informative, given the bi-directional orientation of hair cell bundles along the relevant tonotopic axis of the lizard papilla [18, 19]. The second is more involved and focuses on non-stationary properties of SOAE peaks [1, 2]. For example, SOAE peaks exhibit both amplitude (AM) and frequency modulations (FM), as evident in their filtered responses. Presumably these facets contribute to the variable width of the SOAE peaks themselves, and could in principle arise biomechanically from stochastic effects (e.g., thermal noise) and/or due to coupled dynamics of active elements. Previous studies have examined interpeak AM correlations, finding them in a relatively small fraction (i.e., $15-20\%$) of (intra-ear) SOAE pairs for human [20], as well as for lizards and owls [2]. We seek to expand upon those analyses, hypothesizing that a greater degree of correlation may be present due to the highly (mechanically) coupled nature of the inner ear elements.

METHODS & RESULTS

Interpeak Frequency Relationships

Otoacoustic emissions were measured in lightly anesthetized *Anolis carolinensis* lizards, consistent with previous descriptions [12, 21]. Representative SOAE data from six lizards are shown in Fig.1. For each, peak frequencies were identified by virtue of being at least 2 dB above the noise floor and/or surrounding frequency bins. In all cases, SOAE activity was modulated by the presence of a nearby external tone (e.g., a 30 dB SPL tone would reduce SOAE activity in a localized region; not shown).

Preliminary stationary analysis of these spectra are shown in Fig.2, described in more detail in the figure caption. Note that there is a slight redundancy in checking for intermodulation relationships. That is, assuming $f_1 < f_2 < f_3$, then there is ambiguity as to whether f_1 is $2f_2 - f_3$ or if f_3 is $2f_2 - f_1$. This is addressed in the present analysis by dividing the total counts by two (Fig.2C). Also note that the averaged spectra can exhibit relatively small peaks that did not pass our threshold for inclusion (i.e., at least 2 dB above the nearby noise floor). Time waveforms were collected during experiments that could allow for further analysis of these excluded peaks. Additionally, future analysis will allow for additional (basic) features of SOAE peaks to be objectively characterized, such as peak height versus width [22].

From Fig.2, we can deduce several basic conclusions. First, Fig.2A suggest that there is a relatively low probability that an SOAE occurs at twice the frequency of a lower frequency peak. More typically, the interpeak ratio is smaller and more likely to be close to unity. This observation suggests interpeak frequency spacing is not predominately due to an integer ratio, including the second harmonic. Such may be in contrast to intracellular hair cell potentials [19], as previously predicted [18]. Second, the reciprocal of the interpeak spacing could be interpreted as a delay (e.g., a condition for standing wave interference). As plotted in Fig.2B, these values tend to be on the order of several milliseconds, consistent with delays observed for stimulus frequency OAEs [12]. Figure 2C show the delay plotted in a different fashion (N_{SOAE}, [4]), where there is a clear increase with frequency. Lastly, Fig.2D shows one measure of

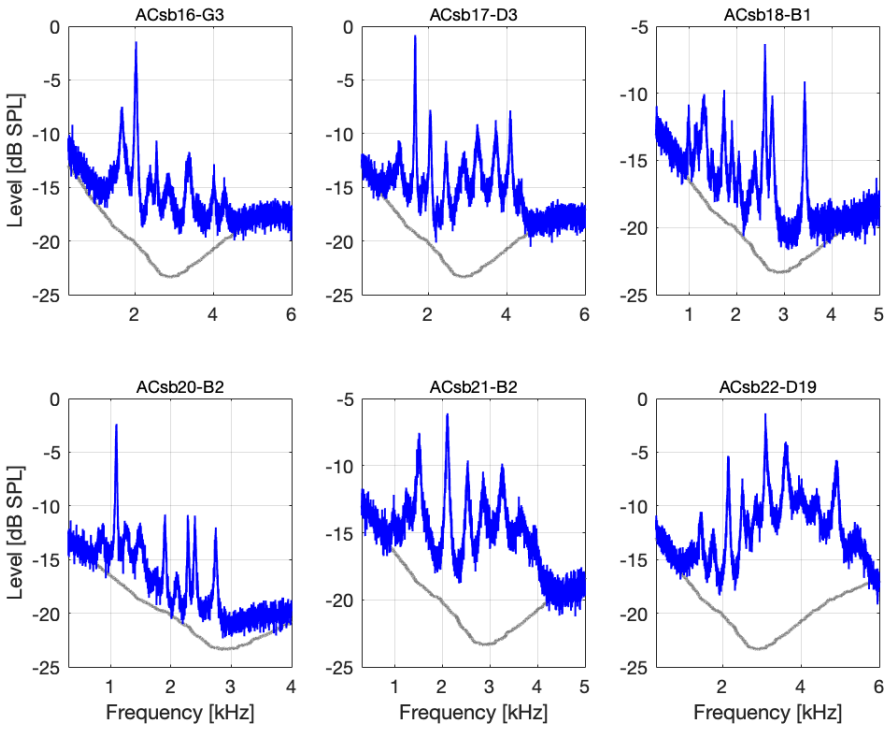


FIGURE 1. Typical SOAE magnitude spectra from six different *Anolis* lizards. An approximate noise floor is shown by the grey curve. Each trace was computed by spectral averaging of 60 sequential time segments (32768 points in length sampled at 44.1 kHz) measured from the microphone coupled to the external meatus. Note that these spectra may have been recorded at different body temperatures, however the temperature was stable in each individual case.

ascertaining whether SOAE peaks exhibit a likelihood that their frequency corresponds to the cubic intermodulation distortion of two nearby peaks. If peaks are cubic intermodulation distortions, this ratio would be expected to be close to unity. While some peaks exhibit a ratio very close to one, the lack of clustering about unity in Fig.2D suggest that is the exception rather than the norm. That is, SOAE peak frequencies tend to be highly idiosyncratic rather than distortions of other peaks.

Interpeak Fluctuation Correlations

To examine correlations in non-stationary properties, 120 s time waveforms were analyzed. Upon initial computation of the spectrum to identify peak location, the signal could then be narrowband filtered about a given peak frequency to obtain $\gamma_n(t)$ (n represents the filter center frequency). The analytic signal could then be computed via the Hilbert transform ($\rightarrow \Gamma_n(t)$), from which the envelope ($|\Gamma_n(t)|$) and instantaneous frequency ($-d\angle\Gamma_n(t)/df$) waveforms could be extracted. These allowed us to assess fluctuations in both the amplitude (AM, or amplitude modulation) and instantaneous frequency (FM). A similar analysis for another peak ($\Gamma_m(t)$) could be done, providing four sets of waveforms for correlation analysis. Similar to a convolution, the correlation analysis could be done in either the time or frequency domains.

Preliminary results for two peaks (Fig.3A) are presented. For both these peaks, they exhibited statistical properties consistent with a self-sustained sinusoid (Fig.3B), verified using a “dip statistic" [23]. Timecourses for both AM and FM are shown in Fig.3C for both peaks. Correlation in AM is shown in Fig.4A. Both temporal and spectral methods yield a weak positive correlation, with the higher frequency peak leading the other by approximately 0.3-0.4 ms. For FM, a correlation was only observed using a spectral domain approach (Fig.4B), the reason for such presently uncertain. A more substantial delay of 4 ms was observed here, again with the higher frequency peak leading. Lastly, intrapeak correlations for AM and FM are shown in Fig.4C. A positive correlation was present for the lower frequency

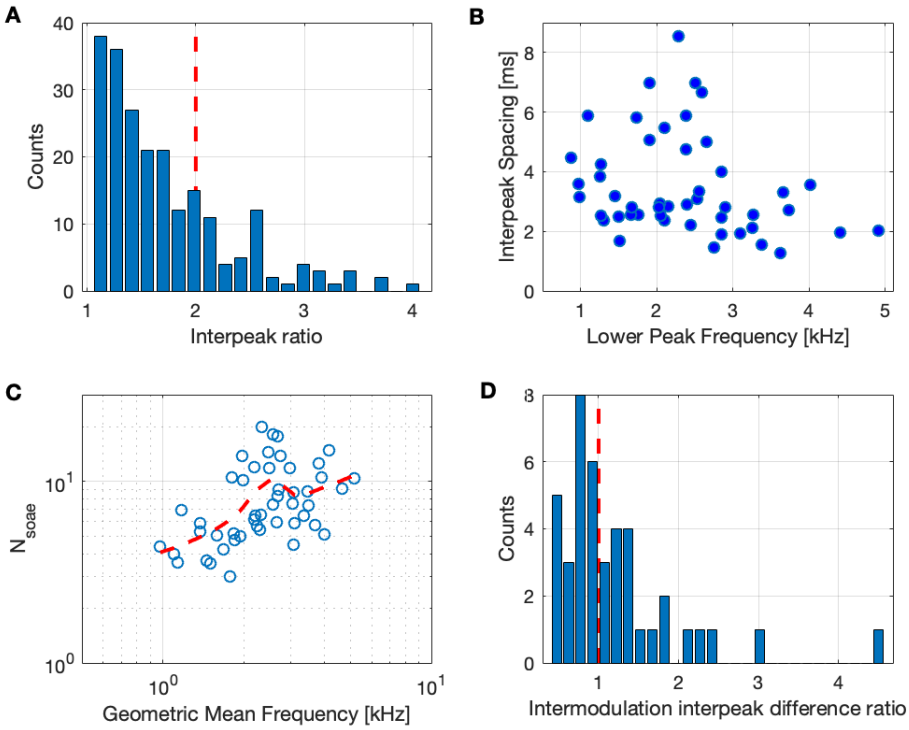


FIGURE 2. Several interpeak relationships stemming from the SOAE spectra shown in Fig.1. **A**- Distribution of interpeak frequency ratios (i.e., SOAE peak frequency divided by other peak frequencies below it). Red dashed line indicates a factor of two for visual reference. **B**- Reciprocal of interpeak frequency spacing (hence units of ms), plotted relative to lower peak frequency of the pair. **C**- Interpeak delay plotted in number of cycles relative to the geometric mean frequency of the pair (N_{SOAE}). A trend line is included for visual reference. **D**- Considering triplets of consecutive peak frequencies ($f_1 < f_2 < f_3$), this plot shows the distribution of the frequency difference ratio $(f_3 - f_2)/(f_2 - f_1)$. The vertical red line, placed for visual reference, shows the ratio value that would result if $f_1 = 2f_2 - f_3$ or $f_3 = 2f_2 - f_1$ (the total count has been divided by two to avoid redundancy).

peak (AM lagging FM by several ms). Less clear is the higher frequency peak, which exhibits a negative correlation with a shorter delay.

For other peak pairs in the same ear (see Fig.3A), the results are varied. In approximately 60% of the pairs, AM correlations are present, whereas FM correlations are exceedingly rare. Analyses for additional animals have yet to be completed. However, initial analyses (not shown) for human and barn owl waveforms [12] indicate correlations can be present, but similarly are exceptional (either within a given ear or across individuals). In short, our preliminary results appear consistent with (and extend) previous studies [1, 2, 20] demonstrating that interpeak AM and/or FM correlations can be observed, but they manifest relatively infrequently.

DISCUSSION & CONCLUSIONS

Several considerations can be drawn from the preliminary SOAE interpeak results described here. First is the characteristic interpeak spacing that is suggestive of delays on the order of several milliseconds (Fig.2). Such appear highly consistent with tuning delays estimated from SFOAEs [12], as well as auditory evoked potential latencies [24]. These observations suggest that mechanical tuning is a critical aspect in determining features observed in SOAE spectra.

Another facet relates to the relative lack of AM and/or FM correlations between SOAE peaks from a given ear. Correlations in *Anolis* SOAE peaks could reasonably be expected due to mechanical coupling, given their relatively small papilla and high hair cell density (e.g., the gap between neighboring hair cells is sub-micron) [16], despite the lack of a tectorial membrane. Further, correlations are present in SOAE from both the left and right ears [14], indicating that weak coupling via the interaural canal is sufficient to synchronize two ears [21]. Yet, our intra-ear results

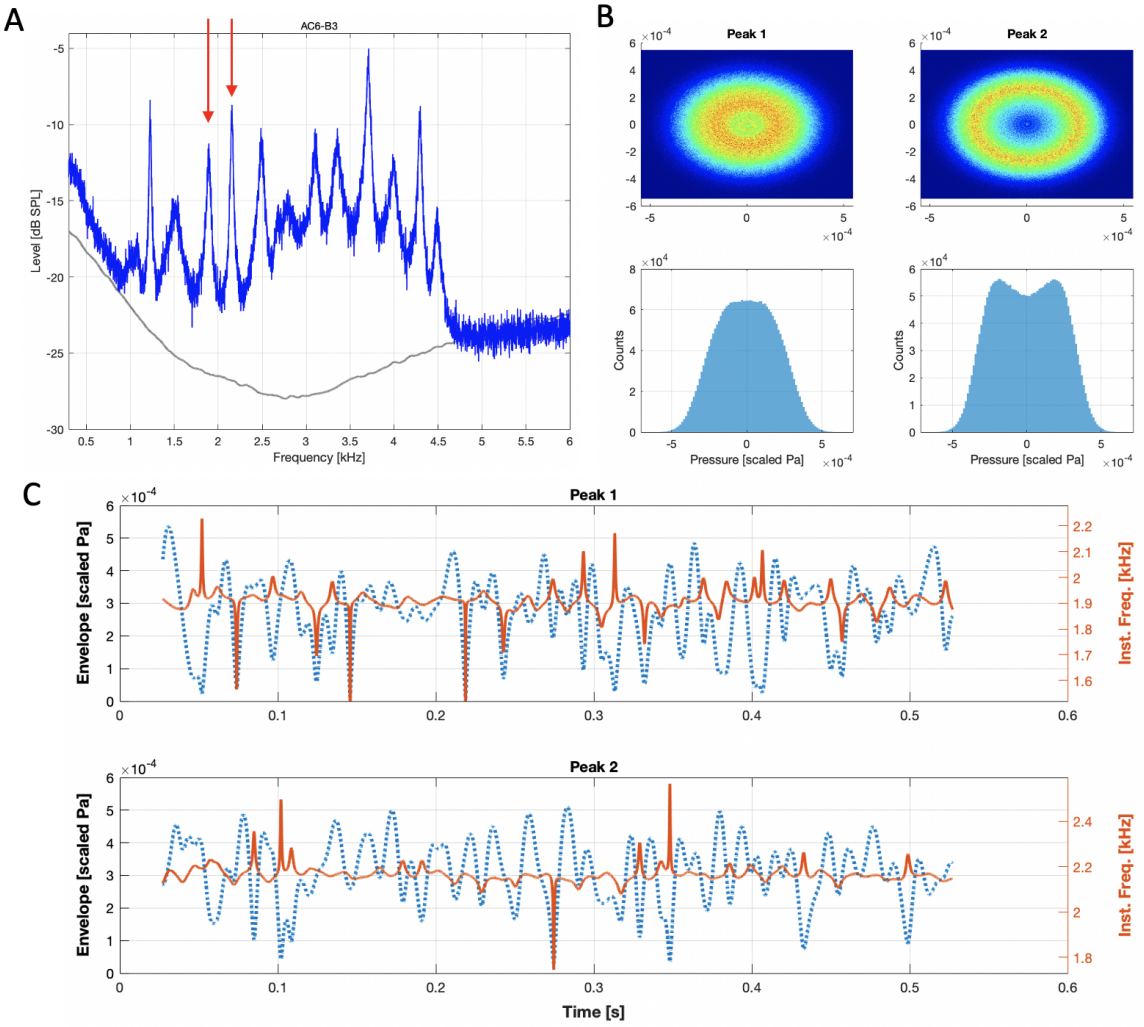


FIGURE 3. SOAE activity from a single *Anolis* lizard. **A**- Averaged spectrum (see Fig.1). Two peaks used for the analysis are shown via the red arrows (peak 1 at 1.9 kHz, peak 2 at 2.164 kHz). **B**- Statistical properties for each peak, shown either as the distribution of the analytic signal (top) or as an amplitude distribution (bottom). Both showed properties consistent with a sinusoid rather than filtered noise. **C**- As described in the text, extracted waveforms for both AM (blue dashed line) and FM (red solid) for each peak. These were the signals used for the correlation analysis shown in Fig.4.

are consistent with previous studies [2, 20] in that interpeak correlations are relatively uncommon. One possibility is that the bulk of potential correlations simply do not exist, or are too weak to be observed otoacoustically (although bilateral correlations suggest otherwise, [14]). Another possibility is that the analysis methodology presently employed is not suitable. This consideration does not however satisfactorily explain however how clear correlative behavior is sometimes present. Lastly, noise may obscure most correlations, perhaps indicative of SOAE activity operating in the vicinity of thermal thresholds for transduction. Regardless, any sort of theoretical framework for SOAE generation ideally should be able to describe this sporadic nature of interpeak correlations. That is, the observed degree of correlation should place tangible constraints upon inter-element coupling in SOAE models.

Ultimately we strive to interpret the data through the lens of elucidating the biomechanics of SOAE generation. Despite experimental advances (e.g., [8, 25]), our current understanding of SOAE is poor. Numerous key questions remain, such as precisely what is oscillating in the inner ear, how those various elements might be plausibly mechanically coupled so to affect one another, and why many species lack SOAE. Related to those considerations is uncertainty as to what the “active" process is in terms of creating force generation that gives rise to SOAE. For example, the notion of how negative damping (e.g., [5]) or a “nonlinear gain medium powered by hair cells" [26]

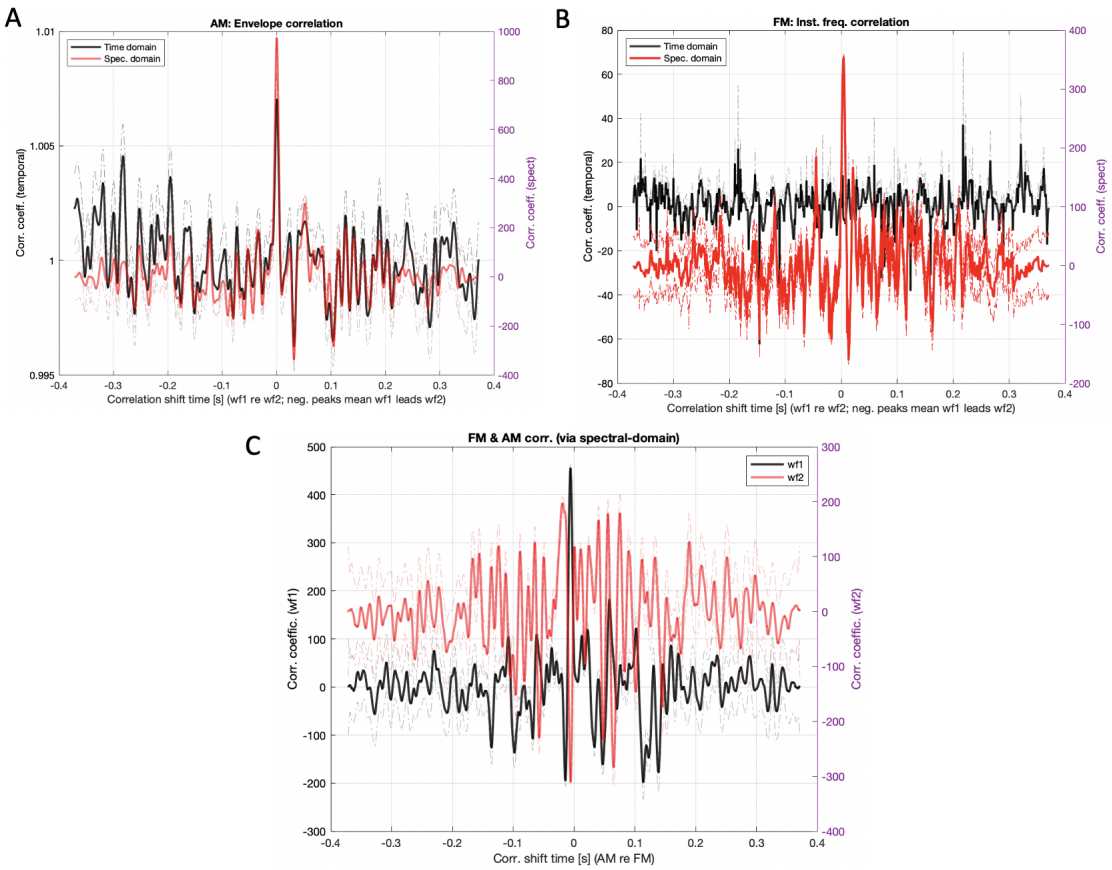


FIGURE 4. Correlation analyses for the two peaks shown in Fig.3. For comparison, traces in panels A and B show correlations computed in both the temporal (black) and spectral (red) domains. Standard error when averaging across time segments is shown via the faint dashed lines. **A**- Interpeak amplitude modulation (AM). **B**- Interpeak frequency modulation (FM). Note that the time domain method does not yield a correlation peak while the spectral domain approach does. **C**- Intrapeak AM and FM correlations (spectral domain only). Peak 1 is shown in black and peak 2 in red.

biophysically arise at audio frequencies remains obscure.

The data presented here provide further context for the discussion around the notion that there are fundamentally different classes of SOAE models. For example, one argument posits that there are "'local-'" and "'global-oscillator' frameworks", going so far to claim that "the two frameworks also quietly espouse radically opposing views of causality" [26]. We argue that such a distinction is a distraction, as common sense indicates that there can be little doubt that the inner ear is a "global" system in that there are many disparate mechanical elements that work together. The pressing question is how they are coupled and how such gives rise to the collective behavior that emerges, especially considering nonlinearity as well as middle ear coupling. By focusing efforts on the relatively simpler morphology of the lizard's inner ear, further study on interpeak relationships should help further bridge the more pressing "two rather different perspectives": in-vitro (saccular) hair cell physiology and in-vivo mammalian cochlear mechanics. Ultimately, to quote Zweig from the 2003 MoH proceedings, the key consideration here is one of *cellular cooperativity* [27].

ACKNOWLEDGMENTS

CB and RW were supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Grant No. RGPIN-430761-2013.

REFERENCES

1. P. van Dijk and H. Wit, "Amplitude and frequency fluctuations of spontaneous otoacoustic emissions," J. Acoust. Soc. Am. **88**, 1779–1793 (1990).
2. P. van Dijk, G. A. Manley, and L. Gallo, "Correlated amplitude fluctuations of spontaneous otoacoustic emissions in six lizard species," J. Acoust. Soc. Am. **104**, 1559–1564 (1998).
3. C. Talmadge and A. Tubis, "On modeling the connection between spontaneous and evoked otoacoustic emissions," in *Biophysics of Hair-Cell Sensory Systems*, edited by H. Duifhuis, J. Horst, P. van Dijk, and S. van Netten (World Scientific Press, 1993) pp. 25–32.
4. C. A. Shera, "Mammalian spontaneous otoacoustic emissions are amplitude-stabilized cochlear standing waves." J. Acoust. Soc. Am. **114**, 244–262 (2003).
5. A. Vilfan and T. Duke, "Frequency clustering in spontaneous otoacoustic emissions from a lizard's ear." Biophys. J. **95**, 4622–4630 (2008).
6. B. Epp, J. L. Verhey, and M. Mauermann, "Modeling cochlear dynamics: Interrelation between cochlea mechanics and psychoacoustics," J. Acoust. Soc. Am. **128**, 1870–1883 (2010).
7. G. A. Manley, "Cochlear mechanisms from a phylogenetic viewpoint," Proc. Nat. Acad. Sci. USA **97**, 11736–11743 (2000).
8. M. Cheatham *et al.*, "Loss of the tectorial membrane protein CEACAM16 enhances spontaneous, stimulus-frequency, and transiently evoked otoacoustic emissions," J. Neurosci. **34**, 10325–10338 (2014).
9. G. Manley, L. Gallo, and C. Köppl, "Spontaneous otoacoustic emissions in two gecko species, *gekko gecko* and *eublepharis macularius*," J. Acoust. Soc. Am. **99**, 1588–1603 (1996).
10. W. Bialek and H. Wit, "Quantum Limits to Oscillator Stability - Theory and Experiments on Acoustic Emissions From the Human Ear," Phys. Lett. A **104**, 173–178 (1984).
11. P. van Dijk, G. A. Manley, L. Gallo, A. Pavusa, and G. Taschenberger, "Statistical properties of spontaneous otoacoustic emissions in one bird and three lizard species," J. Acoust. Soc. Am. **100**, 2220–2227 (1996).
12. C. Bergevin, G. Manley, and C. Köppl, "Salient features of otoacoustic emissions are common across tetrapod groups and suggest shared properties of generation mechanisms." Proc. Nat. Acad. Sci. USA **112**, 3362–3367 (2015).
13. G. Manley, "Evidence for an active process and a cochlear amplifier in nonmammals," J. Neurophysiol. **86**, 541–549 (2001).
14. Y. Roongthumskul, D. Ó. Maoiléidigh, and A. Hudspeth, "Bilateral spontaneous otoacoustic emissions show coupling between active oscillators in the two ears," Biophysical journal **116**, 2023–2034 (2019).
15. M. Miller, "Scanning electron microscope studies of the auditory papillae of some iguanid lizards." Am. J. Anat. **162**, 55–72 (1981).
16. J. Negandhi, C. Bergevin, and R. V. Harrison, "Scanning electron microscopy of the basilar papilla of the lizard (anolis carolinensis)," Can. Acoust **46**, 7–12 (2018).
17. P. van Dijk and H. P. Wit, "Synchronization of cubic distortion spontaneous otoacoustic emissions," J. Acoust. Soc. Am. **104**, 591–594 (1998).
18. G. Manley, D. Kirk, C. Köppl, and G. Yates, "In-vivo evidence for a cochlear amplifier in the hair-cell bundle of lizards." Proc. Nat. Acad. Sci. USA **98**, 2826–2831 (2001).
19. M. Beurg, T. Gamble, A. H. Griffing, and R. Fettiplace, "Atypical tuning and amplification mechanisms in gecko auditory hair cells," Proc. Nat. Acad. Sci. USA **119**, e2122501119 (2022).
20. P. van Dijk and H. P. Wit, "Correlated amplitude fluctuations of spontaneous otoacoustic emissions," J. Acoust. Soc. Am. **104**, 336–343 (1998).
21. C. Bergevin, A. Mason, and N. Mhatre, "Evidence supporting synchrony between two active ears due to interaural coupling," J. Acoust. Soc. Am. **147**, EL25–EL31 (2020).
22. L. McKetton, D. Purcell, V. Stone, J. Grahn, and C. Bergevin, "No otoacoustic evidence for a peripheral basis of absolute pitch," Hear. Res. **370**, 201–208 (2018).
23. J. D. Salvi, D. Ó. Maoiléidigh, B. A. Fabella, M. Tobin, and A. Hudspeth, "Control of a hair bundle?s mechanosensory function by its mechanical load," Proc. Nat. Acad. Sci. USA **112**, E1000–E1009 (2015).
24. E. F. Brittan-Powell, J. Christensen-Dalsgaard, Y. Tang, C. Carr, and R. J. Dooling, "The auditory brainstem response in two lizard species," J. Acoust. Soc. Am. **128**, 787–794 (2010).
25. P. M. Quiñones, S. W. Meenderink, B. E. Applegate, and J. S. Oghalai, "Unloading outer hair cell bundles in vivo does not yield evidence of spontaneous oscillations in the mouse cochlea," Hear. Res. , 108473 (2022).
26. C. A. Shera, "Whistling while it works: Spontaneous otoacoustic emissions and the cochlear amplifier," Journal of the Association for Research in Otolaryngology , 1–9 (2022).
27. G. Zweig, "Cellular cooperation in cochlear mechanics," Biophysics of the Cochlea: From Molecules to Models. AW Gummer, editor. World Scientific, New Jersey, London, Singapore, Hong Kong , 315–329 (2003).

Elliptical Sound-Induced Motion in the Bushcricket Hearing Organ

Anna Vavakou,[1, a)] Jan Scherberich,[2, b)] Manuela Nowotny,[2, c)] and Marcel van der Heijden[1, 2, d)]

[1)]*Department of Neuroscience, Erasmus MC, University Medical Center Rotterdam, 3000 CA, Rotterdam, The Netherlands;*
[2)]*Institute for Zoology and Evolutionary Research, Friedrich Schiller University Jena, 07743 Jena, Germany*

[a)]*Corresponding author: a.vavakou@erasmusmc.nl*
[b)]*Electronic mail: jan.scherberich@uni-jena.de*
[c)]*Electronic mail: manuela.nowotny@uni-jena.de*
[d)]*Electronic mail: m.vanderheyden@erasmusmc.nl*

Abstract. The bushcricket Mecopoda elongata can hear a wide range of frequencies, utilizing less than 50 sensory units. It is known that transduction takes place in the sensory dendrite but the underlying mechanisms are still unexplained. In a recent study we used optical coherence tomography (OCT) to characterize the relative motion in the dorsal-ventral direction between anatomical structures near the transduction site. In the current study, we combined OCT vibrometry with the use of a mirror to measure vibrations of the same structures from two different angles. By developing a method to identify measurement locations across the two viewing angles, we were able to decompose the motion into a anterior-posterior and a dorsal-ventral component. In the structures surrounding the transduction site we observed elliptical sound-induced motion that varied systematically with frequency. The motion in the cap cells (CC) was more tilted towards the anterior-posterior axis than that of the dorsal wall (DW). Finally, the two-dimensional relative motion between the DW and the CC, a candidate for the drive of transduction, was sharply tuned, and for ultrasound it was similar to the neural tuning described in the literature for this species.

INTRODUCTION

Katydids are acoustically communicating insects. They have minuscule hearing organs, smaller than 1 mm, in their front tibia, the crista acoustica (CA). These organs allow them to hear a wide range of frequencies, from few hundreds of Hertz up to ultrasounds [1]. This is achieved by few sensory units (less than 50 in *Mecopoda elongata* [2, 3]) that are tonotopically organized along the proximal-distal axis [4, 5]. The sound input reaches the CA through the acoustic trachea and the two tympana that surround the sensory tissue [6, 7, 8, 9, 10]. Each sensory unit consists of a sensory dendrite, that is located dorsally to the dorsal wall (DW) and a cap cell (CC) that is found on the dorsal tip of the sensory dendrite (Fig. 1 C,D). The sensory dendrite is surrounded by a scolopale cell.

The frequency-selective transduction of mechanical energy to electrical responses is thought to take place through stretching of the sensory dendrite [11, 12]. The neural tuning of the sensors is sharper and shifted towards lower frequencies compared to the mechanical tuning [12]. The actual mechanism of sensory transduction remains unknown. A step towards unraveling the mechanism was taken by Vavakou et al. [13], who used OCT-vibrometry to measure sound-induced vibrations of both the DW and the CC. A crucial aspect of that study was the use, for the first time in insect hearing, of OCT vibrometry, a technique that allows for depth resolution of the recorded tissue. They reported that the CC and the DW mechanical responses to sound stimuli differ in response phase, best frequency, and bandwidth. Further micromechanical exploration of the recorded structures revealed that the relative motion among the DW and the CC is tuned more sharply and to higher frequencies than the individual structures [13].

However, with the customary vibrometry approach, it is only possible to measure the projection of the overall motion of the recorded structure on the recording beam [14, 15]. Vavakou et. al [13] aligned their measurement beam with the dorsal ventral axis of the recorded structures. As a result, any other possible directions of motion are underrepresented in their measurements. As discussed by the authors, transduction could be driven by more complex, two-dimensional motion of the CC relative to the DW, and they concluded that an accurate description of the overall motion requires the measurement of the same structure from different recording angles [13]. Facing a similar challenge, Lee et. al in 2016 [14] disentangled two motion directions in structures within the mouse cochlea, by rotating the preparation relative to the recording beam.

In the current report, we measured sound-induced vibrations of the CA from two different angles, using a mirror (Fig. 1 A). The direct view (DV) was aligned with the dorsal-ventral axis, and the CA was imaged through the dorsal cuticle. Using the mirror, the CA were imaged through the anterior tympanum (Fig. 1 B-D). An important step in the processing of the data is an accurate method of "registration", i.e., finding the corresponding recorded positions across

the viewing angles by embedding them in a three-dimensional coordinate system. Using geometrical calculations, the same target structures were detected in both views. We applied this registration method to characterize the sound-induced vibrations of both the DW and the CC (Fig. 1 E,F). This allowed us to decompose the recorded motion to the dorsal-ventral and anterior posterior components. We report that elliptical motion is prevalent in both the DW and the CC. We further quantify the elliptical character of the recorded motion using geometrical metrics.

METHODS

Animal Preparation In total, 8 bushcrickets of the species *Mecopoda elongata* were used for in vivo sound-induced vibration measurements on the CA. The cuticle was usually surgically removed and replaced by a glass cover slip, to improve the optics and ensure the CA were visible in the OCT image. For more details, look into [13]. In experiment BC0013 the CA were recorded through an intact cuticle. The position of the tibia was optimized so that the dorsal-ventral axis of the CA was aligned with the recording beam. The CA were also measured through the anterior tympanum, using a mirror that was placed close to the anterior tympanum. For that purpose, a 3-mm Protected Gold Coated, N-BK7 Right Angle Mirror 65-844 by Edmund Optics (EO) was used. The mirror was attached to a rigid cable, allowing for finer control during the placement. The rigid cable attached to the mirror was stabilized using plasticine, at the position where the CA image quality through the anterior tympanum was optimal. At the end of the experiment, the animal was anesthetized, and subsequently euthanized by decapitation.

Acoustic stimulation Broadband multi-tone "zwuis" complexes were used, following an open field stimulation paradigm [16]. We used a ScanSpeaker R2904/700005 loudspeaker, placed 35 cm away from the animal, inside a sound-proof booth. The ipsilateral acoustic trachea opening on the thorax was facing the loudspeaker. The surfaces of the OCT probe, animal holder, etc., were covered with cotton wool in order to minimize acoustic reflections. The most commonly used broadband stimulus consisted of 47 spectral components, spanning 0.3 kHz to 38 kHz (average frequency spacing: 817 Hz), presented at 79 dB SPL per component. The components all had equal amplitudes, with levels expressed in dB re 20-μPa (i.e., dB SPL), but stimulus phase was randomized across frequency. Each stimulus was presented for 12 s, and inter-stimulus intervals lasted for approximately 60 s.

OCT imaging and vibrometry Vibration measurements were obtained using OCT. More details on the experimental setup are in ref. [15], [13]. A spectral domain - optical coherence tomography system (Thorlabs Telesto TEL320C1) was used for interferometric imaging and vibration measurements. Its field of view was 10 × 10 mm; its depth range, 3.5 mm. The working distance of the objective was 24 mm. For our dorsal view measurements, the CA was imaged directly, without using a mirror. The OCT measurement beam was manually aligned with the dorsal–ventral anatomical axis of the CA. Therefore, these vibration measurements favor the vibrations along this axis. The depth resolution of the OCT allowed us to record the DW and the CC simultaneously, along one recording beam, at different positions of the CA. Therefore, the DW and CC measurements obtained with a single optical beam share the same distance from the distal end of the CA. Additionally, the CA was imaged through the anterior tympanum, using a mirror (Fig. 1 A,B). Viewed through the anterior tympanum, the DW and the CC can not be obtained from the same recording beam. For that reason, vibration measurements on these structures were obtained separately (see Fig. 1 C). The sample was not moved when recorded through the mirror and directly. Still, the OCT probe had to be adjusted vertically to ensure that the recorded structure was in the focal plane.

Response analysis Mechanical measurements were obtained from (1) the DW imaged along the dorsal-ventral axis, (2) the CC measured along the dorsal-ventral axis, (3) the DW measured along the anterior-posterior axis, (4) the CC measured along the anterior-posterior axis. Responses were analyzed by Fourier transformation of the vibration waveforms derived from contiguous groups of three pixels in each M-scan, each pixel covering a depth of 3 μm. For more details, consult [13]. The mechanical measurements from the CA were normalized to vibrations recorded on the distal end of the organ obtained from the dorsal-ventral view (as in [13]). All the measured structures shared the same mechanical reference. This use of a common reference is analogous to normalization to the vibrations of middle ear ossicles commonly used in the field of cochlear mechanics [17]. It corrects imperfections in the calibration of the acoustics. In some cases (3 experiments), the readjustment of the probe height among the two views altered the microacoustical environment, and the measurement had to be discarded. Usually, this effect could be avoided by using sound absorbing material. An essential methodological aspect of this study is to ensure a proper match ("registration") of measurement points between the direct and the mirror views. That was achieved post-hoc, by first measuring vibrations in a dense array of locations along the length of the CA, using both views, and then registering all the measurement points and the mirror surface plane in a global 3D coordinate system. Successful registration ensured that the same structure was measured from two different angles, obtained from the direct view and the mirror

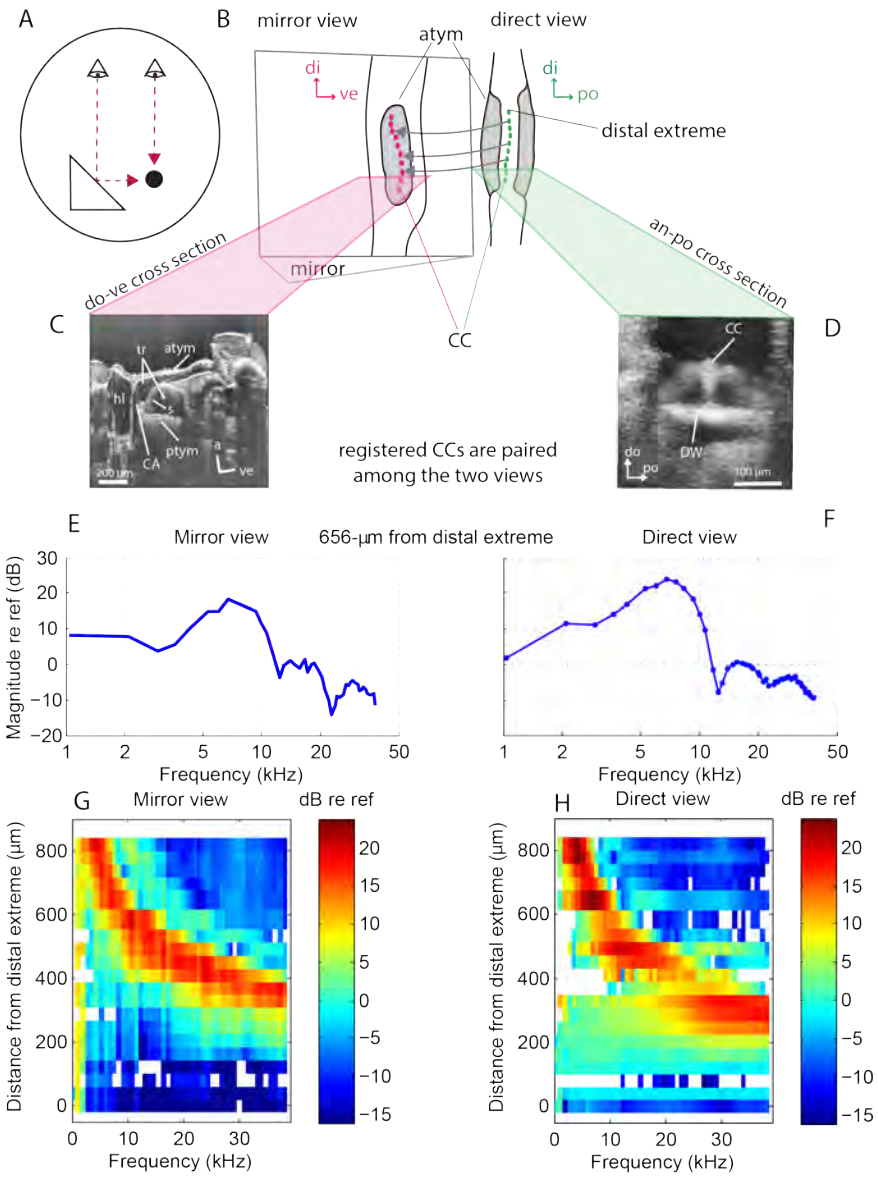


FIGURE 1. Measuring the CA from two different angles. (**A**) Combining the depth resolution of OCT vibrometry with the use of a mirror allowed us to measure vibrations on the same structures from two distinct viewing angles. (**B**) Illustration of the measured sample. The hearing organ was recorded from two different views, the probe facing the dorsal cuticle directly (DV, direct view), and it facing the anterior tympanum indirectly via a mirror (MV, mirror view)(di: distal, ve: ventral, po: posterior, atym: anterior tympanum.) The coordinates of the measured CC and DW positions (magenta and green markers respectively) were registered in 3D space. Corresponding measured positions were paired among the two views. (**C**) Cross-section OCT image ("B-scan") obtained with the mirror view, with the measurement beam (vertical direction in B-scan) directed along the anterior-posterior axis of the hearing organ, and the scanning direction (horizontal direction of B-scan) along the dorsal-ventral direction as shown in panel B. This view favors the measurement of vibrations in the anterior-posterior direction. (do: dorsal, atym: anterior tympanum, ptym: posterior tympanum, tr: trachea, s: septum, CA: crista acustica, hl: haemolymph) (**D**) B-scan obtained with the direct view, with the beam along dorsal-ventral axis and the scanning direction along the anterior-posterior image section plane as shown in panel B. This view favors vibrations in the anterior-posterior direction. (CC: cap cell, DW: dorsal wall). (**E,F**) Tuning of the normalized vibration magnitude of the cap cells located 656 μm from the distal extreme of the CA as imaged mirror view (E) and direct view (F). (**G,H**) Longitudinal vibration maps of paired recorded CC positions between the mirror and the direct views. (**G**) Normalized magnitude responses of the CC across different frequencies and recorded positions along the CA, as recorded from the mirror view. (**H**) Normalized magnitude responses of the CC across different frequencies and recorded positions along the CA, as recorded from the direct view. (Data from experiment BC0013)

view.

RESULTS

CA motion disentangled along the two anatomical axes

In previous measurements along the do-ve axis, [13] reported that the spectral tuning of the DW is shifted at higher frequencies than the CC when recorded at the same longitudinal position. The reported difference in tuning cannot be generalized for the overall motion of the recorded structures, as one recording angle is not sufficient to describe 3D motion. Stated differently, it is possible that the motion direction changes with frequency.

With the use of a mirror, we were able to measure sound-induced motion on the CA from two different angles. That way, we disentangled motion along the do-ve and the an-po axis for both the DW and the CC.

We measured sound-induced vibrations on the CA, covering at least 800 μm of the sensory tissue CA with a spatial resolution of 30-40 μm, while presenting multi-tone broad band stimuli ranging from 0.3-38 kHz, in $\approx$ 817-Hz steps. These measurements allowed us to construct heat maps of the CA vibration magnitude along the do-ve and an-po axis. In summary, the motion direction in the CC varies in a frequency dependent manner, whereas the DW motion was exclusively dominated by a do-ve component.

Elliptical motion in the CA

The CC magnitude and phase at a single location was measured from two different angles, and motion components along the an-po and the do-ve axis were disentangled (Fig. 2A-B) as described in Methods. From these data the cyclic trajectories of motion with respect to the two anatomical axes were reconstructed for each of the frequency components of the stimulus (Fig. 2C).

The 2D motion increased in relative amplitude for frequencies near the best frequency of this recorded position, 11.5 kHz. The elliptical character of the motion was most pronounced near the best frequency. The an-po component took over at frequencies above 13 kHz, where the ellipse flattened. Finally, the 2D motion decreased in amplitude and became more tilted for frequencies higher than 20 kHz.

The normalized magnitude and phase along the an-po and the do-ve axis were also measured for the DW (Fig. 2E-F). Compared to the CC, the elliptical motion in the DW was consistently more uniform and elongated across frequencies (Fig. 2D). For higher frequencies, the direction of the 2D motion along the an-po axis reversed. The directions change occurred at the frequency region where the high frequency plateau has been reported in previous studies [4]. Overall, a frequency-dependent elliptical motion in the do-ve/an-po plane was observed for both the DW and the CC.

Relative motion

We further assessed the relative motion of CC and DW along the do-ve (Fig. 3 A,B,C) and the an-po axis (Fig. 3 B,E,H), by computing the vector difference of their mechanical motions (i.e., taking into account both their relative amplitude and phase). In all cases, we found that the relative motion was shifted towards higher frequencies than the do-ve spectral tuning of the CC. For the do-ve axis this was expected from the measurements of [13]. For motion along the an-po axis, this is reported for the first time. The relative motion of two elliptical motions in the same plane is also an elliptical motion. For that reason, it is possible to calculate the major axis a of the ellipse for vector difference, and to use it as a metric for the amplitude of the overall relative motion (Fig. 3 C,F,I). For all the reported experiments, the a of the relative elliptical motion was sharply tuned.

CONCLUSION

In summary, we recorded sound-induced vibrations in the bushcricket hearing organ from two different angles, using a mirror. First, the recorded structures were registered geometrically, ensuring that the same points were measured from

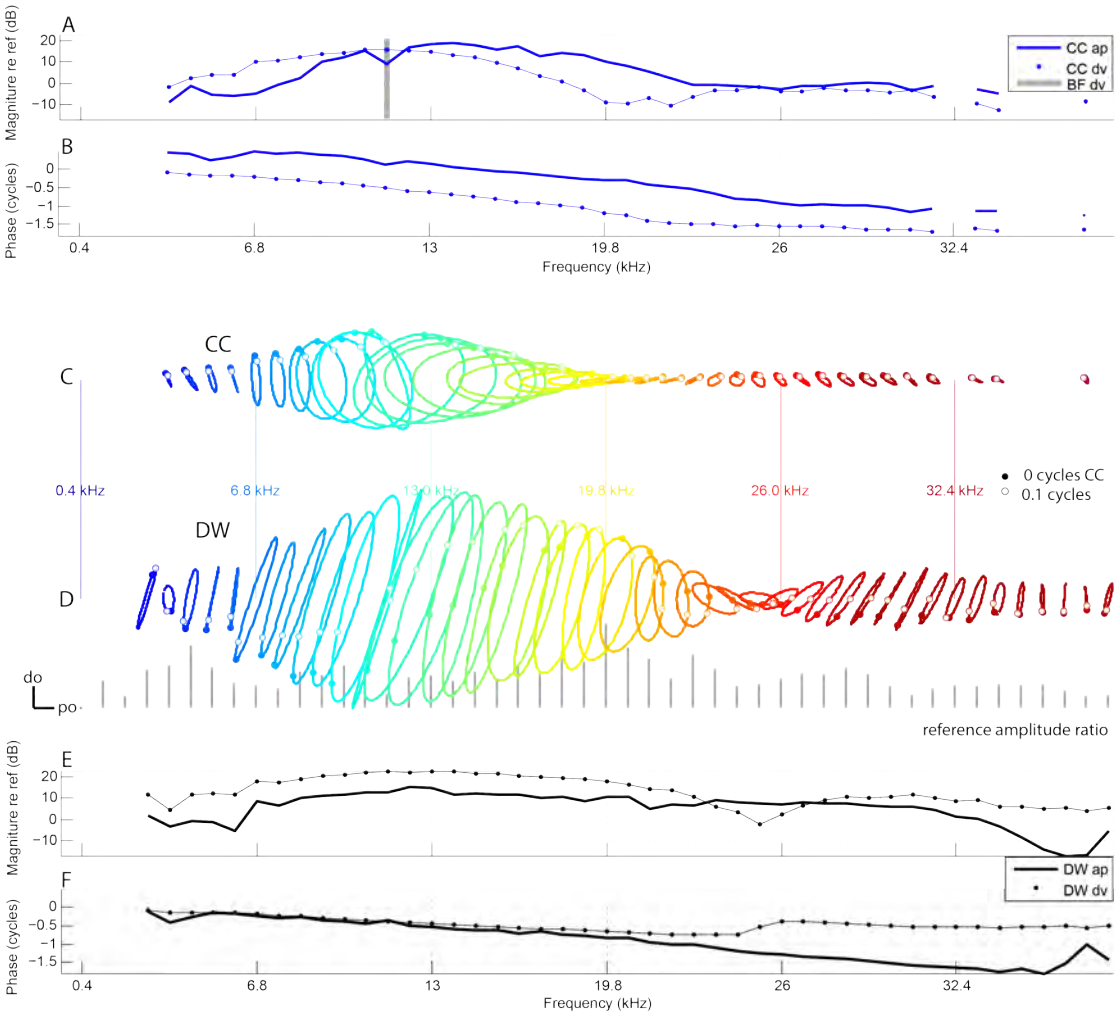


FIGURE 2. Elliptical motion in the CA.(**A,B,E,F**) Normalised displacement magnitude and phase as measured through the mirror and the direct view, for the DW and the CCs. The mirror view and the direct view are related with a known angle delta, that is derived from the OCT images. The direct view is aligned with the dorsal-ventral axis by convention. The dorsal-ventral motion component is identical to the direct view measurement. The anterior-posterior motion component is derived by the obtained measurements through the mirror and the direct view. (**C,D**) Reconstructed elliptical motion for one stimulus frequency cycle. (Data from experiment BC0013. CC and DW were recorded 533 μm away from the distal extreme of the direct view. Best frequency for do-ve CC: 11.5-kHz)

both views (Fig. 1). Secondly, the overall measured motion was decomposed in two vertical components, aligned with the do-ve and the an-po anatomical axis respectively, and we reconstructed the 2D motion of the DW and the CC for a given position (Fig. 2), along the two anatomical axis. Finally, the elliptical relative motion among the DW and the CC was calculated (Fig. 3), and was found to be sharply tuned.

Recent vibration measurements of the DW and the CC along the do-ve axis reveal that the the CC tuning is sharper and shifted towards higher frequencies than the DW [13]. Vavakou et al. discuss in detail whether this difference in tuning among the two structures is specific for the do-ve view, or it is consistent with the overall motion of the two structures (Fig. 6 in [13]). Specifically, they considered the following question: is it possible that the DW motion is dominated by the do-ve component across different frequencies, whereas the direction of CC motion changes in a frequency dependent manner?

Hummel et al., in 2016, performed simultaneous neural and mechanical measurements of the bushcricket hearing organ [12]. Their mechanical measurements were performed on the CC along the do-ve direction. Their results indicate that the neural characteristic and best frequency is shifted towards higher frequencies compared to the mechanical

tuning, especially for frequencies above 20 kHz ([12], Fig. 1E,F). This discrepancy between neural and mechanical tuning could possibly be explained by considering the mechanical coupling that underlies transduction.

Our measurements of the CA along two anatomical axes show that the relative motion of the DW and the CC is elliptical and sharply tuned at frequencies higher than that of the individual structures. It is likely that the relative motion of the DW and the CC is driving transduction, by stretching of the sensory dendrite, especially for ultrasounds.

This study sets the methodological basis for disentangling different directions of sound-induced motion in a hearing organ by using a mirror. Although this is a powerful tool to unravel the geometry of sound-induced motion and its contribution to transduction, it is important to consider some of the limitations of this approach.

First, the mirror allows us to perform these measurements without rotating the animals among measurements. That is critical for open-field stimulation paradigms, as it avoids changes in the microacoustical environment surrounding the sensory organ between measurements. (Note that, when studying the bushcricket, it is not advisable to use a closed stimulation paradigm, as the sound reaches the CA through both the spiracle and the tympanum [6], [7], [8].) Still, the probe height needs to be adjusted for around 300 μm among measurements, to ensure that the recorded side is within the focal plane of the OCT. This can have an effect on the microacoustical environment. Especially if the animal holder position is placed vertical to the probe, there can be pronounced reverberations. The experimenter needs to make sure that the position of the animal holder does not facilitate reverberations, and that sufficient sound-insulating material is used to minimize acoustic reflections. The need for adjusting the probe height could be eliminated by using additional mirrors to equalize the optical path lengths corresponding to the two viewing angles. Finally, motion along

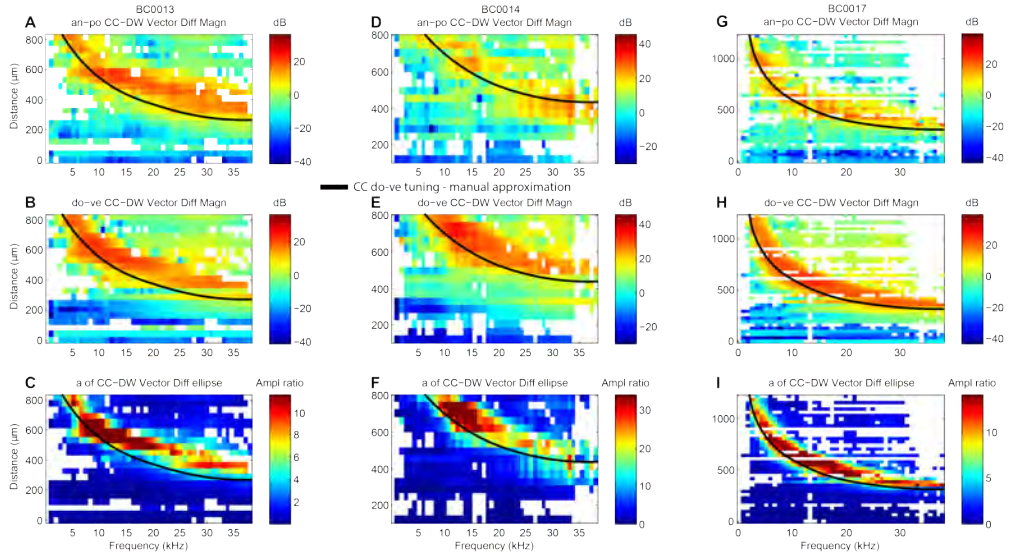


FIGURE 3. The relative motion of structures within the CA. (*A,D,G*) The relative motion of the DW and the CC along the do-ve axis for different experiments. (*B,E,H*) The relative motion of the DW and the CC along the an-po axis. (*C,F,I*) The *a* of the elliptical relative motion of among the DW and the CC along two dimensions. (Data from experiments BC0013, BC0014 and BC0017). A manual interpretation of the CC spectral tuning along the do-ve axis is replicated among all the plots per experiment for comparison.

the pr-di axis is not disentangled following the current methodology. Still, it is quite likely that it has contributed to our measurements. The way to minimize this contribution is to ensure that the measurement beam is perpendicular to the pr-di axis of the CA.

Hummel et al. in 2016 [12] report phase delays in the mechanical response of the organ that coincided with the sensory cell responses. They hypothesize that there might be a tilt of the CC along the longitudinal direction that drives transduction.

To achieve motion decomposition along 3 anatomical axes, one needs to place an additional mirror (or a prism) above the leg. This opens the way to systematically quantify any longitudinal components of motion. One methodological complication in this configuration is that the measured positions might fall outside the focal plane of the OCT, because the pathlength of the recording beam increases as the angle between the measuring beam and the pr-di axis decreases.

This study has shown that it is possible to disentangle two directions of motion using a mirror and also that there is

elliptical motion in the bushcricket hearing organ. The sample size of this study is quite small, and to put our findings on a firmer footing, there need to be more measurements that reproduce the current findings in a reliable manner. For that purpose, for future experiments we recommend to perform spatially dense measurements (30 B-scans in 10-μm steps), thus making sure that the same structure is measured from two different angles. If possible, these blocks of measurements should be repeated across different locations on the CA along the pr-di axis. With this method one can test if the discrepancy between neural and mechanical data along the pr-di axis that Hummel et al. report ([12], Fig. 2C,D) can be explained by different directions of motion of the CC.

For the current dataset, the recording beam was aimed at the middle of the CA along the an-po axis. That is not an ideal measurement point to disentangle the an-po motion if the CC pivots around that point. In future measurements, one should perform additional measurements of the CC and the DW towards the anterior and posterior extremes of the CA. The contribution of an-po motion to the recorded motion will thus more pronounced, whereas the overall results should be consistent with our current findings.

Our findings underscore that it is crucial to consider the overall motion when reporting the mechanical tuning of measured structures, including its spatial direction. That is a general consideration when studying the mechanics of hearing organs. In the mammalian cochlea there already are indications [15] and measurements [14] of longitudinal and radial motion, respectively. Given the elaborate anatomy of the mammalian cochlea, it is expected that there exist complex patterns of motion along multiple anatomical axes of the sensory organ, and that these patterns change in nontrivial ways when varying the frequency and intensity of the stimulus. It is crucial to describe in detail the micromechanical motion of hearing organs in order to generate realistic models that are consistent with mechanical data.

REFERENCES

1. J. Hummel, K. Wolf, M. Kössl, and M. Nowotny, "Processing of simple and complex acoustic signals in a tonotopically organized ear," Proceedings of the Royal Society B: Biological Sciences **281**, 20141872 (2014).
2. J. Strauß, G. U. Lehmann, A. W. Lehmann, and R. Lakes-Harlan, "Spatial organization of tettigoniid auditory receptors: insights from neuronal tracing," Journal of Morphology **273**, 1280–1290 (2012).
3. J. Hummel, M. Kössl, and M. Nowotny, "Morphological basis for a tonotopic design of an insect ear," Journal of Comparative Neurology **525**, 2443–2455 (2017).
4. A. Palghat Udayashankar, M. Kössl, and M. Nowotny, "Tonotopically arranged traveling waves in the miniature hearing organ of bushcrickets," PLoS One **7**, e31008 (2012).
5. B. Oldfield, "Tonotopic organisation of auditory receptors in tettigoniidae (orthoptera: Ensifera)," Journal of comparative physiology **147**, 461–469 (1982).
6. D. Lewis, "The physiology of the tettigoniid ear: Ii. the response characteristics of the ear to differential inputs: Lesion and blocking experiments," Journal of Experimental Biology **60**, 839–851 (1974).
7. E. Hoffmann and M. Jatho, "The acoustic trachea of tettigoniids as an exponential horn: theoretical calculations and bioacoustical measurements," The Journal of the Acoustical Society of America **98**, 1845–1851 (1995).
8. A. Michelsen, K. Rohrseitz, K.-G. Heller, and A. Stumpner, "A new biophysical method to determine the gain of the acoustic trachea in bushcrickets," Journal of Comparative Physiology A **175**, 145–151 (1994).
9. R. Heinrich, M. Jatho, and K. Kalmring, "Acoustic transmission characteristics of the tympanal tracheae of bushcrickets (tettigoniidae). ii: comparative studies of the tracheae of seven species," The Journal of the Acoustical Society of America **93**, 3481–3489 (1993).
10. J. Hummel, M. Kössl, and M. Nowotny, "Sound-induced tympanal membrane motion in bushcrickets and its relationship to sensory output," Journal of Experimental Biology **214**, 3596–3604 (2011).
11. J. E. Yack, "The structure and function of auditory chordotonal organs in insects," Microscopy research and technique **63**, 315–337 (2004).
12. J. Hummel, S. Schöneich, M. Kössl, J. Scherberich, B. Hedwig, S. Prinz, and M. Nowotny, "Gating of acoustic transducer channels is shaped by biomechanical filter processes," Journal of Neuroscience **36**, 2377–2382 (2016).
13. A. Vavakou, J. Scherberich, M. Nowotny, and M. van der Heijden, "Tuned vibration modes in a miniature hearing organ: Insights from the bushcricket," Proceedings of the National Academy of Sciences **118** (2021).
14. H. Y. Lee, P. D. Raphael, A. Xia, J. Kim, N. Grillet, B. E. Applegate, A. K. Bowden, and J. S. Oghalai, "Two-dimensional cochlear micromechanics measured in vivo demonstrate radial tuning within the mouse organ of corti," Journal of Neuroscience **36**, 8160–8173 (2016).
15. N. P. Cooper, A. Vavakou, and M. van der Heijden, "Vibration hotspots reveal longitudinal funneling of sound-evoked motion in the mammalian cochlea," Nature communications **9**, 1–12 (2018).
16. M. van der Heijden and P. X. Joris, "Cochlear phase and amplitude retrieved from the auditory nerve at arbitrary frequencies," Journal of Neuroscience **23**, 9194–9198 (2003).
17. W. S. Rhode, "Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique," The Journal of the Acoustical Society of America **49**, 1218–1231 (1971).

Nonlinearity in Hearing: The Role of Inner-Hair-Cell Saturation in Neural Coding

Laurel H. Carney[1, a)] Braden N. Maxwell[1, 2, b)]
and Virginia M. Richards[3, c)]

[1]*Departments of Biomedical Engineering and Neuroscience, University of Rochester, Rochester, NY 14642*
[2]*Department of Music Theory, Eastman School of Music, University of Rochester, Rochester, NY 14604*
[3]*Department of Cognitive Science, University of California-Irvine, Irvine, CA 92697*

[a)]Corresponding author: Laurel.Carney@Rochester.edu
[b)]bmaxwel2@u.rochester.edu
[c)]v.m.richards@uci.edu

Abstract. Psychophysical models often begin with a filterbank and assume that energy in the filter outputs is the basis for coding complex sounds. Filter bandwidths are estimated based on masked detection thresholds and are referred to as critical bands. While this conceptual model structure can explain many psychophysical phenomena, it also fails for many tasks. We propose a different conceptual model: a bank of peripheral filters, each followed by a saturating nonlinearity that represents transduction by inner hair cells (IHCs). The metric of interest is the amplitude of low-frequency fluctuations in hair-cell responses. Low-frequency fluctuations in the IHC voltage drive fluctuations in auditory-nerve (AN) responses, referred to here as "neural fluctuations" (NF). NFs are a temporal response feature distinct from phase-locking to temporal fine structure.

The saturating nonlinearity that follows each peripheral filter results in neural fluctuations that differ markedly from both the envelope of the stimulus and the envelope of a bandpass-filtered stimulus representing the basilar membrane response. A key difference between neural fluctuations and stimulus-related envelopes occurs when IHCs tuned near a spectral peak are "captured" (or dominated) by the spectral peak (Deng and Geisler, 1987; Zilany and Bruce, 2007). A peripheral channel that is captured has relatively small fluctuations in its response, whereas channels that are not captured can have strongly fluctuating responses, due to beating between multiple stimulus components. The profile of neural fluctuation amplitudes along the characteristic-frequency (CF) axis encodes key features of the stimulus spectrum. For example, in the case of masking paradigms, a local decrease in the neural-fluctuation amplitude profile encodes the presence and frequency of the target tone.

The role of the saturating IHC transduction nonlinearity in the proposed model was explored using a computational model for the periphery that includes the interaction between the saturating nonlinearity and compressive cochlear gain. We demonstrate that NF profiles in response to psychophysical masking paradigms can predict performance in several tasks for which the classical critical-band model fails (as well as those for which it succeeds.) Our goal is to determine whether a single model can predict results for several masked-detection tasks, including band-widening paradigms, with or without roving-levels, and co-modulation masking release.

The profile of NFs along the tonotopic axis is of interest because most midbrain neurons are excited or suppressed over a range of amplitude-modulation frequencies (Kim *et al*., 2020). Fluctuation profiles that are set up in the periphery are thus transformed into rate profiles in the auditory midbrain. Rate profiles across populations of model midbrain neurons were the basis for the psychophysical predictions described below. Finally, the degree of contrast in the NF profiles is reduced by sensorineural hearing loss, suggesting that fluctuation profiles may provide insight into deficits associated with hearing loss.

INTRODUCTION

Masked-detection paradigms are often modeled using specialized approaches for each paradigm. The goal of this study was to test the ability of a model based on neural fluctuations (NFs) (Carney, 2018) to predict thresholds in several classical masked-detection experiments. The experiments include some that are traditionally explained using a power-spectrum model (PSM), such as the band-widening experiment (e.g., Fletcher, 1940) and experiments that cannot be explained by the PSM, such as detection of tone in a roving-level paradigm (Kidd *et al*., 1989) and co-modulation masking release (Hall *et al*., 1984; Haggard *et al*., 1990). A previous study (Maxwell *et al*., 2020) showed that NFs can explain thresholds for detection of a tone gated on and off with a notched-noise, with and without a roving-level paradigm, as well as for a profile-analysis task (Green, 1987; Lentz *et al*., 1999).

The NF model shares some properties with modulation-filterbank models, which have been used to explain results in some of these paradigms based on the envelope statistics of the stimuli (e.g., Dau *et al*., 1997; Jepson *et al*., 2008). Unlike modulation-filterbank models, though, the NF model explicitly includes the saturating nonlinearity that describes inner-hair-cell (IHC) transduction (Fig. 1). IHC saturation has a strong effect on the temporal fluctuations in auditory-nerve (AN) responses. Furthermore, the NF model focuses on fluctuation frequencies in the range where the midbrain has the greatest sensitivity, e.g., 50-100 Hz (Kim *et al*., 2020), whereas modulation-filterbank models typically include lower-frequency amplitude-modulations. Here we show that the NF model provides a common mechanism that has the potential to explain masked thresholds across several paradigms.

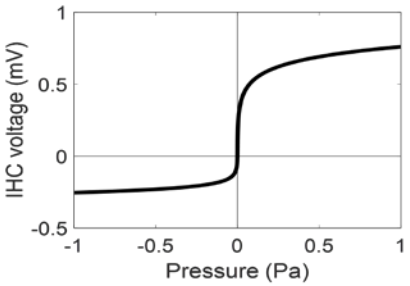


Figure 1. Input-Output nonlinearity that describes IHC transduction (Zhang et al., 2001).

The NF model was implemented using a computational model for AN fiber responses (Zilany *et al*., 2014), which includes a saturating IHC, along with basilar-membrane compression, suppression, and neural adaptation. The fluctuations of AN responses vary with CF for the stimuli in these masking paradigms. As the temporal information in the AN fluctuation profile ascends the auditory pathway, it is converted into a profile of average rate vs. CF across the population midbrain neurons, which are sensitive to fluctuations in amplitude and frequency of their inputs. A simple model for subcortical processing up to the level of the midbrain was used to simulate responses of neurons in the inferior colliculus (IC) (Carney *et al*., 2015; Carney and McDonough, 2019). One population of neurons in the IC is *excited* by fluctuations over a band of modulation frequencies, the so-called band-enhanced neurons (BE) (Kim *et al*., 2020). These neurons are often assumed to be the neural equivalent of modulation filters, and have been previously studied in psychophysical modeling studies (e.g., Maxwell *et al*., 2020). Here, the NF model was further explored using rate profiles for model band-suppressed (BS) IC neurons, which have rates that are *suppressed* by stimuli over a band of modulation frequencies to a rate below that in response to an unmodulated sound (Kim *et al*., 2020).

In response to a noise masker, AN responses have large temporal fluctuations, driven by the narrowband signals that result from cochlear filtering (Fig. 2A, top). These AN fluctuations persist even when the *average* discharge rates of AN fibers are saturated due to synaptic saturation at high sound levels. IHCs tuned near the target frequency gradually saturate over the range of moderate to high target sound levels used in most masking paradigms. As an IHC saturates, the target frequency increasingly dominates, or "captures," its response. Subsequently, an AN fiber that is driven by a captured IHC will have a response that is dominated by the target tone. AN temporal responses that are dominated by one stimulus component, such as the target tone in a masking task, are uniform over time, with fluctuation amplitudes that are reduced, e.g., in comparison to the response to the masker alone (Fig. 2A).

Fluctuations in AN responses are conveyed up the auditory pathway to the midbrain, the first stage of the pathway where neurons are tuned to fluctuation frequency (Joris *et al*., 2004). Here, we will focus on IC BS neurons (Kim *et al*., 2020). At low sound levels, for which neither the IHC voltage nor the IHC-AN synapse would be saturated, AN discharge rates encode the spectrum of the masker, and a change in rate for AN fibers tuned near the target frequency indicates the presence of the target tone in a masking task. In the case of such low-level stimuli, the IC BS cell rate profile is simply driven by the rate profile of its inputs, similarly following the profile of the stimulus spectrum. At moderate to high sound levels, such as those typically used in masking tasks, IC BS cells respond more strongly to inputs with reduced fluctuations than to inputs with large fluctuations, which suppress their responses.

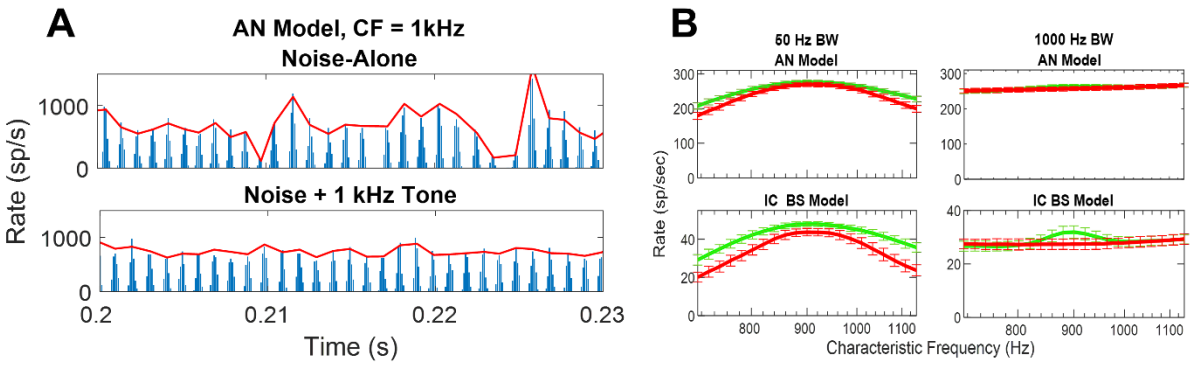


Figure 2. A) PSTH for noise alone (top) and noise+1kHz tone (bottom) for AN model with CF=1000 Hz. 70-dB SPL noises had 1kHz bandwidth, centered at 1 kHz with a tone at 70 dB SPL. Binsize = 0.1 ms. B) Model AN and IC BS cell population average rate profiles for noise-alone (red) and noise + 900 Hz tone (green). Masker bandwidth = 50 Hz (left) and 1000 Hz (right). Error bars are standard deviations across 100 trials. Noise spectrum level = 35 dB re 20 μPa. Tones: 55 dB SPL.

Thus, IC BS neurons tuned near the target frequency have relatively strong responses when the tone is present, because of the reduced fluctuations on their inputs (Fan *et al.*, 2021). In contrast, BS neurons tuned at frequencies away from the target tone have responses that are suppressed by the neural fluctuations in response to the masker, due to lack of IHC saturation. Therefore, even for stimuli that would saturate the neural discharge rates across the population of AN fibers due to synaptic saturation, the rate profile across the population of IC BS neurons has a peak for CFs tuned near the target frequency. Overall, the rate profile of BS neurons provides a potential basis for a decision variable related to detection over a wide range of stimulus levels.

The results presented here show that the rate-profile across the population of model IC BS neurons is able to predict thresholds, as well as trends in thresholds across masker parameters, for several paradigms. Changes in detection threshold as a function of masker bandwidth have often been used to infer the "critical bandwidth" in the context of the PSM. Traditionally, it was assumed that this bandwidth represents the bandwidth of peripheral filters (*e.g.*, Fletcher, 1940). In the context of the NF model, this "bandwidth" is determined by a combination several factors: peripheral bandwidth, IHC saturation, cochlear gain, and sensitivity to NFs in the central nervous system. Future directions for extending this study, including re-evaluating the effects of sensorineural hearing loss, will be discussed.

METHODS

Computational Models

Predictions of psychophysical thresholds were based on response of model AN fibers (Zilany *et al.*, 2014) and model BS cells (Carney *et al.*, 2015; Carney and McDonough, 2019). A population of 21 model high-spontaneous-rate AN fibers had log-spaced CFs centered on the target tone frequency. The range of CFs was determined based on the thresholds for the fixed-level 5-Hz bandwidth masker (see Fig. 3). Thresholds for this condition depended primarily on spread-of-excitation, therefore they were directly impacted by the model's CF range. Based on this condition, the CF range was set to 2/3 octaves, and it was then fixed for all other simulations. Ten independent AN model simulations were averaged for each CF channel. The averaged responses were the inputs to the brainstem stage of the IC BS model (Carney and McDonough, 2019). Each BS model inherited its CF from its input; thus the BS model cells also had 21 log-spaced CFs. Each BS model had a band-suppressed amplitude-modulation transfer function (MTF) with a trough frequency of 100 Hz, which falls within the distribution of trough frequencies of BS neurons in rabbit (Kim *et al.*, 2020). No internal noise was added to manipulate the model thresholds for any stimulus condition; the only source of internal noise was the time-varying spontaneous rate of the AN model (Zilany *et al.*, 2009).

Stimuli and Psychophysical Studies to be Modeled

Thresholds were predicted for three masking paradigms. The first was detection of tones in flat-spectrum noise maskers that varied in bandwidth and frequency (van de Par and Kohlrausch, 1999). Tone frequencies were 250,

500, 1000, 2000, and 4000 Hz. Masker bandwidths were set to 5, 10, 25, 50, 100, 250, 500, 1000, 2000, and 4000 Hz, limited to bandwidths twice the tone frequency. Maskers were created in the frequency domain with a flat magnitude spectrum and random phase, and stimulus waveforms were generated using an inverse fast Fourier transform. Maskers and tones were gated synchronously, with a duration of 300 ms including 50-ms raised-cosine ramps. Maskers were scaled to an overall level of 70 dB SPL; tone levels were varied to find threshold. Model thresholds were compared to the diotic results in van de Par and Kohlrausch (1999).

The second experiment was detection of a 900-Hz tone in a Gaussian noise masker that varied in bandwidth, with and without a roving-level paradigm (Kidd *et al.,* 1989). Thresholds were estimated for geometrically centered maskers with bandwidths of 10, 50, 100, 300, 1000, and 3000 Hz. The spectral skirts described in Kidd *et al*. (1989) were mimicked with digital filters. Maskers and tones were simultaneously gated; stimuli were 250 ms in duration, with 20-ms raised-cosine ramps. Maskers for this experiment had a mean spectrum level of 35 dB re 20 µPa. The stimulus levels were randomly varied from interval to interval; the levels were drawn from uniform distributions with ranges of 0 (no variation), 4, or 32 dB.

The third experiment considered the co-modulation-masking release, band-widening paradigm of Hall *et al*. (1984). This paradigm involved detection of a tone in either Gaussian noise or in multiplied noise. The multiplied noise was a coherently modulated noise created by multiplying a wideband Gaussian noise by a 50-Hz lowpass-filtered Gaussian noise. The Gaussian or multiplied noises were then filtered to create maskers with bandwidths of 50, 100, 200, 400, 1000, and 2000 Hz. The filter slopes described in Hall *et al.* (1990) for the low-pass noise and for the bandpass noises were approximated digitally. Maskers and a 1000-Hz tone were simultaneously gated, with 300-ms duration and 50-ms raised-cosine ramps. Masker spectrum levels were fixed at 40-dB SPL re 20 µPa; tone levels were varied to estimate detection thresholds.

Methods for Estimating Model Thresholds

Thresholds for each condition in the above paradigms were estimated using methods similar to those in Maxwell *et al*. (2020), except as mentioned below. Notably, the previous study focused on BE IC model cells, whereas here we used only BS models. Briefly, population responses for the AN model provided inputs to the brainstem stage for a population of model IC BS cells. AN and IC BS model responses were simulated for two-interval trials; half of the trials had the target in the first interval, and half in the second to accommodate any effects of synaptic adaptation, included in the AN model. The silent duration between intervals was 500 ms. For each condition, model AN and IC templates were computed by averaging 500 responses to the masker alone. Stimuli for each test condition covered a range of signal-to-noise ratios (SNRs) spanning detection threshold. Population responses were computed for 100 trials for each SNR. For each trial, the Mahalanobis (1936) distances between the response profile for each of the two intervals and the template were computed (Duda *et al*., 2001; Maxwell *et al*., 2020). The interval with the response having a larger distance from the template was selected as the target interval, and correct/incorrect responses were tallied. Psychometric functions for %-correct as a function of SNR were computed. The psychometric functions were fit with a logistic and interpolated to estimate the SNR for 70.7% correct, corresponding to the threshold in a two-down-one-up adaptive track, as used in van de Par and Kohlrausch (1999) and Kidd *et al*. (1989) (Hall *et al.* (1984) used a three-down-one-up procedure).

RESULTS

The general pattern of detection thresholds across tone frequencies and masker bandwidths reported by van de Par and Kohlrausch (1999) was replicated by the NF model responses (Fig. 3). Because the overall level of the masker was held constant, the spectrum level changed with bandwidth. As a result, the location of the cues that supported model thresholds varied across the conditions. For narrowband maskers, which had high spectrum levels, the IHCs tuned near the target frequency were saturated for both intervals; thus, the detection was primarily possible due to "spread of excitation" of the population response to fibers tuned below and above the target frequency (Fig. 2B). For wideband maskers, spread of excitation was obscured by the masker. However, the lower spectrum levels of the wideband maskers did not saturate the IHCs, and as a result, addition of the tone caused robust changes in the NFs for AN fibers tuned near the tone. The success of this transition of cues in predicting the overall pattern of thresholds makes it clear that the band-widening experiment with fixed overall level is not a simple reflection of the bandwidth of peripheral filters. Note that the simulation results in Fig. 3 show that the AN-rate based thresholds are slightly

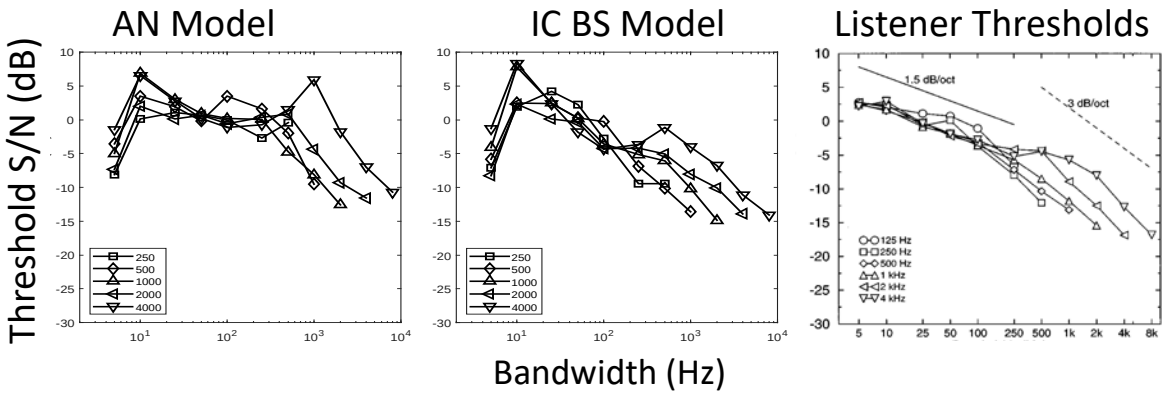


Figure 3. Simulations of AN model (left) and IC model (center) thresholds, compared to listener thresholds (right) from van de Par and Kohlrausch (1999) (permission requested).

elevated, but for this task with fixed-level maskers, the trends in AN thresholds are similar to those in the listener thresholds. The next paradigm studied provided a greater challenge for the AN-rate-based model.

Listener thresholds for masked detection in a roving-level tone-in-noise detection paradigm are surprisingly little affected by large rove ranges, a result that challenges the PSM (Kidd *et al*., 1989). The NF model successfully captured this result for masker bandwidths greater than or equal to 100 Hz. However, for the narrowband maskers, the variance in rate with sound level elevated model thresholds (Fig. 4). In the Kidd et al. (1989) experiment, the mean spectrum level was fixed at 35 dB SPL, so the narrowband maskers had low overall sound levels, i.e., below the levels that yield robust NF cues. The reduction of the NF cues for the narrowband, low-level stimuli explains the elevated model thresholds in the roving-level condition. Note that a more complete AN model with actively controlled cochlear gain (*e.g*., Farhadi *et al*., 2021) would potentially extend the level range (and thus the bandwidth range) over which the model is robust to the roving-level paradigm. Additionally, inclusion of BE model cells, in addition to the BS cells, could potentially take advantage of beating cues that arise when tones are added to low-level maskers at low SNRs, resulting in enhanced performance in the roving-level paradigm (Maxwell *et al*., 2020). These hypotheses will be tested in future work.

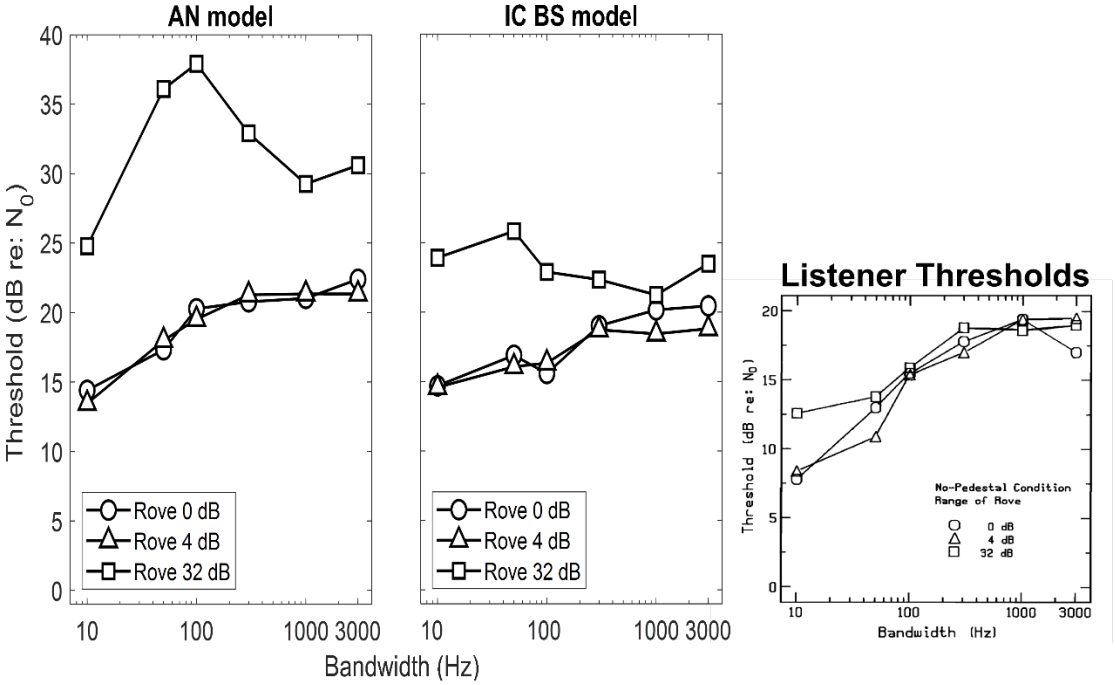


Figure 4. AN model (left) and IC model (center) thresholds, compared to listener thresholds (right) from Kidd *et al*. (1989).

Another challenge to the PSM model is the classical band-widening co-modulation release (CMR) paradigm (e.g., Hall *et al*., 1984; Haggard *et al*., 1990). In this experiment, thresholds for detection as a function of masker bandwidth are compared for two conditions: a Gaussian-noise masker, and a Gaussian noise that is multiplied by a low-pass filtered (Fc = 50 Hz) noise, referred to as multiplied noise. The slow fluctuations in the multiplied noise result in a noise stimulus that has correlated, or co-modulated, envelopes across frequency channels (Hall *et al.*, 1984). Multiplied-noise stimuli result in a larger difference in the amplitude of envelope fluctuations between stimuli with or without an added tone. Thus, a single-channel model can explain much of the CMR in the band-widening paradigm (Verhey *et al*., 1999). The NF model's thresholds based on the BS rate profile (Fig. 5B) demonstrate the same general trends, including approximately the same amount of masking release, as in the psychophysical results (Fig. 5C). Thresholds based on AN rate profiles showed similar trends (Fig. 5A), albeit with elevated thresholds.

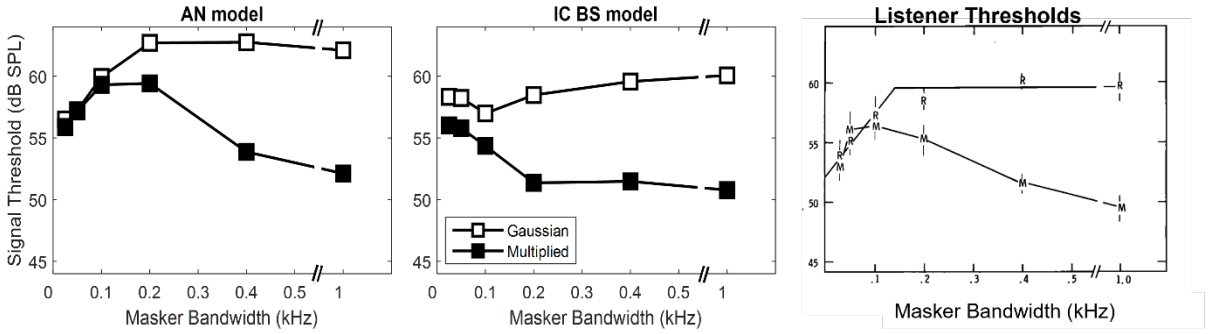


Figure 5. AN (left) and IC (center) model thresholds, compared to listener thresholds (right) from Hall *et al*. (1984).

DISCUSSION

The goal of this study was to challenge the NF model to predict thresholds in three different masked-detection paradigms. The paradigms varied in several ways – The first experiment focused on detection of tones in flat spectrum noise using fixed-level maskers over a wide range of frequencies and bandwidths (van de Par and Kohlrausch, 1999). The bandwidth variation of the fixed-level masker resulted in a wide range of spectral levels, a challenge for physiologically-based detection models due to the dynamic-range limitations of many neural cues. The Kidd *et al*. (1989) study used maskers that were randomly varied, from interval to interval, around a mean spectrum level. The small changes in listeners thresholds in the roving-level paradigm are difficult for the power-spectrum model to explain. Finally, the co-modulation masking release paradigm of Hall *et al*. (1984) was used to determine whether the fluctuation-based midbrain rate profiles in the NF model predicted appropriate amounts of masking release for multiplied vs. Gaussian-noise maskers.

In general, the NF-model detection thresholds were reasonably close to psychophysical thresholds, with appropriate trends across bandwidth, level, and frequency. Exceptions were for narrowband maskers, for one of two reasons: (1) When overall level was fixed across bandwidth, narrowband maskers were relatively intense, such that both the standard and test stimuli saturated the IHCs. In this case, fluctuation cues were minimized, and the model depended on spread-of-excitation cues across a range of CFs. (2) When the spectrum level was fixed, narrowband maskers had relatively low levels. At levels well below IHC saturation, NF cues are relatively weak, and, whereas changes in AN rate upon addition of the tone could support detection, rate-based cues are vulnerable to the roving-level paradigm. Both of these situations would be partially addressed by an AN model that included efferent control of cochlear gain. Future work will take advantage of an AN model that includes both level-related and fluctuation-related efferent gain control, based on anatomical pathways from both the cochlear nucleus and IC to the medial olivocochlear system (Farhadi *et al*., 2020).

The NF model used in this study focused on BS IC neurons. Previous models that depend on fluctuation-related cues tend to use modulation filters or neural models that are excited by fluctuations, as opposed to neural models for cells that are *suppressed* by modulations. However, the BS neurons have an interesting advantage: these neurons represent moderate to high-level spectral peaks, or target tones, based on reduced fluctuations associated with captured AN responses, and they represent low-level peaks based simply on the excitation pattern that ascends to the midbrain. The features of these neurons thus support the NF model over a relatively wide range of levels. In addition to the

efferent gain control mentioned above, a more comprehensive NF detection model could include both BS and BE neurons, as well as neurons having a range of best modulation frequencies (Maxwell *et al*., 2020).

The NF model achieved thresholds in the correct range and with appropriate trends, except for the narrowest band maskers, as mentioned above. However, it is worth noting that for this model, the predictions of the basic band-widening results are not explained by the shape of an underlying peripheral filter, nor are they constrained to frequency channels near the target tone frequency. Depending on the level of the stimulus, the cues used by the NF model may occur in channels near the tone frequency, primarily in the form of fluctuation cues, but they may also occur in channels far from the tone frequency, in the form of "spread of excitation" cues. Furthermore, the "on-CF" cues are only partly explained by the shape of the underlying filter. These cues are also strongly affected by IHC saturation, by cochlear gain that determines the operating point of the nonlinear IHC, and by synaptic saturation. Additionally, the separate saturation of the AN discharge rates is beneficial to the NF model because it effectively normalizes the AN average-rate cues, making the fluctuation cues more robust across levels.

In the framework of the NF model, it is interesting to consider that the effect of sensorineural hearing loss on detection thresholds may not be simply to broaden peripheral filters, but rather to distort the fluctuation profile. In particular, reduced cochlear gain affects the operating point of the IHC nonlinearity, resulting in stronger overall fluctuations and reduced capture (Miller *et al*., 1997). These changes associated with hearing loss would reduce the contrast in the NF profile along the tonotopic axis of AN fibers, and ultimately reduce the contrast across the population rate profile in the midbrain. To address these deficits would require a different strategy from the typical algorithms in hearing aids and cochlear implants, which typically focus on gain control designed to achieve audibility.

ACKNOWLEDGMENTS

Supported by NIH-NIDCD-R01010813.

REFERENCES

1. L. H. Carney, *JARO* **18**, 331-352 (2018).
2. L. H. Carney, L. H., T. Li, and J. M. McDonough, *Eneuro* **2**(4) (2015)
3. L. H. Carney and J. M. McDonough, *Attention, Perception, & Psychophysics* **81**, 1034-1046 (2019).
4. T. Dau, B. Kollmeier and A. Kohlrausch, JASA **102**, 2892-2905 (1997).
5. L. Deng, C. D. Geisler and S. Greenberg, *JASA* **82**, 1989-2000 (1987).
6. R. O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification*, 2nd ed (Wiley, New York), p. 35. (2001).
7. L. Fan, K. S. Henry and L. H. Carney, *Hearing Research* **409** 108328, 1-14 (2021).
8. A. Farhadi, S. G. Jennings, E. A. Strickland and L. H. Carney, in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing* (IEEE, 2021), pp. 291-295.
9. D. M. Green, *Profile analysis: Auditory intensity discrimination* (No. 13) (Oxford University Press, 1988).
10. M. P. Haggard, J. W. Hall III and J. H. Grose, *JASA* **88**, 113-118 (1990).
11. J. W. Hall, M. P. Haggard and M. A. Fernandes, *JASA* **76**, 50-56 (1984).
12. M. L. Jepsen, S. D. Ewert and T. Dau, *JASA* **124**, 422-438 (2008).
13. P. X. Joris, C. E. Schreiner and A. Rees, *Physiological reviews* **84**, 541-577 (2004).
14. G. Kidd, Jr., C. R. Mason, M. A. Brantley and G. A. Owen, *JASA* **86**, 1310-1317 (1989).
15. D. O. Kim, L. H. Carney and S. Kuwada, *Journal of Neurophysiology* **124**, 1198-1215 (2020).
16. J. J. Lentz, V. M. Richards and M. R. Matiasek, *JASA* **106**, 2779-2792 (1999).
17. P. C. Mahalanobis, National Institute of Science of India (1936).
18. S. van de Par and A. Kohlrausch, *JASA* **106**, 1940-1947 (1999).
19. B. N. Maxwell, V. M. Richards and L. H. Carney, *JASA* **147**, 3523-3537 (2020).
20. R. L. Miller, J. R. Schilling, K. R. Franck and E. D. Young, *JASA* **101**, 3602-3616 (1997).
21. J. L. Verhey, T. Dau and B. Kollmeier, *JASA* **106**, 2733-2745 (1999).
22. X. Zhang, M.G. Heinz, I.C. Bruce and L.H. Carney, *JASA* **109**, 648-670 (2001).
23. M. S. Zilany, I. C. Bruce, *JASA* **122**, 402-417 (2007).
24. M. S. Zilany, I. C. Bruce and L. H. Carney, *JASA* **135**, 283-286 (2014).
25. M. S. Zilany, I. C. Bruce, P. C. Nelson and L. H. Carney, *JASA* **126**, 2390-2412 *(2009)*.

Convergent Otoacoustic Tuning Estimates in the Anole Lizard?

Rebecca Whiley,[1, a)] and Christopher Bergevin[2, b)]

[1)]*Department of Biology, York University, Toronto, Ontario, Canada*
[2)]*Department of Physics and Astronomy, York University, Toronto, Ontario, Canada*

[a)]*Corresponding author: rwhiley@yorku.ca*
[b)]*Electronic mail: cberge@yorku.ca*

Abstract.
Spontaneous otoacoustic emissions (SOAEs) are one hallmark of an active ear and provide insight into the underlying biomechanics. SOAEs interact readily with external tones and correlate with stimulus-frequency emissions (SFOAEs) evoked by low-level stimuli ($\leq$ 40 dB SPL). Such findings are consistent with a canonical model of emission generation based upon the principle of coherent reflection, which predicts relationships between OAE-derived tuning estimates. These predictions, in part, have been empirically validated even for non-mammalian vertebrate ears that lack the classic (mammalian) basilar membrane travelling wave. However, inconsistencies persist and require additional research, such as the incompatible tuning estimates between SOAE "suppression-tuning curves" (STCs) and stimulus frequency emission (SFOAE) phase-gradient delays (N_{sf}). Further complications stem from the nonlinear, stimulus level-dependent nature of SFOAE delays. To reconcile these uncertainties and elucidate the relationship between otoacoustic tuning estimates, we examined correlations between SOAE and SFOAE tuning estimates in the green anole lizard (*Anolis carolinensis*). Specifically, we characterize the level-dependence of N_{sf} across frequency for swept-tone stimuli while extracting SOAE "interaction tuning curves" (ITCs) measured simultaneously. These allow us to quantify both N_{sf} and Q_{10dB} values of the ITCs from the same ear using the same stimuli. Preliminary results suggest that it is possible to obtain convergent otoacoustic tuning estimates from the same data.

BACKGROUND

Spontaneous otoacoustic emissions (SOAEs) occur across tetrapod classes (e.g., [1]) and provide an objective, non-invasive method for investigating the biomechanics underlying active hearing. Further, SOAEs are closely tied to peripheral auditory function, correlating with psychoacoustic measures such as threshold microstructure [2, 3, 4]. SOAEs interact readily with external tones, and this characteristic can be exploited to create "suppression tuning curves" (STCs; [5]) that are comparable to auditory nerve fibre tuning in a variety of species (e.g., barn owls [6]; Bobtail skinks [7]; macaque monkeys [8]). Clinical applications of SOAEs are scant because of their inherent idiosyncrasy; while unique to an individual ear when present, SOAEs are absent from many ears [9]. Fortuitously, healthy ears typically still produce evoked otoacoustic emissions (eOAEs), such as stimulus-frequency otoacoustic emissions (SFOAEs).

SFOAE magnitudes and phase-gradient delays (expressed in number of stimulus cycles, N_{sf}) show clear correlations with SOAE peaks when evoked with stimuli below 40 dB SPL [10, 11]. Both N_{sf} and STC-derived quality factor (Q) values correlate with tuning measured directly from auditory nerve fibres (N_{sf} and neural Q [12, 13]; STC Q and neural Q [14, 15]). Additionally, there are correlations between N_{sf} values and N_{SOAE} values (an SOAE tuning estimate based on interpeak spacing; [14]). Conversely, SFOAE phase-gradients do not produce consistent tuning estimates with STCs when measured in humans [16]. Rather, Manley and van Dijk [16] reported greater similarities between STC-derived Q values and psychophysical simultaneous masking paradigms [17]. The lack of a relationship between STC-derived Q values and N_{sf} is difficult to interpret given documented correlations between N_{sf} and other tuning measures, including neural tuning and forward masking [18, 19].

To reconcile some of the discrepancies in otoacoustic tuning estimates, we simultaneously measured SOAEs and SFOAEs in green anole lizards (*Anolis carolinensis*). Anole emissions are robust and qualitatively similar to those of mammals, despite disparate inner ear morphologies, providing an opportunity to assess tuning correlations in a relatively simpler model. Building upon previous work (e.g., [12]), we compared SFOAE delays (N_{sf}) to Q_{10dB} values from SOAE tuning curves, expecting compatible tuning estimates since both measures come from one response to swept-tone stimuli. We refer to these SOAE tuning curves as "interaction" tuning curves (ITCs) since a stimulus may not necessarily suppress SOAE activity per se but instead entrain it. Here, we present preliminary SOAE and SFOAE tuning estimates from a single individual to examine the extent to which these values converge.

METHODS AND RESULTS

Characteristics of Emissions

For brevity, we report data here from a single animal. There was nothing exceptional about this particular animal's OAEs, and thus we anticipate the results reported here to be representative. The lizard was lightly anesthetized and otoacoustic measurements were acquired following a methodology consistent with previous descriptions [14, 20]. SFOAEs were evoked by swept tone stimuli from 0.5 to 5.3 kHz (2.4-second sweep duration) and extracted via a suppression paradigm [21]. Probe levels (L_p) ranged from 10 to 60 dB SPL in 5 dB steps. Suppressor stimuli were 40 Hz higher than probe frequencies and 15 dB higher in level (L_s). Thirty-four temporal averages were obtained at each level.

SOAE activity was determined in a two-fold manner. First, SOAEs were recorded in the absence of external stimuli between SFOAE presentations (at different levels). These measurements indicated that peak characteristics were stable and showed nominal variation throughout the experiment. For this animal, seven peaks were identified in the SOAE spectrum, with frequencies from 1.2 to 4.1 kHz and amplitudes from -8 to 0 dB SPL (Fig.1, blue trace). Second, since SOAEs were present during the swept-tone stimulation, those waveforms provided the chief means by which we quantified SOAE activity in this report. Peak waveforms were calculated from a zero-phase bandpass filtered region around a suspected peak that was inverse transformed to compute the analytic signal via the Hilbert transform [6]. Peak frequency and level were defined as the instantaneous frequency and envelope, respectively, and used as reference SOAE parameters for interaction analyses described below. SOAE structure showed qualitative similarities with SFOAE recorded at low stimulus levels (15 to 40 dB SPL). Localized increases in SFOAE amplitudes and non-monotonic growth across stimulus levels occurred near SOAE peak frequencies (Fig.1).

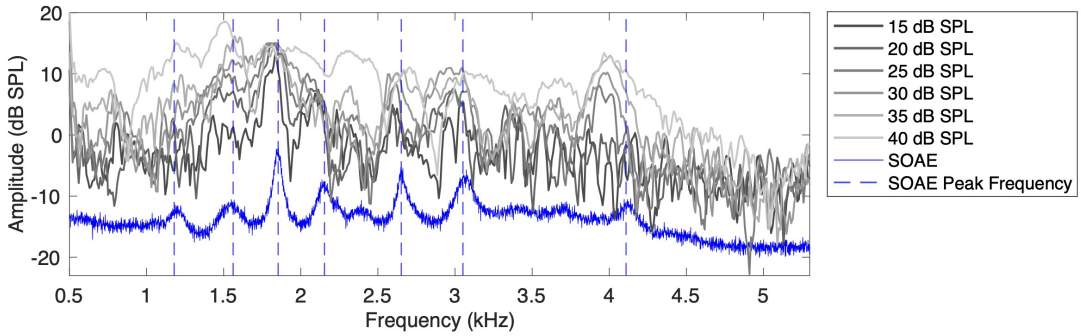


FIGURE 1. Structural similarities between an SOAE (blue) and SFOAE amplitudes in an anole lizard (lighter shades of grey reflect higher stimulus levels). SFOAEs were evoked by swept tone stimuli (0.5 to 5.3 kHz) from 15–40 dB SPL. Peaks in the SOAE (n = 7) were identified via a custom peak-picking routine (vertical dashed blue lines; see text for details). Note that while SFOAE phase is not shown here, it is used for determination of N_{sf} as shown in Fig.3.

SOAE Interactions with External Tones

To evaluate swept tone interactions with SOAE peaks, we created spectrograms from the SFOAE responses using short-time Fourier transforms with an 8192-point time window with a sliding overlap of 97% (responses to stimuli between approximately 0.68 to 5.16 kHz were evaluated from hereon; see spectrograms in Figure 1B from [14] for example). We used the residual SFOAE response (i.e., $residual = vp + vs - vps$) to ensure that we calculated only the SOAE interaction with the swept tone and did not include the stimulus itself [21]. Filtered spectrograms were composed of 403 points in the time domain, each representing one of the stimulus frequencies presented, providing a frequency resolution of approximately 10 Hz in the stimulus domain. We interpret our sweep duration to be slow enough such that the response to each stimulus frequency can be considered steady-state.

SOAE interaction spectrograms were quantified as the reference SOAE peak level subtracted from the measured response level at the reference SOAE peak frequency across all stimulus frequencies and levels presented (total of 4433

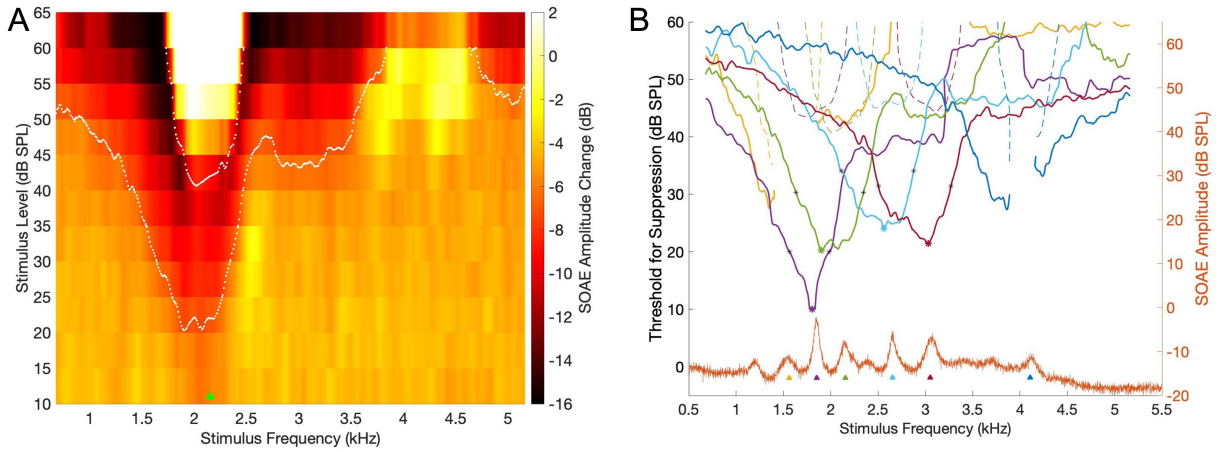


FIGURE 2. Characterizing SOAE interactions with swept tone stimuli to create tuning curves. (**A**) Representative SOAE interaction tuning spectrogram at SOAE peak frequency 2.15 kHz (indicated by green triangle). Colour bar represents changes in response amplitudes at 2.15 kHz in response to stimulus of a given level and frequency minus the reference amplitude (-4.4 dB SPL). Spectrogram demonstrates regions of facilitation and suppression (white markers outline the boundaries of the inner facilitation region and the outer suppression region). There is a central region of facilitation about 2.2 kHz (light yellow shading outlined by inner white markers) and a second, subharmonic region of facilitation when the stimulus is twice the peak frequency, approximately 4.3 kHz (bright yellow region that creates a discontinuity in the outer white markers). Suppression regions surround the central facilitation region (darker red shading outlined by outer white markers). The tuning curve at 2.15 kHz (green line in panel B) was obtained at the boundaries of the outer suppression region at a threshold value of -7.5 dB. (**B**) Interaction tuning curves (ITCs) at SOAE peak frequencies (represented by triangles along the x-axis in the same colour) plotted above the SOAE spectrum (orange). Suppression regions of the ITC are plotted as solid lines and corresponding facilitation regions for that ITC are plotted as dashed lines. For ITCs where the tip region was characterized, tip frequency is identified by stars at the lowest point of the tuning curve in the region of the corresponding peak frequency. Black stars on each ITC represent the frequencies along the curve that correspond to 10 dB above the ITC tip level. These bandwidth values are used to calculate Q_{10dB} for each tuning curve.

stimulus conditions). Since peak frequency can fluctuate as much as $\pm$ 0.2 kHz due to interactions with the stimulus, spectrogram amplitudes were smoothed by 4- and 2-point moving averages in the frequency and level dimensions, respectively. Although SOAE interaction spectrograms can be created across the response spectrum using continuous swept-tone stimuli, we focussed on spectrograms at the reference SOAE peak frequencies (see the example for reference peak at 2.15 kHz in Figure 2A). The interaction spectrogram is composed of a central region of facilitation about the reference peak frequency (i.e., calculated responses exceed reference peak amplitude), corresponding to the SFOAE that emerges at that frequency, and a surrounding suppression region where the underlying SOAE peak is suppressed by external stimuli (i.e., calculated responses are below reference peak amplitude). Additionally, a second subharmonic region of facilitation emerges when the stimulus frequency is twice the reference peak frequency (at approximately 4.3 kHz in Figure 2A).

SOAE ITCs were calculated from the interaction spectrograms using the *contour* function in Matlab. Threshold values for determining the ITC differed across peaks and were defined as the unsmoothed mean response envelope at the reference peak frequency at the lowest stimulus level (10 dB SPL) subtracted from the reference peak envelope to normalize the noise floor. An additional criterion of 3 dB suppression was then added, producing a mean threshold value of -7.68 dB. Because the software can return multiple contours for each interaction spectrogram (due to noise in the recording, or subharmonic facilitation), only contours that contained at least 50 values were used to define the ITC. ITCs were created for six of the seven peaks identified in the reference SOAE spectrum (Figure 2B). In most tuning curves, the suppression region of the ITC (solid lines) could be clearly distinguished from the facilitation region (dashed lines). Broader, lower amplitude peaks in the spectrum were less sensitive to suppression than higher amplitude peaks (e.g., compare ITCs at 4.11 versus 1.85 kHz). For lower amplitude peaks, the shape of the suppression region of the ITC at low levels could be masked by the facilitation region, resulting in a "split" suppression region that was discontinuous about the SOAE peak (see peaks at 1.56 [yellow] and 4.11 kHz [dark blue] in Figure 2B).

Comparing Otoacoustic Tuning Estimates

Here, we focus on the suppression region of the ITC to facilitate comparisons with previous studies and SFOAE phase-gradient delays. ITC tip frequency and level were characterized to calculate quality factor Q_{10dB} for the four ITCs with continuous suppression regions. SFOAEs from 0.5 to 4.5 kHz at 20 to 40 dB SPL were used to calculate N_{sf} values fitted with locally weighted regression trends to evaluate tuning estimates across the spectrum (Figure 3; [21]). Data that passed an SNR criterion of 8 dB were included in calculations, excluding responses above 4.5 kHz and other stimulus levels at which SFOAEs were measured. N_{sf} trends exhibited level-dependence and decreased with increasing stimulus levels. Although Q_{10dB} values were only available for a limited subset of peaks, these values fall within the regime of N_{sf} values around 2 to 3 kHz. While the Q_{10dB} appear to most closely match N_{sf} for the 35 dB SPL probe level condition, we note however that the equivalent-rectangular-bandwidth quality factor (Q_{ERB}) is approximately 1.8 times larger than Q_{10dB}. Thus, the SOAE ITC tuning estimates as Q_{ERB} would most closely match the 25-30 dB SPL stimulus level delays.

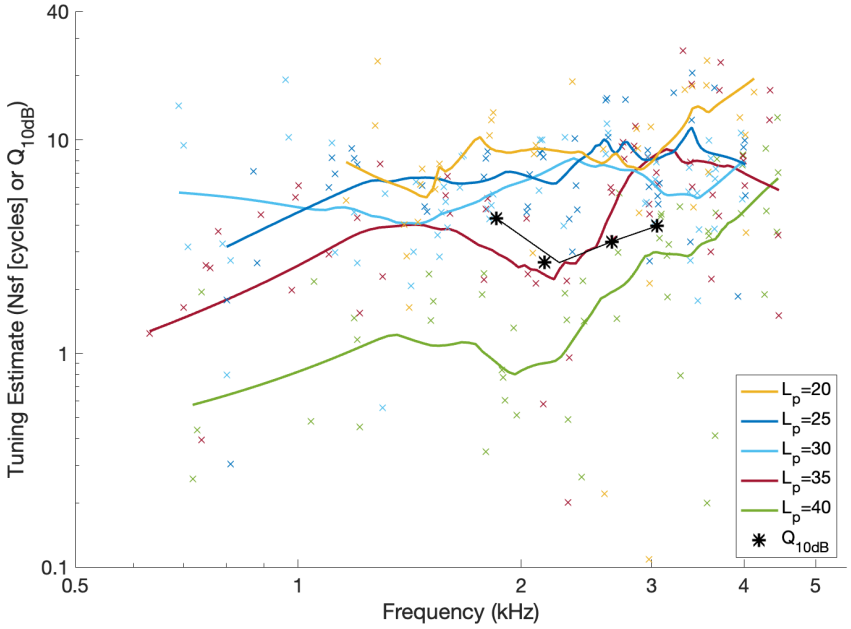


FIGURE 3. Comparing otoacoustic tuning estimates via SFOAE N_{sf} and SOAE ITC Q_{10dB} in an anole lizard. Individual N_{sf} values for probe levels at 20, 25, 30, 35, and 40 dB SPL are plotted as "x" and fitted with locally weighted regression trends (solid lines). Only values that exceeded the noise floor by 8 dB were used to calculate curves. Q_{10dB} values calculated from continuous suppression regions of ITCs are plotted at corresponding SOAE peak frequencies (black stars) and fitted with a locally weighted regression trend (solid black line).

DISCUSSION

We evaluated the level-dependence of SFOAE phase-gradient delays (as measured in number of stimulus cycles, N_{sf}), which demonstrated patterns that were consistent with past findings in anoles [14] and humans [10, 22]. For a fixed probe level, delays increased with stimulus frequency, similar to direct measures of tuning [14, 15]. We compared Q_{10dB} to SFOAE tuning estimates at and below 40 dB SPL because correlations between SFOAEs and SOAEs are well-documented at low levels, in line with theoretical framework predictions related to coherent reflection [10, 23]. Contrary to past reports from humans [16] and barn owls [6], our preliminary results in anole suggest that simultaneous ITC Q_{10dB} and N_{sf} tuning estimates match up well, although further study is needed.

A few features of our ITC methodology could explain discrepancies with other studies. First, we obtained characterizations of tuning estimates at a higher frequency resolution than previous work by using swept-tone stimuli. We covered 4433 stimulus frequencies between 0.5 to 5.3 kHz compared to 70 stimulus frequencies between 0.5 to 10 kHz in humans [16]. Since SOAE peaks are related to variations in threshold microstructure [2], measurements at a higher frequency resolution might better characterize interaction tuning curves. This approach also provides the opportunity to extract tuning curves beyond SOAE regions to extend our tuning estimates across the spectrum. Second is that our method estimates from one individual, whereas most past experiments (e.g., [6, 16]) rely on aggregate tuning estimates from multiple individuals. Given the relationship between SOAE and SFOAE structure, particularly discontinuities in SFOAEs around SOAE peaks and the idiosyncrasy of SOAEs, grouped data could mask correlations between tuning estimates within individuals. Lastly, we obtained SOAE- and SFOAE-based tuning measures simultaneously. SOAEs are inherently entwined with evoked emissions like SFOAEs because of their robust interactions with external stimuli [12]. Consequently, these emissions are closely related, as previously demonstrated by correlations in emission structure [10, 11] and some tuning estimates [14]. However, SOAEs have nonlinear interactions with stimuli, making it difficult to disentangle them from SFOAEs and complicating our interpretation of their relationship. The ITC methodology provides a means of accounting for the non-linearity of SOAE interactions with external tones by deriving it from our measurements of SFOAEs. Our preliminary results suggest that convergent otoacoustic tuning estimates might be obtainable by controlling for these interactions via simultaneous measurements.

ACKNOWLEDGEMENTS

CB and RW were supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Grant No. RGPIN-430761-2013.

REFERENCES

1. C. Bergevin, D. M. Freeman, J. C. Saunders, and C. A. Shera, "Otoacoustic emissions in humans, birds, lizards, and frogs: Evidence for multiple generation mechanisms," Journal of Comparative Physiology A **194**, 665–683 (2008).
2. G. R. Long and A. Tubis, "Investigations into the nature of the association between threshold microstructure and otoacoustic emissions," Hearing Research **36**, 125–138 (1988).
3. C. L. Talmadge, A. Tubis, G. R. Long, and P. Piskorski, "Modeling otoacoustic emission and hearing threshold fine structures," The Journal of the Acoustical Society of America **104**, 1517–1543 (1998).
4. R. R. Baiduc, J. Lee, and S. Dhar, "Spontaneous otoacoustic emissions, threshold microstructure, and psychophysical tuning over a wide frequency range in humans," The Journal of the Acoustical Society of America **135**, 300–314 (2014).
5. E. Schloth and E. Zwicker, "Mechanical and acoustical influences on spontaneous oto-acoustic emissions," Hearing Research **11**, 285–293 (1983).
6. S. Engler, C. Köppl, G. A. Manley, E. de Kleine, and P. van Dijk, "Suppression tuning of spontaneous otoacoustic emissions in the barn owl (*Tyto alba*)," Hearing Research **385**, 107835 (2020).
7. G. A. Manley and C. Köppl, "What have lizard ears taught us about auditory physiology?" Hearing Research **238**, 3–11 (2008).
8. C. A. Shera, C. Bergevin, R. Kalluri, M. Mc Laughlin, P. Michelet, M. van der Heijden, and P. X. Joris, "Otoacoustic estimates of cochlear tuning: testing predictions in macaque," in *AIP conference proceedings*, Vol. 1403 (American Institute of Physics, 2011) pp. 286–292.
9. C. Talmadge, G. Long, W. Murphy, and A. Tubis, "New off-line method for detecting spontaneous otoacoustic emissions in human subjects," Hearing Research **71**, 170–182 (1993).
10. C. Bergevin, A. Fulcher, S. Richmond, D. Velenovsky, and J. Lee, "Interrelationships between spontaneous and low-level stimulus-frequency otoacoustic emissions in humans," Hearing Research **285**, 20–28 (2012).
11. G. Martin, B. Lonsbury-Martin, R. Probst, and A. Coats, "Spontaneous otoacoustic emissions in a nonhuman primate. I. basic features and relations to other emissions," Hearing Research **33**, 49–68 (1988).
12. C. Köppl and G. A. Manley, "Spontaneous otoacoustic emissions in the bobtail lizard. II: Interactions with external tones," Hearing Research **72**, 159–170 (1994).
13. G. A. Manley, "Evidence for an active process and a cochlear amplifier in nonmammals," Journal of Neurophysiology **86**, 541–549 (2001).
14. C. Bergevin, G. A. Manley, and C. Köppl, "Salient features of otoacoustic emissions are common across tetrapod groups and suggest shared properties of generation mechanisms," Proceedings of the National Academy of Sciences **112**, 3362–3367 (2015).
15. C. J. Sumner, T. T. Wells, C. Bergevin, J. Sollini, H. A. Kreft, A. R. Palmer, A. J. Oxenham, and C. A. Shera, "Mammalian behavior and physiology converge to confirm sharper cochlear tuning in humans," Proceedings of the National Academy of Sciences **115**, 11322–11326 (2018).
16. G. A. Manley and P. van Dijk, "Frequency selectivity of the human cochlea: Suppression tuning of spontaneous otoacoustic emissions," Hearing Research **336**, 53–62 (2016).
17. B. R. Glasberg and B. C. Moore, "Derivation of auditory filter shapes from notched-noise data," Hearing Research **47**, 103–138 (1990).
18. C. A. Shera, J. J. Guinan, and A. J. Oxenham, "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements," Proceedings of the National Academy of Sciences **99**, 3318–3323 (2002).

19. C. A. Shera, J. J. Guinan, and A. J. Oxenham, "Otoacoustic estimation of cochlear tuning: Validation in the chinchilla," Journal of the Association for Research in Otolaryngology **11**, 343–365 (2010).
20. C. Bergevin, A. Mason, and N. Mhatre, "Evidence supporting synchrony between two active ears due to interaural coupling," The Journal of the Acoustical Society of America **147**, EL25–EL31 (2020).
21. R. Kalluri and C. A. Shera, "Measuring stimulus-frequency otoacoustic emissions using swept tones," The Journal of the Acoustical Society of America **134**, 356–368 (2013).
22. C. L. Talmadge, A. Tubis, G. R. Long, and C. Tong, "Modeling the combined effects of basilar membrane nonlinearity and roughness on stimulus frequency otoacoustic emission fine structure," The Journal of the Acoustical Society of America **108**, 2911–2932 (2000).
23. C. A. Shera, "Mammalian spontaneous otoacoustic emissions are amplitude-stabilized cochlear standing waves," The Journal of the Acoustical Society of America **114**, 244–262 (2003).

A Comparison of Implantable Microphones Constructed Around a Piezoelectric Polymer

John Z Zhang,[1,2] Aaron J Yeiser,[1,2] Annesya Banerjee,[1,3,4] Benjamin G Cary,[2] Christopher I McHugh,[4] Lukas Graf,[4] Ioannis Kymissis,[5] Elizabeth S Olson,[5,6] Hideko Heidi Nakajima,[3,4] and Jeffrey H Lang[2]

[1)] *Authors contributed equally to this work*
[2)] *School of Engineering, Massachusetts Institute of Technology*
[3)] *Speech and Hearing Bioscience & Technology Program, Harvard*
[4)] *Otolaryngology-Head and Neck Surgery, Harvard Medical School and Massachusetts Eye and Ear*
[5)] *School of Engineering and Applied Science, Columbia University*
[6)] *Otolaryngology-Head and Neck Surgery, Columbia University*

Abstract. Cochlear implants are neural prosthetic devices that restore hearing in individuals with severe-to-profound sensorineural hearing loss. As of 2019, an estimated 736,900 cochlear implants had been deployed throughout the world. In conventional cochlear implants, the microphone and speech processing electronics are located externally. This external placement has functional and cosmetic drawbacks. Generally, cochlear implant users must remove the external components for showering, swimming, or participating in certain sports. Users typically do not wear the device during sleep, preventing them from detecting auditory danger signals and influencing auditory development in children. Additionally, the acoustics of the outer ear and ear canal provide information that enhances sound localization and signal discrimination in noisy environments. These shortcomings could be addressed by a fully implantable device. With technological advances in batteries, electronics and wireless charging, all components could potentially be implantable. However, internalizing the microphone remains a challenge. We have been developing implantable microphones based on a piezoelectric material, polyvinylidene fluoride (PVDF) and its co-polymer PVDF-trifluorethylene (PVDF-TrFE), available as a thin film. The internal placement of the microphones takes advantage of natural enhancement and filtering of the outer and middle ear. We have developed three prototype devices: (1) the *coch-mic*, which is inserted into the cochlea much as a cochlear implant, senses cochlear pressure as a hydrophone using a strip of silicone-coated PVDF-TrFE film, (2) the *drum-mic*, which senses the motion of the umbo (the tip of the malleus), has a circular PVDF membrane mounted onto a cylindrical base inserted between the umbo and the cochlear promontory, and (3) the *cantilever-mic*, which is a substrate-mounted PVDF beam where one end senses umbo motion, and the other end is anchored to the bone of the skull.

INTRODUCTION

Current cochlear implants use state-of-the-art *external* electret condenser microphones. To achieve totally implantable cochlear implants, the challenge of a fully implantable microphone must be overcome. Assistive hearing systems combining a fully-implantable microphone and electronics with a cochlear implant would enhance directional and focused hearing by taking advantage of outer and middle ear mechanics. They would be usable in all environmental conditions throughout the day and night. Various approaches to the fully implantable microphone have been explored, including MEMS-based, piezoelectric, and optical strategies [1]. However, current implantable microphones suffer from unstable mechanics, poor signal-to-noise ratio (SNR), low bandwidth, and/or internal body noise.

Piezoelectric films, such as polyvinylidene fluoride (PVDF), are a promising candidate for the sensitive transducer material needed in the miniaturization and implantation of microphones. The direct piezoelectric effect allows us to create sensors that convert mechanical motion into charge [2]. The mechanical deformation can be produced through a pressure boundary condition or a displacement boundary condition [3]. PVDF is a semi-crystalline polymer which is piezoelectric in the beta phase [4]. Although piezofilms exhibit weaker piezoelectric coupling than ceramic piezoelectric materials, PVDF's advantages for implantable microphones are that it is safe, easily obtainable commercially, and is flexible due to its form as a thin film [5, 6]. Furthermore, PVDF can provide better electromechanical coupling to the structures of the middle and inner ear due to its mechanical impedance. Through analytical derivations we have previously shown that the co-polymer, PVDF-Trifluoroethylene (PVDF-TrFE), can show higher sensitivity when used as a pressure sensor in fluid [7]. Charge amplifiers are preferred for use with piezoelectric sensors as their sensitivity is more stable than voltage amplifiers [8]. We have developed a custom differential charge amplifier with miniature coaxial connectors that is designed specifically to maximize SNR and attenuate electromagnetic interference (EMI) when used with a shielded design such as the *cantilever-mic*. Amplifier design considerations are the subject of another paper. Here we describe the development of a PVDF-based microphone with our three microphone types.

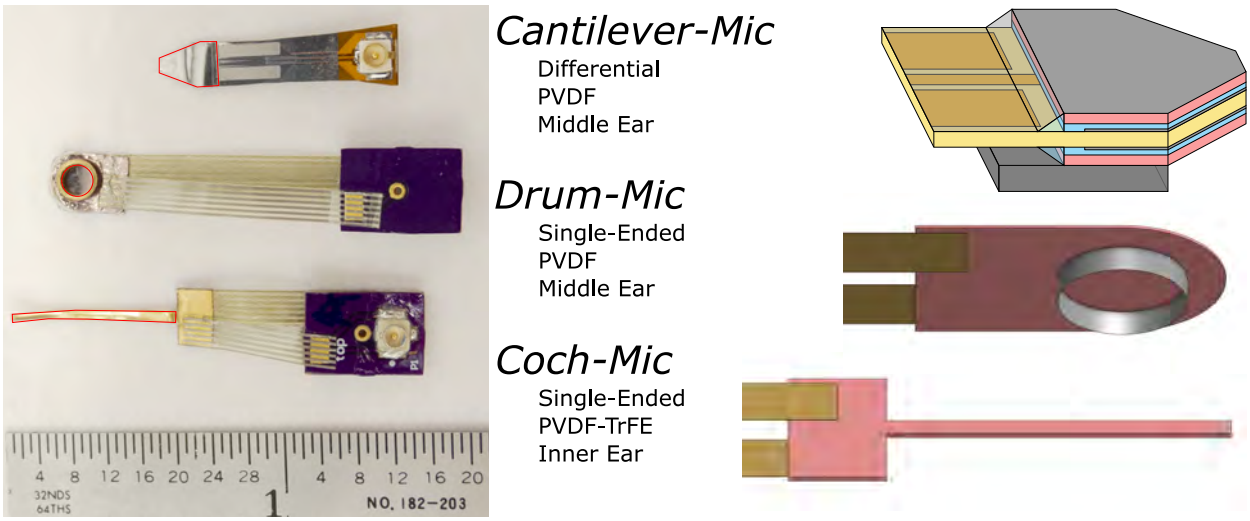


FIGURE 1. Comparison of our implantable microphone styles. The scale numbers mark off inches. The active element of each sensor is demarcated in red. We show stack-ups of the three microphones on the right with the active element of PVDF or PVDF-TrFE colored in pink.

The *coch-mic* is designed to produce charge in response to fluid pressure, whereas the *drum-mic* and *cantilever-mic* produce charge in response to stretching and/or bending. The charge amplifier is designed to enhance SNR of these piezoelectric sensors and is the subject of a forthcoming paper. These sensors have been modeled analytically and numerically, bench-tested, and tested in human cadaveric specimens. Results have been presented in conference proceedings, journal papers, and theses [7, 9, 10, 11, 12],. In this paper, we review the devices and compare their operation in terms of ease of placement, sensitivity, and the ability to improve performance through engineering enhancements. We illustrate the different microphone geometries, construction, and implantation in Figure 1.

IMPLANTABLE MICROPHONE TAXONOMY

We have developed microphones for implantation in both the middle ear and the inner ear. In this section, we discuss briefly the fabrication and operation of these microphones. We also describe the procedures for collecting data to characterize performance in a human cadaveric temporal bone. All three sensors use U.FL miniature coaxial connectors to interface with the differential charge amplifier. When connected to the *coch-mic* and *drum-mic* the charge amplifier is used in single-ended mode. The *coch-mic* and *drum-mic* use hot-bar bonded heat seal flex wires to connect to the PVDF without damaging or depoling the material. The PVDF and PVDF-TrFE are 50 µm in thickness with 100 nm Nickel-Copper metallized electrodes, purchased from PolyK Technologies (State College, PA).

Intracochlear Microphone

The *coch-mic* consists of a 15 mm by 0.5 mm strip of 50 µm PVDF-TrFE. There is an additional 3 mm by 5 mm area of desensitized PVDF-TrFE that is used to bond the heat-seal flex wire connectors. The outline of the sensor is cut using a femto-second laser cutting service provided by the PVDF vendor. The sensor is coated using an aerosol 422B Silicone Modified Conformal Coating. Two layers are applied to both the flex wire, the sensor, as well as the small PCB with the coax connector. The conformal coating is necessary to prevent shorting of the sensor when submerged in saline, which will rail the output of the charge amplifier resulting in no signal. During TB experiments, the strip part of the sensor is inserted into the scala tympani via the round window. After successful insertion, the round window is sealed using alginate impression material.

Drum Microphone

The *drum-mic* is effectively a miniature drum with a piezoelectric PVDF drum head. Motion of the umbo causes the PVDF to stretch, generating charge that is detected by our custom low-noise charge amplifier. We build the *drum-mic* by tacking a single piece of 50 µm PVDF to a 3 mm-diameter brass ring a small amount of cyanoacrylate glue, and then securing the outside of the drum with UV-curing epoxy. The drum's height of about 1.3 mm is chosen to ensure that the drum fits snugly between the promontory and the umbo. Flex wires are hot-bar bonded to a 4×6mm tab of the PVDF drum head, electrically connecting the drum to a PCB with a coaxial connector. Both the flex wires and the drum head are cut on a Silhouette Cameo vinyl cutter.

Cantilever Microphone

The *cantilever-mic* is a roughly triangular piezoelectic cantilever approximately 3 mm long, with the tip resting on the umbo. The sensor consists of two layers of 50 µm PVDF sandwiching an inert 100 µm Kapton flex-PCB core. Umbo motion causes the cantilever to bend, compressing one layer of PVDF and tensioning the other, generating a differential signal transmitted through U.FL connectors soldered to the flex-PCB. The sensor is sputter-coated with 200 nm of chromium, which serves both as a ground shield and as a ground electrode for the PVDF, providing excellent EMI resistance when paired with the sensor's differential output. Since the charge-sensitive electrodes are sealed inside the sensor between the PVDF and the Kapton, no additional waterproofing is needed.

MICROPHONE PERFORMANCE

We implant our microphones in human cadaveric temporal bones as shown in Figure 2 to determine microphone performance. We use a Beyerdynamic DT-48 Headphone Driver as a speaker coupled to the ear canal using through a flexible hose. A Knowles EK-3103 microphone is used to measure the ear canal pressure used as reference. A Polytec Laser Doppler Vibrometer (LDV) is used to measure the umbo motion for *drum-mic* and *cantilever-mic* experiments and stapes motion for *coch-mic* experiments. The LDV measurement is critical in determining linearity and functionality of the ear when the microphones are implanted in the cadaveric specimen. A National Instruments PXI computer with an Analog Input 16-bit and Output Module 24-bit is used as the Data Acquisition System (DAQ). All measurements are made using a measurement software written in LabVIEW. We summarize performance of the different microphones in Figure 3 and Figure 4. We compare the equivalent input noise of our microphones with the Sonion 65GC31, a state-of-the-art commercial electret microphone used by some cochlear implants.

DISCUSSION

Implantation

The *coch-mic* is designed to be integrated with or implanted alongside a traditional cochlear implant. This implantation process is straightforward—the sensor is inserted into the cochlea through the round window, and the round window is sealed to improve the pressure gain from the ear canal to the scala tympani. The *drum-mic* is designed to be similarly practical to implant. Originally designed to rest on the promontory right below the umbo, the *drum-mic* would directly pick up vibrations from the ear drum. However, the Jacobson's nerve runs over the surface of the promontory, and placement of the device in that location requires a rigid platform, which would be anchored to the bone. This rigid platform must be thin enough to fit through the facial recess and practically adjustable by the surgeon. Once the mastoid bone is drilled out as if to install a cochlear implant, a rigid platform is screwed into the ear canal wall and the sensor is attached to this platform. The design of such a platform is in development, and a stiff, adjustable platform is feasible. The *cantilever-mic* also requires anchoring, and implantation is similar to that of the *drum-mic*. All of these implantation methods have downsides. Care must be taken to avoid kinking the *coch-mic* or scratching its surface, which may damage the waterproof coating or the electrodes underneath.

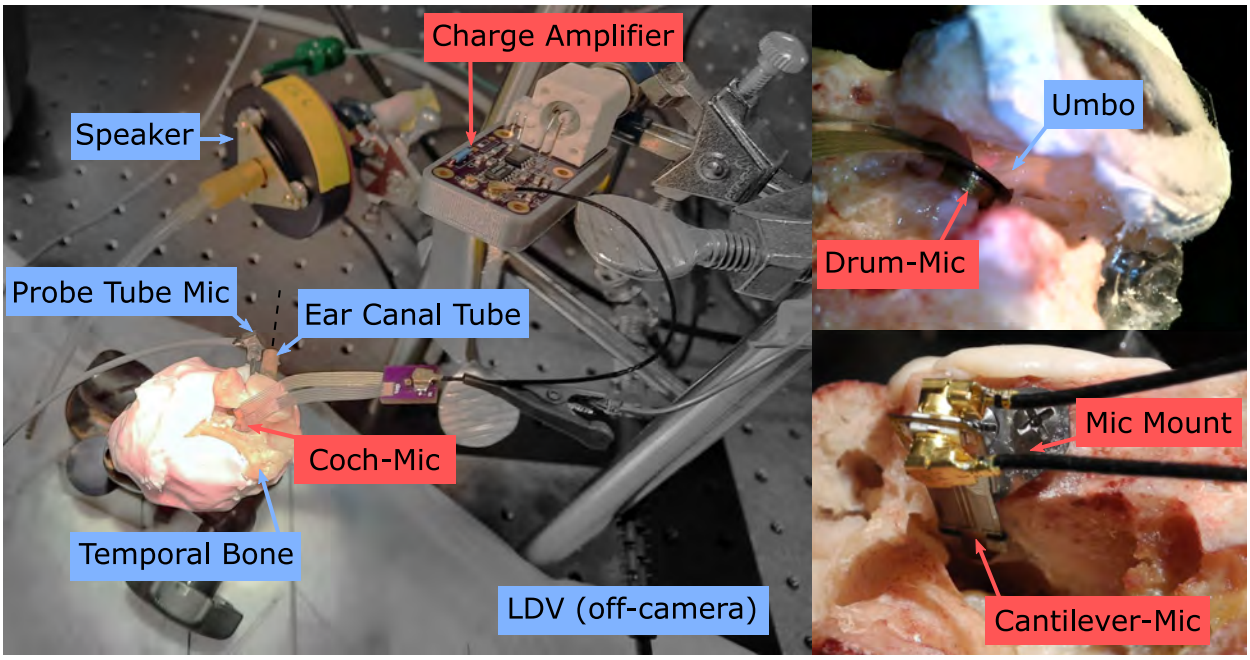


FIGURE 2. Experimental setup for our microphones in human cadaveric specimens. Our sensors, charge amplifier, and custom mounting hardware are highlighted in red. *Left:* Experiments involve an external acoustic stimulus provided by a speaker, a probe tube microphone to measure ear canal pressure, a Laser Doppler Velocimetry measurement of either the umbo or stapes motion, and the sensor-under-test (here a *coch-mic* inserted through the round window). Note that the speaker is disconnected from the ear canal tube in this photo. *Top-Right:* A *drum-mic* is inserted into the middle ear under contact with the umbo. *Bottom-Right:* A *cantilever-mic* is mounted using prototype surgical hardware screwed to the ear canal wall.

Mechanical Impedance Matching

One of the main advantages of PVDF as compared to typical ceramic piezoelectrics is its lower Young's modulus. This allows for the practical fabrication of durable microphones with mechanical impedance comparable to the middle ear structures. In particular, the cantilever's mechanical impedance is comparable with the impedance of the umbo. This allows for significant conversion from acoustic energy in the ear canal to elastic strain energy in the PVDF, resulting in high sensitivity. This mechanical impedance is relatively linear and controllable, reducing the effect of variability in microphone placement. The *drum-mic* also has reasonably good mechanical impedance matching. However its mechanical impedance depends heavily on static offset, and so tuning the impedance is more difficult. Even so, it produces a high signal output.

Characterizing the *coch-mic*'s mechanical impedance is somewhat more difficult. Based on the bulk modulus of PVDF and the input impedance of the cochlea (Frear 2018), the stiffness of the *coch-mic* is several orders of magnitude larger than that of the scala tympani—that is, when the ear is stimulated, the compression of the PVDF will be much smaller than the motion of the cochlear fluids [13]. It is also possible that motion of the intracochlear fluid results in bending motion of the microphone. Such coupling through bending motion is potentially significant, but it is very difficult to simulate or characterize. In any case, the current *coch-mic* design does not respond to bending motion, and such a coupling mechanism can be safely ignored.

Shielding and Waterproofing

The charge amplifier we designed has a measured input-referred noise floor of approximately 385 electrons over its bandwidth of 200 Hz to 20 kHz when measured with a *cantilever-mic*. Such a low noise floor means that great care must be taken to protect the microphone's signal path from the outside environment. The effect of non-ideal environmental conditions on sensor performance is dramatically illustrated by the noise curves in Figure 3. Although

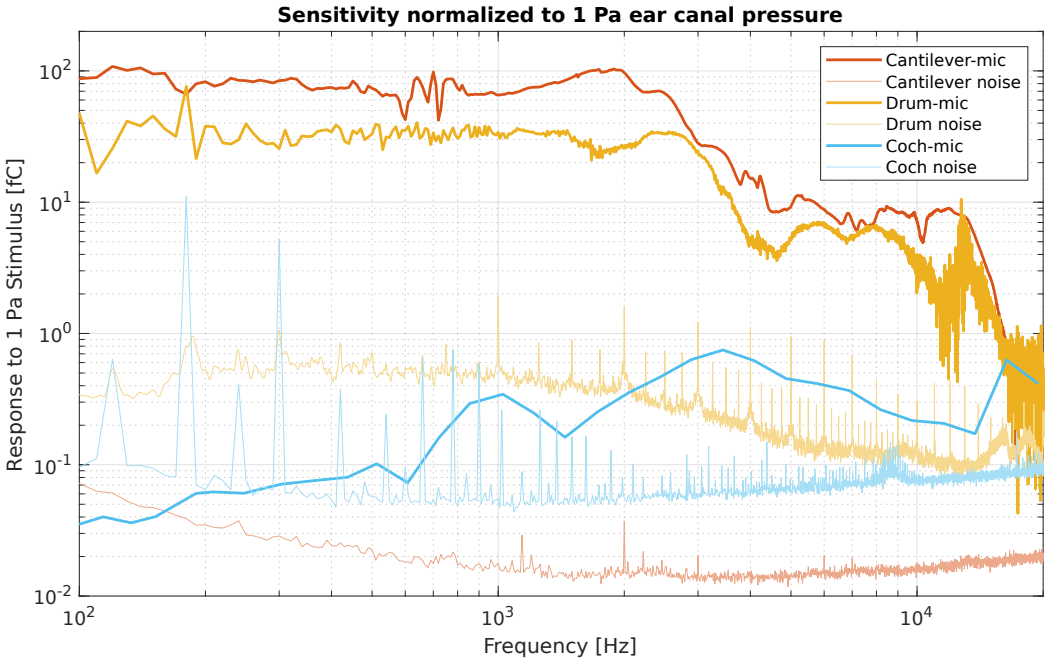


FIGURE 3. Typical measured frequency response curves. The *cantilever-mic* does not exhibit 60 Hz or 1 kHz harmonics since it is shielded from EMI (light black curve). Note that these responses do not include the influence of the outer ear since no pinna was present during the human cadaveric experiment. All responses are referenced to the ear canal pressure.

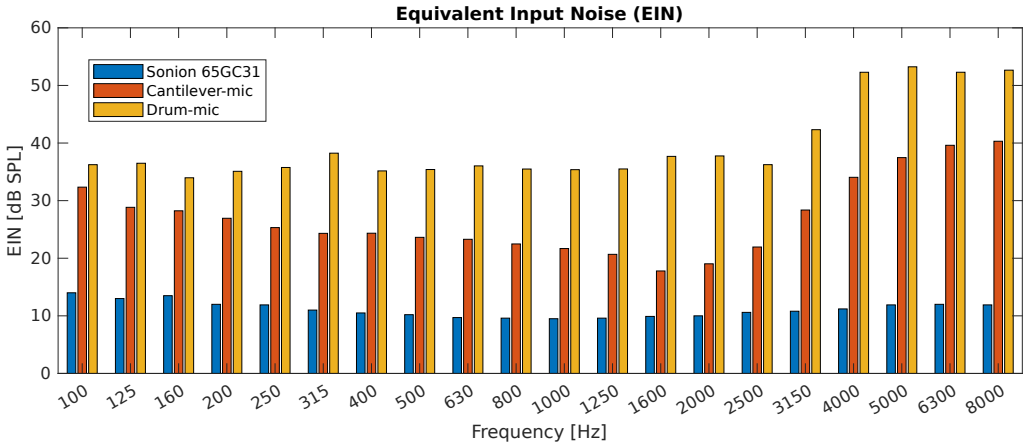


FIGURE 4. Typical 1/3 octave equivalent input noise (EIN). EIN was computed by finding the noise floor in volts in 1/3 octave bandwidths and dividing these values by the sensitivity of the device in volts/Pa, where the pressure was measured in the ear canal. The result is expressed in dB SPL. The EIN of the *cantilever-mic* was calculated using the data from Figure 3; the EIN of *drum-mic* was calculated using bench-tested noise data. The *coch-mic* has a significantly higher equivalent input noise due to its lower sensitivity and has been omitted from this plot for readability.

these noise curves were taken with the same amplifier board, the noise performance differs by orders of magnitude.

The most obvious discrepancy is the presence of large spikes corresponding to multiples of 60 Hz present in the drum and *coch-mic* noise floor, but absent in the cantilever microphone, even though these experiments were conducted in an electrically shielded room with the lights out. The *drum-mic* also picks up harmonics of 1 kHz, possibly from a piece of lab equipment. The cantilever design is fully electrically shielded and differential–mode, leading to at least 50 dB of EMI reduction in comparison to a single–ended unshielded device.

In addition to electromagnetic interference, the parasitic capacitance and leakage conductance of the sensors leads to increased noise levels. In particular, the *coch-mic* has a thin waterproofing coating, which induces a significant parasitic capacitance when immersed with the relatively conductive intracochlear fluid acting as a ground electrode. The *drum-mic* has no such coating and thus has a significant parallel leakage conductance, contributing to a higher noise floor. Coating the drum in an insulating material like silicone or parylene could help mitigate this problem. Older designs of the cantilever microphone had significant leakage conductance through the edge of the sensor especially in humid conditions—this was solved by completely sealing the charge sense electrodes inside the device.

Future Directions

Although these piezoelectric sensors are promising prototypes, a large amount of work is needed to make them practically implantable in humans. The microphone designs and amplifier must be completely waterproofed, repackaged in biocompatible materials, and integrated with a cochlear implant processor or other active auditory prosthetic. Although PVDF is biocompatible, the coatings and adhesives used are not, and would have to be replaced or hermetically sealed away from living tissue. The charge amplifier makes an excellent tool for lab bench testing, but is too bulky and power-hungry to be implanted. One possibility is to integrate a low-noise preamplifier with the *cantilever-mic*'s flex PCB, similar to the integral JFET preamplifier used in many commercial electret microphones [14]. The *cantilever-mic* also requires an adjustable rigid platform for implantation, which is in the early stages of development. The *coch-mic* will need to be integrated with the existing electrode arrays of cochlear implants. One possible design will be a cylindrical wrapper of PVDF around the electrode array. One additional challenge will be implementing shielding or active cancellation of the signal from the electrode array for the *coch-mic* during operation.

Our goal is to bring fully implantable microphones into the realm of practicality. Such microphones would improve the experience of cochlear implant users and open the door for many other novel implantable hearing assistive devices currently impractical from lack of good implantable microphones.

ACKNOWLEDGMENTS

This work was supported by the NIH through grant R01DC016874, the NSF through GRFP Grant Number 1745302, the Foundation of Prof. Dr. Max Cloëtta, the Research Fund of the University of Basel, and the NIH/NIDCD T32DC000020-29A1.

REFERENCES

1. D. Calero, S. Paul, A. Gesing, F. Alves, and J. A. Cordioli, "A technical review and evaluation of implantable sensors for hearing devices," Biomedical engineering online **17**, 1–26 (2018), iSBN: 1475-925X Publisher: Springer.
2. A. Meitzler, H. Tiersten, A. Warner, D. Berlincourt, G. Couqin, and F. Welsh III, "Ieee standard on piezoelectricity," (1988).
3. H. F. Tiersten, *Linear Piezoelectric Plate Vibrations: Elements of the Linear Theory of Piezoelectricity and the Vibrations Piezoelectric Plates* (Springer, 2013).
4. L. Li, M. Zhang, M. Rong, and W. Ruan, "Studies on the transformation process of pvdf from α to β phase by stretching," Rsc Advances **4**, 3938–3943 (2014).
5. Y. Roh, V. V. Varadan, and V. K. Varadan, "Characterization of all the elastic, dielectric, and piezoelectric constants of uniaxially oriented poled pvdf films," IEEE transactions on ultrasonics, ferroelectrics, and frequency control **49**, 836–847 (2002).
6. H. Wang, Q. Zhang, L. E. Cross, and A. O. Sykes, "Piezoelectric, dielectric, and elastic properties of poly (vinylidene fluoride/trifluoroethylene)," Journal of applied physics **74**, 3394–3398 (1993).
7. J. Z. Zhang, B. G. Cary, A. Yeiser, C. I. McHugh, H. H. Nakajima, J. H. Lang, I. Kymissis, and E. S. Olson, "A pvdf-trfe intracochlear hydrophone and amplifier for totally implantable cochlear implants," in *2022 IEEE 35th International Conference on Micro Electro Mechanical Systems Conference (MEMS)* (IEEE, 2022) pp. 408–411.
8. J. Sirohi and I. Chopra, "Fundamental understanding of piezoelectric strain sensors," Journal of intelligent material systems and structures **11**, 246–257 (2000).
9. B. G. Cary, J. Z. Zhang, C. I. McHugh, I. Kymissis, E. S. Olson, H. H. Nakajima, and J. H. Lang, "An implantable umbo microphone for fully-implantable assistive hearing devices," in *2021 IEEE Sensors* (IEEE, 2021) pp. 1–4.
10. S. Park, X. Guan, Y. Kim, F. P. X. Creighton, E. Wei, I. Kymissis, H. H. Nakajima, and E. S. Olson, "Pvdf-based piezoelectric microphone for sound detection inside the cochlea: toward totally implantable cochlear implants," Trends in hearing **22**, 2331216518774450 (2018).
11. J. Z. Zhang, *An Intracochlear Hydrophone and Amplifier*, Master's thesis, Massachusetts Institute of Technology (2021).

12. B. G. Cary, *Design of a surgically-viable umbo microphone for implantable assistive hearing devices*, Master's thesis, Massachusetts Institute of Technology (2021).
13. D. L. Frear, X. Guan, C. Stieger, J. J. Rosowski, and H. H. Nakajima, "Impedances of the inner and middle ear estimated from intracochlear sound pressures in normal human temporal bones," Hearing Research **367**, 17–31 (2018).
14. J. W. Dally, W. F. Riley, and K. G. McConnell, *Instrumentation for Engineering Measurements* (Wiley India Pvt. Limited, 2006).

Design and Testing of Ultraminiature MEMS Middle Ear Accelerometers

Panagiota Kitsopoulos,[1, a)] Alison E. Hake,[1, b)] Emily Z. Stucken,[2, c)] Christopher M. Welch,[2, d)] and Karl Grosh[3, e)]

[1)]*Department of Mechanical Engineering, University of Michigan, G.G. Brown Building, 2350 Hayward St., Ann Arbor, MI 48109-2125, U.S.A.*
[2)]*Department of Otolaryngology-Head and Neck Surgery, University of Michigan, Taubman Center, 1500 E Medical Center Dr, Floor 1 Reception A, Ann Arbor, MI 48109-5312, USA.*
[3)]*Department of Mechanical Engineering, University of Michigan, G.G. Brown Building, 2350 Hayward St., Ann Arbor, MI 48109-2125, U.S.A.*

[a)]*Corresponding author: pkitsop@umich.edu*
[b)]*Electronic mail: aehake@umich.edu*
[c)]*Electronic mail: estucken@med.umich.edu*
[d)]*Electronic mail: welchch@med.umich.edu*
[e)]*Electronic mail: grosh@umich.edu*

Abstract. The goal of this work is to develop a completely implantable ossicular vibration sensor utilizing microelectromechanical systems (MEMS) technology in combination with integrated-circuit approaches as part of a larger effort to enhance auditory prostheses by eliminating their external components. To develop a completely implantable device, sensors are required to replace the external microphones used in traditional systems (e.g., cochlear implants and hearing aids). Present-day implantable sensors do not meet the stringent requirements for acoustic performance or size to fit in the middle ear; however, in our lab, a dual-resonance design has been developed that holds the potential to achieve these results. In this paper, we show the analytical model used to predict and understand the behavior of these sensors. We discuss preliminary cadaveric temporal bone results, show how a fabricated MEMS proof-of-concept prototype mounted on a printed circuit board (PCB) along with the amplifying electronics and covered with a high-resolution 3D printed lid results in a packaged size small enough to fit in the middle ear, and outline future testing protocols.

INTRODUCTION

Hearing loss is a debilitating condition that affects over 5% of the world's population and can negatively impact an individual's quality of life by hindering their ability to communicate, socialize, receive education, and work [1]. In the United States, approximately 37.5 million people report some hearing disability, making it one of the most prevalent conditions in the country [2]. Depending on the severity of the hearing loss and the location of the damage within the ear, devices such as hearing aids (HAs) and cochlear implants (CIs) are beneficial by amplifying speech and environmental sounds to audible levels or stimulating hearing electrically [2]. Although HAs can significantly improve the lives of those suffering from hearing loss, their adoption rate is less than 14.2%. Surveys and analyses have indicated several key issues associated with the low HA adoption rates. These include cost (Medicare does not cover any HAs), lack of perceived effectiveness, and a multitude of issues associated with the external processing components. HA and CI users express the inconvenience of the externally worn, detachable parts that make them self-conscious of their disability and can easily be lost, damaged or stolen [2, 3, 4]. Additionally, these external elements cannot be worn underwater, during intense physical activities, or sleep. Thus, the long-term goal of this work is to utilize MEMS technology to build ultraminiature accelerometers that can be packaged to fit within the middle ear and effectively replace and eliminate those external components.

In typical externally worn auditory prostheses, sounds are detected by microphones and analyzed by signal processors. To create a completely implantable auditory prosthesis (CIAP), implantable sensors must be developed that meet certain operating requirements (e.g., size, bandwidth) -specifically match or even exceed the performance of current microphone capabilities between the frequency range of 100 Hz and 8 kHz. No present-day implantable sensors meet these specifications [4, 5]. Our approach includes the use of piezoelectric accelerometers as acoustic sensors whose performance holds the potential to address this issue by utilizing the dual resonance design, inspired by our first generation of MEMS prototypes [6]. Our objective is to fabricate and test these dual resonance devices. As preparations for manufacture are underway, we are creating new testing protocols using newly miniaturized printed circuit boards (PCBs) with amplification circuitry, 3D printed cover lids, and the first generation single resonance

sensors. Preliminary experiments with cadaveric temporal bones provided additional insight on the maximum size of the PCB-sensor-lid assembly in order to be attached on the umbo and have space to freely vibrate. Below we will briefly describe the theory behind the single and dual resonance model, provide some theoretical and experimental results, and outline our testing plans for the near future.

DESIGN OF A PIEZOELECTRIC MEMS ACCELEROMETER

Analytical Model

The accelerometer is modeled as a piezoelectric bimorph cantilever beam that produces a measurable voltage output when deflected due to sound-induced ossicular vibrations. In Fig. 1, we display the diagram of the sensor that sits on top of a PCB which includes the necessary electronics to extract and amplify the voltage output. The modeled beam consists of two aluminum nitride (AIN) layers and three electrode layers (on the top, in the middle, and on the bottom). A proof mass is attached flush against the end of the bimorph. The structure is subjected to either transverse (z-axis, W_{bz}), longitudinal (x-axis, U_{bx}), or rotational (about y-axis, θ_{by}) base excitation and the two translation acceleration sensitivities and the angular acceleration sensitivity are calculated for each respective case.

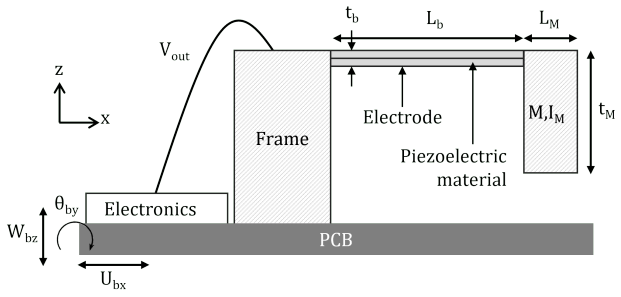


FIGURE 1. The accelerometer is comprised of a piezoelectric cantilever bimorph beam. The design consists of two aluminum nitride (AIN) layers of length L_b, thickness t_b, and width Y_b. Three electrodes are on the top, middle, and bottom of the bimorph. The proof mass sits flush against the beam and has length L_M, thickness t_M, width W_M, mass M, and moment of inertia I_M. The entire device sits on a PCB that includes the signal processing electronics for the voltage output V_{out} from the piezoelectric material. W_{bz} is defined as the prescribed base displacement in the z-direction, U_{bx} is defined as the prescribed based displacement in the x-direction, and θ_{by} is defined as the prescribed base rotation about the y-axis. Not to scale.

To model the response of the vibrating beam, the Euler-Bernoulli beam hypothesis is assumed, resulting in the following governing equations for transverse and longitudinal beam deflection:

$$\overline{EI}\frac{d^4W(x)}{dx^4} - \omega^2\rho AW(x) = 0 \tag{1}$$

$$EA\frac{d^2U(x)}{dx^2} + \omega^2\rho AU(x) = 0. \tag{2}$$

In these equations, $W(x)$ and $U(x)$ are the displacements in the z- and x-direction, $\overline{EI}$ is the beam bending stiffness that is determined via the derivations described in [7, 8], ω is the frequency, A is the cross-section area of the beam, and ρ and E are the density Young's Modulus of the AIN material respectively. Additionally, we assume a $e^{j\omega t}$ time dependence.

The boundary conditions used to solve the unknown coefficients in Eq. 1 and Eq. 2 are the following:

$$W(0) = W_{bz}$$
$$\frac{dW(0)}{dx} = \theta_{by}$$
$$\overline{EI}\frac{d^2W(L_b)}{dx^2} = \frac{1}{2}ML_M\omega^2 W(L_b) + I_M\omega^2\frac{dW(L_b)}{dx} + M\omega^2\left(\frac{t_M}{2} - t_{piezo}\right)U(L_b) - \frac{Y_b d_{31} t_{piezo}}{s_{11}}V_{out}$$
$$\overline{EI}\frac{d^3W(L_b)}{dx^3} = -M\omega^2 W(L_b) - \frac{1}{2}ML_M\omega^2\frac{dW(L_b)}{dx}$$
$$U(0) = U_{bx}$$
$$\frac{A}{s_{11}}\frac{dU(L_b)}{dx} = M\omega^2 U(L_b) + M\omega^2\left(\frac{t_M}{2} - t_{piezo}\right)\frac{dW(L_b)}{dx}$$
$$Q_p = C_p V_{out} + \frac{d_{31}Y_b t_{piezo}}{s_{11}}\frac{dW(L)}{dx}$$

The solutions to Eq. 1 and Eq. 2 are:

$$W(x) = C_1\sin(\gamma x) + C_2\cos(\gamma x) + C_3\sinh(\gamma x) + C_4\cosh(\gamma x)$$
$$U(x) = C_5\sin(\beta x) + C_6\cos(\beta x).$$

Here Q_p is the output charge from the layered beam for a parallel electrical connection. $\gamma = (\frac{\rho A\omega^2}{\overline{EI}})^{\frac{1}{4}}$, $\beta = \omega\sqrt{\rho s_{11}}$, and $C_p = \frac{2\hat{\varepsilon_{33}}Y_b L_b}{t_{piezo}}$ is the capacitance, where $\hat{\varepsilon_{33}} = \varepsilon_{33} - \frac{{d_{31}}^2}{s_{11}}$ and ε_{33} is the permittivity, s_{11} is the elastic compliance, V_{out} is the voltage output, Y_b is the width of the beam, d_{31} is the piezoelectric coefficient calculated experimentally for low frequencies as described in [8], and $t_{piezo} = 0.5t_b$ as there are two layers of AlN. The location of the center of gravity of the proof mass is below the neutral axis of the beam which results in the transverse and longitudinal coupling in the boundary conditions. The charge equation does not include a longitudinal extension component because the voltage from each extending layer is canceled in a parallel electrical connection. Sample results for the transverse (z-axis) and longitudinal (x-axis) sensitivities are plotted in Fig. 2. The sensitivity is the voltage output in each case normalized to the excitation acceleration. The first resonance occurs at 1140Hz which limits the operational bandwidth of the sensor below 1kHz. The transverse and longitudinal sensitivities are out of phase which is important to keep in mind when considering the orientation and position of the sensor as sensor output could either increase or decrease depending on the phase of W_{bz}, U_{bx}, and θ_{by}. The dimensions and material property values used for these results are included in Table I.

TABLE I. Dimensions and material properties used in Fig. 2.

L_b	125 μm
Y_b	300 μm
t_b	1 μm
L_M	75 μm
W_M	300 μm
t_M	400 μm
M	2.1x10^{-12} kg
I_M	1.1x10^{-15} m^4
s_{11}	2.9x10^{-12} m^2/N
ε_{33}	$10.12 * 8.854$x10^{-12} F/m
A	3.0x10^{-10} m^2
d_{31}	-2.2x10^{-12} C/N

Previous work [9] has demonstrated and verified the correlation between the analytical model acceleration sensitivity in the z-direction and experimental data. Future work will include the testing and verification of the sensitivities in the x- and y-axes due to base motion and rotation about each respective direction.

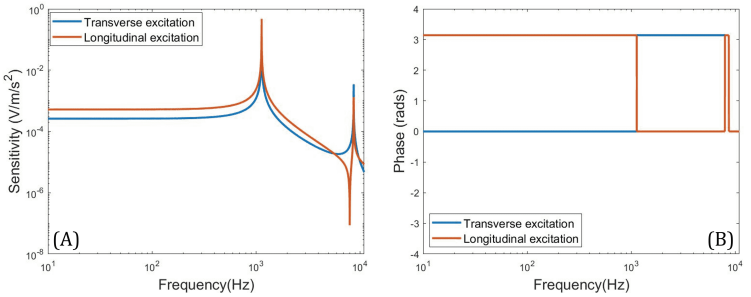


FIGURE 2. (A) Modeled acceleration sensitivities vs. frequency due to base motion in the z- (transverse) and x- (longitudinal) directions. (B) Corresponding phase plots for each case.

A Single Bandwidth is Not Enough!

In previous work, we computed the minimum detectable sound pressure level (SPL) of the accelerometer based on the sensitivity, input referred noise, and experimental data from the literature relating input SPL to the middle ear acceleration; we reaffirmed that the a single resonance accelerometer (i.e. one beam or sensing element) cannot meet the required performance specifications for a CIAP [9]. Thus, we have proposed a new theoretical design combining two sensing elements creating a dual resonant accelerometer. As seen in the schematic top-view in Fig. 3, two piezoelectric bimorph cantilever beams will be arranged in the MEMS die. One sensing element will be designed to operate over low frequencies, and the other sensing element will be designed for high frequencies. Fig. 3 shows the theoretical transverse sensitivity of each sensing element using the analytical model discussed above. We predict that the dual resonant accelerometer, designed with increased low-frequency sensitivity, can simultaneously meet the noise floor, bandwidth, and size requirements for a CIAP. However, in addition to the necessary amplification, further filtering is needed to obtain the desired voltage output and manage the transition between elements.

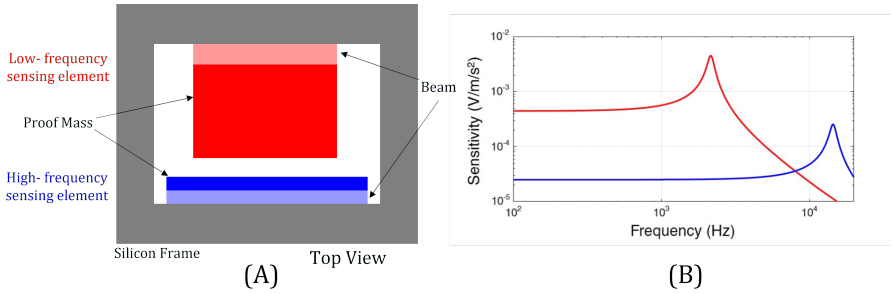


FIGURE 3. (A) The dual resonance accelerometer diagram. Each beam is made out of two aluminum nitride (AIN) layers, contains three electrodes (top, middle, and bottom), and has a proof mass flush against each beam tip. Not to scale. (B) Acceleration sensitivity responses for each of the two piezoelectric cantilever bimorph beams. Filtering will be used to manage the transition between sensing elements to span the entire frequency range.

TESTING, PACKAGING, AND ASSEMBLY

Preliminary Cadaveric Temporal Bone Testing

In order to develop our experimental testing protocols on cadaveric settings, preliminary testing was undertaken. Temporal bone samples were obtained and initial surgery and preparations were completed at the University of Michigan Temporal Bone Dissection Lab in Michigan Medicine. A commercially available voice accelerometer (Vesper Technologies Inc., VA1200) was tested in both benchtop and preliminary cadaveric temporal bone experi-

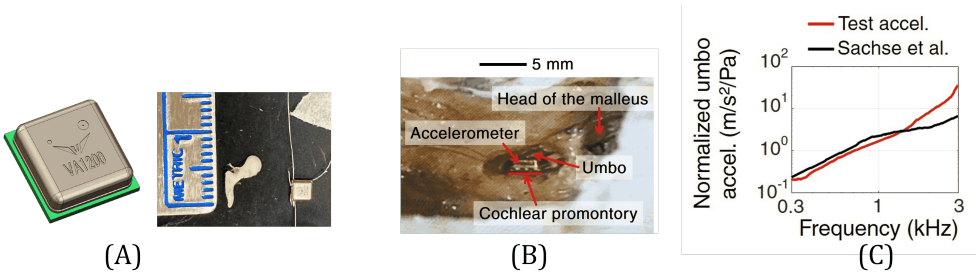


FIGURE 4. (A) Commercial accelerometer (Vesper, VA1200) compared next to a human malleus bone. (B) The accelerometer is shown attached on the umbo of the malleus accessed through a mastoidectomy with posterior tympanotomy and attached using super glue, and (C) preliminary umbo acceleration sensitivity results compared to literature data [11].

ments. These demonstrated the testing procedures to be followed with the current MEMS sensor and aided in the development of new testing protocols. The VA1200, as seen in Fig. 4A, has the extremely small dimensions of 2.9mm x 2.76mm x 0.9mm which was ideal as we attempted to fit the sensor in the middle ear space. Dr. Stucken and Dr. Welch accessed the middle ear cavity of a cadaveric temporal bone through a mastoidectomy with posterior tympanotomy. As a preliminary method, super glue was used to attach the accelerometer to the umbo of the malleus (Fig. 4B). Preliminary testing demonstrated the limited space available within the middle ear cavity which significantly informed the choice of dimensions for our own PCB-sensor assembly design for future cadaveric experiments. Furthermore, the procedure and acoustic set-up used were non-ideal as we lacked a good access point for obtaining umbo acceleration data via the laser doppler vibrometer (LDV). The experimental setup for future cadaveric testing will be modified and improved following [10]. Fig. 4C displays the preliminary umbo acceleration sensitivity results measured as acceleration normalized by pressure through a probe microphone at the tympanic membrane. At frequencies below approximately 1.5 kHz, the results match well with literature data [11].

Sensor Packaging and Circuitry

We fabricated a package (PCB and 3D printed cover lid) that will house and protect the single resonance accelerometer and amplifying circuitry (LMV1015-25) during testing. The PCB was designed with the dimensional restrictions of the sensor, LMV1015-25, and limited capabilities of the manufacturing company (PCBUniverse). The resultant dimensions were set to 3.56mm x 3.21mm x 1.0mm. During preliminary cadaveric testing, it was observed that a height between 0.9mm -1.1mm is small enough to be attached to the umbo and vibrate without hitting the promontory for many temporal bone preparations. Fig. 5A displays the LMV1015-25, and sensor mounted and wirebonded on the board. A 3D printed cover lid is used to encapsulate the PCB-sensor assembly as seen in Fig. 5B. The testing

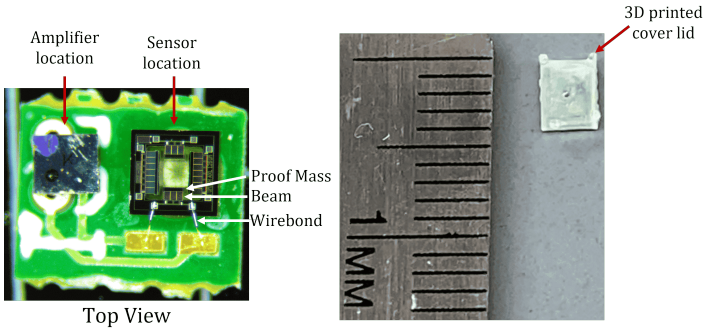


FIGURE 5. (A) Amplifier LMV1015-25 and single resonance sensor mounted on minimized printed circuit board (PCB). Sensor is wirebonded to the PCB. (B) PCB-sensor assembly covered by a 3D printed lid. The dimensions of the assembly are 3.56mm X 3.21mm X 1.0mm.

procedure for evaluating the sensor operation includes actuation and subsequent sensitivity experimentation as is described in [9]. Fig. 6 shows the sensitivity testing set-up that utilizes a shaker to provide controlled vibrations and the LDV to record and compare data to that received by the accelerometer. For the verification of the x- and z-axis sensitivities, separate shaker fixtures will be used to orient the sensor in the appropriate directions. It is necessary to understand the operation of the sensor in these orientations in order to optimize the choice of location for attachment in the cadaveric temporal bone and to especially ensure sensitivities are maximized considering the mode of vibration on the middle ear bone.

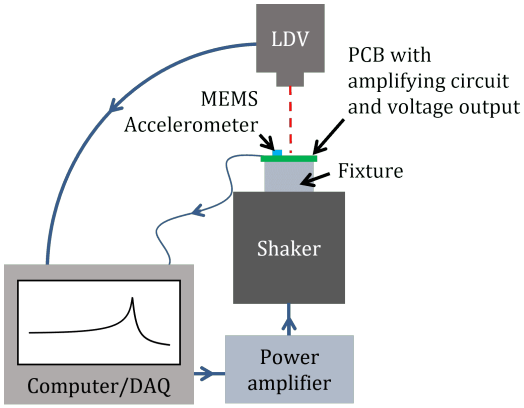


FIGURE 6. Sensing testing set-up. The procedure includes the use of a shaker that provides controlled vibrations to the PCB-sensor assembly. A laser doppler vibrometer (LDV) focused on the frame of the die is used to record the acceleration of the assembly and compare it to the output received by the sensor assembly itself with the DAQ collecting system. The sensitivities from excitations in different directions can be determined by this test.

Once all benchtop testing is completed and the functionality of the sensor is confirmed, we will advance to cadaveric testing by covering the PCB-sensor assembly with the 3D-printed lid - sealing and securing all mounted components. The modified cadaveric set-up will utilize a pressure sensor (probe mic, ER-7c) and configuration that will allow for a better quantification of sensitivity results, optical access for LDV measurements, and further special bone preparation [10]. Based on the modeled and experimental results, the accelerometer will be secured at the most advantageous location possible.

DISCUSSION AND CONCLUSIONS

In this paper, we briefly described the analytical model for the single resonance accelerometer and provided the modeled sensitivities due to certain base excitations. Although the sensitivity response due to the base displacement in the z-direction has been verified experimentally [9], the x- and about y-directions have yet to be completed. Accomplishing this task will allow for better understanding of the MEMS die behavior and for optimization on the attachment location within the middle ear for best sensitivity results. As the benchtop and cadaveric testing for the single resonant sensor are underway with the new miniaturized PCB-lid assembly, steps are taken to fabricate the dual-resonant devices that can potentially achieve the required specifications for a CIAP. Current experimentation allows for the iteration and improvement of the housing and mounting units that will ultimately be used for these dual resonant devices.

ACKNOWLEDGEMENTS

We would like to thank Dr. Sajad Arabnejad from the Scaffold Tissue Engineering Group at University of Michigan and Brendan Casey, from ProtoConnect LLC for their help. We would also like to acknowledge the NIH training grant (T32 DC00011) that partially funded this research.

REFERENCES

1. World Health Organiation, "Deafness and hearing loss," (2021).
2. The National Institute on Deafness and Other Communication Disorders (NIDCD), "2017-2021 NIDCD Strategic Plan," (2017).
3. L. L. Cunningham and D. L. Tucci, "Hearing loss in adults," New England Journal of Medicine **377**, 2465–2473 (2017), pMID: 29262274, https://doi.org/10.1056/NEJMra1616601.
4. D. Calero, S. Paul, A. Gesing, F. Alves, and J. A. Cordioli, "A technical review and evaluation of implantable sensors for hearing devices," BioMedical Engineering OnLine **17**, 23 (2018).
5. F. Creighton, X. Guan, S. Park, I. Kymissis, H. Nakajima, and E. Olson, "An intracochlear pressure sensor as a microphone for a fully implantable cochlear implant," Otology & neurotology: official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology **37** (2016), 10.1097/MAO.0000000000001209.
6. A. E. Hake, Piezoelectric MEMS Accelerometers for Sensing Ossicular Vibration, Ph.D. thesis, University of Michigan (2021).
7. M. Krommer, "On the correction of the bernoulli-euler beam theory for smart piezoelectric beams," Smart Materials and Structures **10**, 668–680 (2001).
8. R. Littrell and K. Grosh, "Modeling and characterization of cantilever-based mems piezoelectric sensors and actuators," Journal of Microelectromechanical Systems **21**, 406–413 (2012).
9. A. E. Hake, C. Zhao, W.-K. Sung, and K. Grosh, "Design and experimental assessment of low-noise piezoelectric microelectromechanical systems vibration sensors," IEEE Sensors Journal **21**, 17703–17711 (2021).
10. H. H. Nakajima, W. Dong, E. S. Olson, S. N. Merchant, M. E. Ravicz, and J. J. Rosowski, "Differential intracochlear sound pressure measurements in normal human temporal bones," Journal of the Association for Research in Otolaryngology **10**, 23–36 (2008).
11. M. Sachse, W. Hortschitz, M. Stifter, H. Steiner, and T. Sauter, "Design of an implantable seismic sensor placed on the ossicular chain," Medical Engineering & Physics **35**, 1399–1405 (2013).

Bio-inspired, Adaptive Acoustic Sensor: Sensing Properties in Dependence of Feedback Parameters

Kalpan Ved,[1] Claudia Lenk,[1, a)] Tzvetan Ivanov,[1] Philipp Hoevel,[2] and Martin Ziegler[1, 3]

[1)]*Dept. Micro- and Nanoelectronic Systems, Technische Universität Ilmenau, 98693 Ilmenau, Germany*
[2)]*School of Mathematical Sciences, University College Cork, Western Road, Cork T12 XF64, Ireland*
[3)]*Institute of Micro- and Nanotechnologies (IMN) MacroNano, Technische Universität Ilmenau, 98693 Ilmenau, Germany*

[a)]*Corresponding author: claudia.lenk@tu-ilmenau.de*

Abstract. Pre-processing of the sound signal during sensing is an integral functionality of the cochlea, the part of human hearing responsible for sound sensing. This pre-pocessing, which is integrated into the sensing stage directly, enables the remarkable properties of human hearing. Similarly, integrating some of these pre-processing functionalities in technological speech processing systems strongly improves their recognition performance.
We developed a bio-inspired, adaptive acoustic sensor with pre-processing capabilities like nonlinear amplification and frequency filtering functionality. The sensor is composed of a single clamped silicon beam with integrated deflection sensing and thermo-mechanical actuation, subjected to a real-time feedback. While the resonance frequency and bandwidth are determined by the geometry of the sensor beam, its transfer characteristics can be switched dynamically from linear to nonlinear regime by changing the feedback parameters. In the linear regime, the feedback controls the sensitivity and bandwidth of the sensors. Here, we elaborate on the influence of the sign of feedback strength and offset on the sensor behaviour. Changing the sign of the feedback parameters switches between amplification and damping behaviour, enabling the change of sensitivity by 44 dB. Thereby, complex oscillation modes are observed for feedback parameters with similar polarity.

INTRODUCTION

The process of human hearing is full of intriguing properties like a dynamic range of 120 dB sound pressure level (SPL), resolving tones only 0.2% apart in frequency and understanding speech in noisy environments [1, 2]. These properties originate partly from the sound capturing and pre-processing capabilities of the cochlea [3]. The cochlea is the organ for sound sensing, and it incorporates a number of signal conditioning and feature extraction steps in the sensing process as well. Thereby, the most important steps are a frequency decomposition and amplification of the input signal, compressive (nonlinear) transfer characteristics and adaptation of amplification and transfer characteristics properties based on local and efferent feedback. These steps take place before and during the transduction process. The nonlinear amplification characteristics are discussed to stem from the sensor system acting as a Hopf oscillator [4, 5]. This hypothesis is further underlined by the fact, that the ear itself can produce tones without sound input, the so-called oto-acoustic emissions [6].

Introducing some of these steps into a technological speech processing system strongly improves the performance of the system, as was recently shown by Araujo et al. [7]. Thereby in particular, the nonlinear transfer characteristics improve the clustering of data, as was tested for spoken digits, which simplifies the classification task for the subsequent neural network. Most speech processing systems nowadays include, after the actual sensing step, a signal conditioning and feature extraction step, typically calculating the envelope and the frequency of the signal and transforming it into a frequency-time representation, often referred to as spectrogram [8]. Thereby, software-based realisations are quite common, but hardware-based implementations, often referred to silicon or FPGA cochlea, exist as well (cf. Refs. [9, 10, 11, 12]). Despite the remarkable increase in performance of speech processing systems, these still struggle in particular with noisy conditions, i.e., low signal-to-noise ratios, or multi-source situations. To address this issue, a bio-inspired approach targets the integration of the signal conditioning and feature extraction step into the sensing process, similar to the operation of the cochlea, rather than after the sensing process, as it is done in current technological realisations.

Several bio-inspired acoustic sensors have been developed, which differ mainly in the integrated functionalities and the number of sensors necessary to cover the frequency range [13]. The most common approach is the integration of frequency decomposition into the sensor system. However, these are the level used of bio-inspired sensors, which has been tested and applied already in speech processing systems [14]. Other approaches of bio-inspired sensors include nonlinear transformation and/or adaptation.

We developed an acoustic sensor which integrates nonlinear transfer characteristics, amplification and frequency filtering functionality into the sensor and incorporates dynamic adaptation [15, 16]. These functionalities are achieved by combining a passive transducer with a high-speed (real-time) feedback. Here, we present our study on the sensing properties in dependence of the sign of the feedback parameters after a short introduction of the sensor system.

BIO-INSPIRED, ACOUSTIC SENSOR SYSTEM

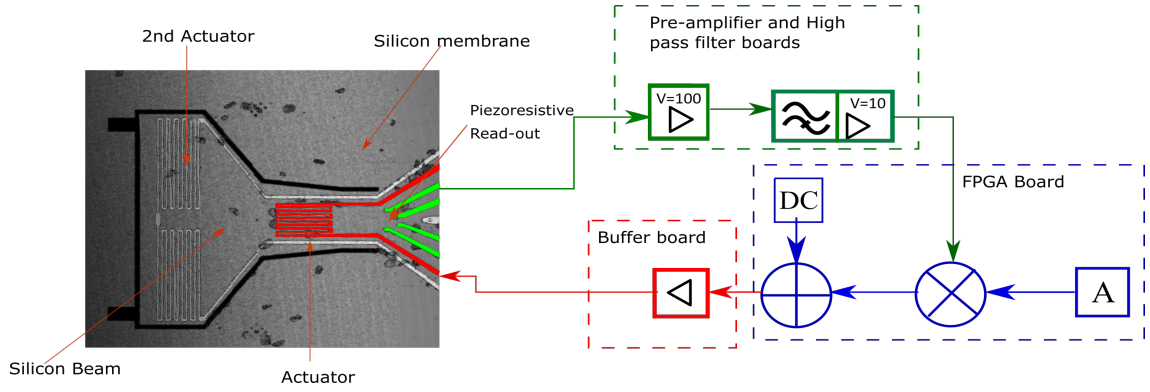


FIGURE 1: Schematic representation of the sensor system with microscope image (left) of silicon-based transducer with integrated thermo-mechanical actuator (red) and piezo-resistive deflection sensor (green) and the feedback algorithm (right) used to tune sensing properties of transducer.

To obtain the above described sensing functionalities, the developed bio-inspired sensor system consists of a transducer in combination with a high-speed feedback, as schematically shown in fig 1. Thereby, a single-side mounted silicon beam with integrated piezo-resistive deflection sensing is used as transducer. The dimensions of the silicon beam are 775 µm−1354 µm (length), 775 µm (width) and $2-10$ µm thickness. The transducer is surrounded by a silicon membrane to decrease the acoustic short circuit effect and thus increase the sensitivity. The beam is operated in resonance mode with a quality factor of $100-200$ and a gain factor of ≈ 3 V/Pa upon excitation with sound [17]. Through the resonant operation of the sensor, it effectively acts as band-pass filter with a bandwidth of $20-100$ Hz, depending on the design. The frequency f of each sensor is determined by its geometrical dimensions:

$$f = \frac{\omega_0}{2\pi} = \delta_n^2 \frac{d_{\mathrm{Si}}}{2\pi l_{\mathrm{Si}}^2} \sqrt{\frac{E_{\mathrm{Si}}}{12\rho_{\mathrm{Si}}}}$$

with l_{Si} the length of the beam, d_{Si} its thickness, E_{Si} the elastic module of silicon, ρ_{Si} the density of Silicon, and δ_n a positive pre-factor, depending on the oscillation mode. Thus, the resonance frequency can be tuned by the length and the thickness. However, the majority is tuned by changing the length of the beam [17]. To cover the audible range and implement frequency decomposition, an array of sensors with different resonance frequencies is necessary. Despite the small bandwidth, a spectrogram-like graph can be obtained from a relatively small number of sensors as shown in figure 2. Therefore, the envelope of the sensor signals, obtained from sensors with different resonance frequencies, are plotted as time series. Each signal is drawn on the y-axis according to the resonance frequency of its beam. Thereby, the sensor signals were simulated using a previously described model of the sensor system [18], which calculates the beam deflection in dependence of the introduced heating of the beam by the applied feedback voltage (see below). The input for the sensor array was a recording of a spoken digit (here: 'zero') [19]. A spectrogram of the same input, obtained from short-time fast Fourier filtering of overlapping time intervals of 5 ms length, is plotted for comparison in figure 2. In both cases, the frequency scale was taken according to the mel-scale with 16 channels for the frequency range of 100 Hz to 4 kHz. As can be seen, the beam-based spectrogram resembles the Fourier-based spectrogram quite well, despite the fixed bandwidth of 50 Hz.

Furthermore, implementation of high-speed feedback, as schematically shown in figure 1, is enabled by the thermo-mechanical actuator integrated into the silicon beam as well as the piezo-resistive read-out for deflection sensing [20,

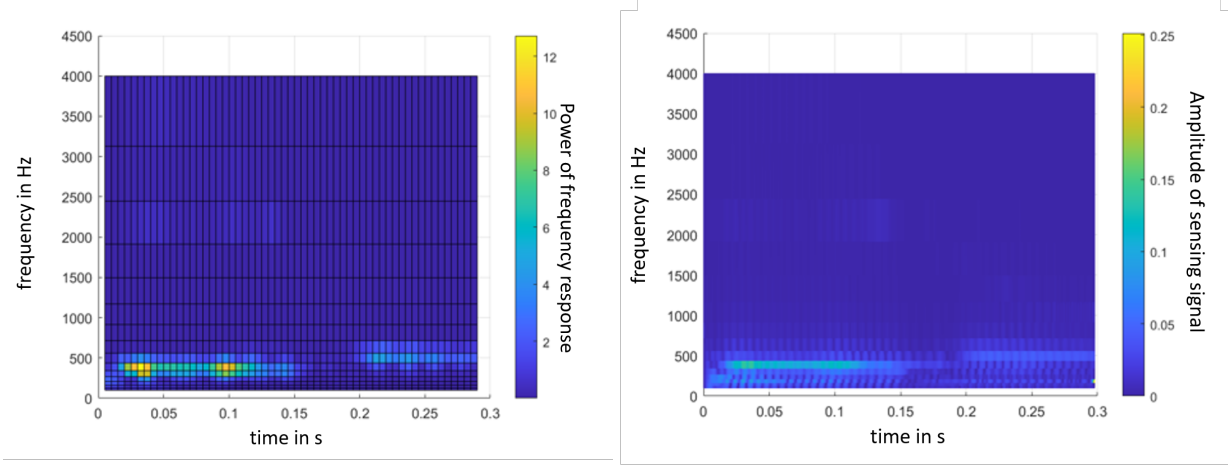


FIGURE 2: Spectrogram of spoken digit 'zero' obtained with short-time fast Fourier filtering (left graph) and from envelope of sensor signal from simulated beam array (right graph). Thereby, the color code refers to the power of the signals for the Fourier-based spectrogram (left) or the amplitude of the simulated sensing signals in the case of the beam array (right). In both cases, 16 channels distributed according the mel-scale were used for the frequency range of 100 Hz to 4 kHz.

21]. The feedback algorithm is given by

$$u_{\mathrm{act}} = a_{\mathrm{f}} u_{\mathrm{sens}} + u_{\mathrm{DC}} \tag{1}$$

and was realized in FPGA and analog circuits [16, 22]. Here, u_{act} refers to the actuation voltage applied to the actuator, u_{sens} is the sensing voltage obtained from the piezo-resistive read-out after amplification and high-pass filtering, a_{f} is the feedback strength and u_{DC} a bias offset. The feedback is used to tune the sensing properties of the system. In particular, one obtains three operation modes of the sensor depending on the feedback strength a_{f} or the bias voltage u_{DC}:

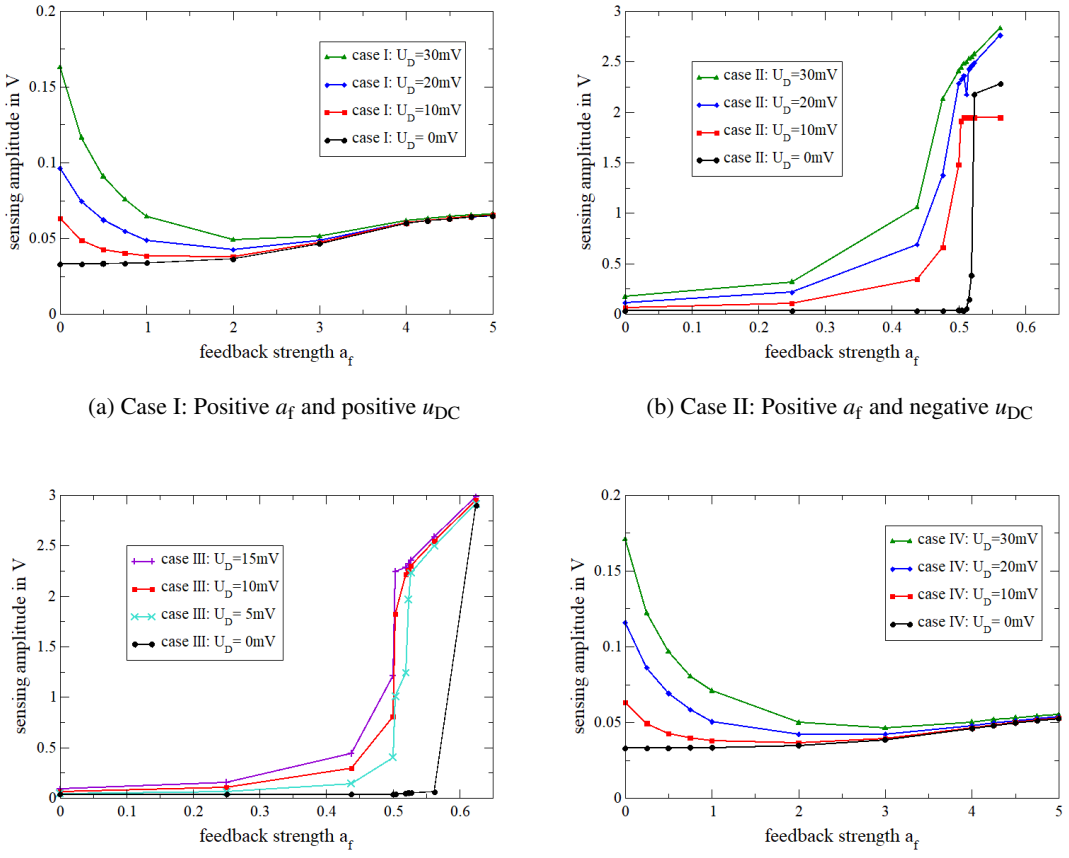
- (i) a passive mode with linear transfer characteristics, if both feedback parameters are zero,
- (ii) an active, linear mode for intermediate feedback strengths or DC values, which is characterised by an increased gain and sensitivity compared to the passive mode and linear transfer characteristics and
- (iii) an active, nonlinear mode for high feedback strengths or small DC values, which exhibits a gain depending on the sound pressure level (nonlinear transfer characteristics).

Further increase of the feedback strength or further decrease of the u_{DC} value will result in self-excited oscillations due to the system undergoing a Hopf bifurcation [18]. Interestingly, measurements with a reference microphone (MT Gefell MV203), which was placed 5.5 cm beside our developed bio-inspired microphone in an anechoic chamber, revealed that these self-excited oscillations produce measurable sounds. This is quite similar to the oto-acoustic emissions from the inner ear, which are thought to occur due to parts of the cochlea acting as Hopf oscillators as well [4]. An additional mechanism implemented in the sensor system is adaptation to the input or the acoustic environment [22]. Thereby, an amplitude-based switching of the feedback strength a_{f} was implemented to model sensor adaptation to constant sound inputs, to increase the dynamic range (for sensing with high gain values) and to enhance the sound onset for subsequent feature extraction.

Summarising, the developed, bio-inspired sensor system demonstrates functionalities observed from the inner ear in human hearing [22, 23]. These are (i) frequency decomposition of the input, (ii) amplification, (iii) nonlinear transfer characteristics, (iv) dynamic adaptation of sensing properties and (v) emission of sounds (without input). Thereby, amplification and nonlinear properties are obtained for positive feedback strength a_{f} and negative offset u_{DC}. However, for the inner ear a frequency-selective reduction of the gain is discussed as well besides the above discussed amplification [24, 25]. It is assumed to arise from damping of the activity of outer hair cells controlled by efferent

feedback. Since outer hair cells are believed to provide the amplification, dampening their activity leads to a reduction of the effective gain of the inner ear. Since outer hair cells are activated by the basilar membrane, which performs a frequency decomposition of the sound signal by creating a frequency-place coding, the amplification and damping is specific for each frequency bands. This effect is thought to enhance speech processing in noisy situations by reducing the response for the noisy signals [26]. Here, we study if similar effects, like frequency-selective amplification and dampening of the sensor response are possible by investigating the dependence of the dynamic and sensing properties in dependence of the sign of the feedback parameters a_f and u_DC.

DEPENDENCE OF SENSOR SYSTEM PROPERTIES ON SIGN OF FEEDBACK PARAMETERS



(a) Case I: Positive a_f and positive u_DC

(b) Case II: Positive a_f and negative u_DC

(c) Case III: Negative a_f and positive u_DC

(d) Case IV: negative a_f and negative u_DC

FIGURE 3: Amplitude of sensing signal in dependence of feedback strength a_f for different driving voltages of the loudspeaker (see legends), corresponding to different SPL.The microbeam has a resonance frequency of 3.63 kHz. The panels represent the four different feedback cases, as summarized in table I.

Since the feedback signal u_act depends on a_f and u_DC (see eq. 1), , four possible combinations of the signs of the feedback parameters are possible, i.e. both parameters positive, both parameters negative and two cases with one negative and one positive. However, since we use a thermo-mechanical actuation principle, we expect only two different cases, as explained in the following. For thermo-mechanical actuation, the feedback signal is used to heat the beam due the resistance of the aluminium wire on top of the beam, at which the feedback voltage is applied.

This heating results in a beam deflection based on the bimorph effect [21]. Since the electrically induced temperature increase of the beam is linearly depending on the introduced electrical power, the effective driving signal for the beam, i.e., the power of the electric signal, is proportional to the square of the feedback signal

$$P_h \approx (a_\mathrm{f} u_\mathrm{sens} + u_\mathrm{DC})^2 = a_\mathrm{f}^2 u_\mathrm{sens}^2 + 2 a_\mathrm{f} u_\mathrm{sens} u_\mathrm{DC} + u_\mathrm{DC}^2.$$

Thus, combinations of similar signs of a_f and u_DC will result in only positive signs of the introduced power, whereas opposite signs of a_f and u_DC will yield a negative term.

To study for the four different feedback cases (summarized in table I) the dependence of sensor properties on the sign of feedback parameters, the response of a sensor with a resonance frequency of 3.63 kHz to acoustic excitation was measured. The results are shown in figure 3 (a)-(d). Thereby, a single tone signal of 3.63 kHz, generated by a signal generator, was applied to drive the loudspeaker. Different driving voltages are used from $U_D = 0$ mV (no sound) to 30 mV, corresponding to different sound pressure levels (SPL). Note that the sound pressure is linearly depending on the amplitude of the driving signal. For each combination of feedback strength a_f and amplitude of the driving signal U_D, one time series was recorded and the maximal amplitude of the sensing signal was determined, occurring at the resonance frequency. In all cases $|u_\mathrm{DC}| = 200$ mV was used.

As discussed above, it is clearly visible from Fig. 3 that always two feedback cases yield similar responses. The case $sign(a_\mathrm{f}) = sign(u_\mathrm{DC})$ yields a decrease of amplitude with increasing feedback strength, whereas the case of opposite signs leads to an increase of sensing amplitude with increasing feedback strength a_f. The initial sensitivity, given by the slope of $u_\mathrm{sens}/U_\mathrm{D}$, is ≈ 5 for the passive mode in all four cases. It increases up to ≈ 80 for $a_\mathrm{f} = 0.5$ for the cases II and III (opposite signs) and decreases to ≈ 0.5 for $a = 2$ for the cases I and IV (similar signs). This corresponds to a total change in sensitivity by 44 dB. Thus by changing the sign of the feedback parameters, we can control the sensitivity of the sensor in a wide range. This enables not only to increase the sensor gain, as was previously shown (see e.g. [27]), but also to dampen the sensors response. The latter effect can be used to implement frequency-selective and sound processing based damping of unwanted signals, as discussed for the inner ear in the frame of efferent feedback [24, 25].

Besides this, for the cases II and III with opposite signs, a strong increase for a certain feedback strength ($a_\mathrm{f} \approx 0.55$) is observed. This is the indication of the Hopf bifurcation [16, 18]. As a result of this, self-excited oscillations with sine wave shape are observed for larger feedback strengths with a frequency of 3.63 kHz corresponding to the resonance frequency of the beam. For the cases with similar signs, this strong increase in amplitude is observed as well, if very large feedback strengths are considered. This is visible from figure 4, which shows results from simulations of the sensor system for the cases III and IV.

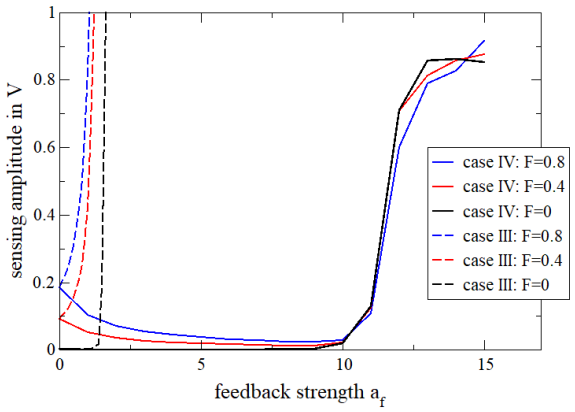
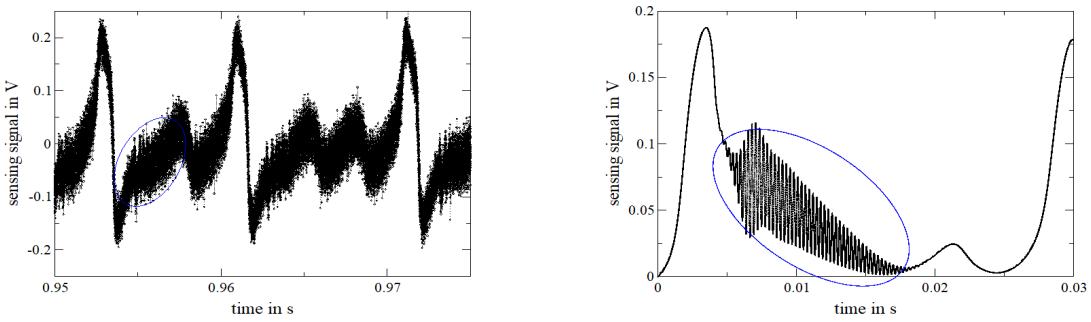


FIGURE 4: Amplitude of sensing signal in dependence of feedback strength a_f for case III and case IV obtained from numerical simulations of a beam with resonance frequency of 3.63 kHz.

Thereby, the feedback strength at bifurcation, i.e., the critical feedback strength a_c, is much larger for case IV (similar signs) than for case III (opposite signs). The strong increase in amplitude observed for $|a_\mathrm{f}| > a_c$ in fig. 4 corresponds to the onset of self-excited oscillations, similar to case II and III. Thereby, particularly interesting is the



(a) Case IV: Measured time series for $a_f = -5$ and $u_{DC} = -200\,\text{mV}$.

(b) Case II: Simulated time series $a_f = 11$ and $u_{DC} = 200\,\text{mV}$.

FIGURE 5: Time series of sensing signal for case IV obtained from measurements and for case II obtained from numerical simulations. In both cases, two oscillations are visible: a slow one with a frequency of ≈ 100 Hz and a fast one (marked by blue ellipse), on the slope of the slow one, with a frequency 3.63 kHz corresponding to the first mode response of the cantilever.

shape of the self-excited oscillations in case IV and II (see figure 5). In these cases, a complex oscillatory pattern is observed, composed of a slow oscillation ($\approx 100\,\text{Hz}$) with large amplitude and a faster oscillation (3.63 kHz) with small amplitude, visible in the rise/fall of the slow oscillation. However, if this behaviour is due to a Hopf bifurcation as well (as in the cases II and III), and if these oscillations correspond to a limit cycle and their nature requires further research.

CONCLUSION

We presented a bio-inspired acoustic sensor that provides several functionalities of the cochlea. These are frequency decomposition and amplification of the sound signal, nonlinear transfer characteristics and dynamic adaptation of sensing properties to the input and acoustic environment. Those functionalities are obtained by resonant operation of the transducer in combination with a high-speed feedback. The transducer is a silicon beam with integrated actuation and deflection sensing. The feedback loop gives precise control for the sensitivity and the linearity of the sensor. The polarity of the feedback parameters a_f and u_{DC} determines if the sensitivity and gain are increased (opposite sign of feedback parameters) or decreased (similar sign of feedback parameters) compared to the passive case without feedback, as summarized in table I. Combining both polarities, the sensitivity can be tuned in a large range by up to 44 dB. If the feedback parameters cross a certain threshold, self-excited oscillations occur, which can be addressed to the system undergoing a Hopf bifurcation. Interestingly, in the case of opposite polarity, the self-excited oscillations exhibit a sine-wave shape and the sensor can emit sound. This resembles the oto-acoustic emissions of the inner ear. However, in the case of similar polarity of the feedback parameters, the self-excited oscillations exhibit a complex dynamic behaviour, which contains slow and fast oscillations.

TABLE I: Feedback cases in dependence of the sign of feedback strength a_f and offset u_{DC} and their effect on the gain of the sensor system for excitation by sound.

sign of a_f \ sign of u_{DC}	positive	negative
positive	case I: decrease of gain with feedback (damping)	case II: increase of gain with feedback (amplification)
negative	case III: increase of gain with feedback	case IV: decrease of gain with feedback

ACKNOWLEDGMENTS

Funded by the Deutsche Forschungsgemeinschaft (German Research Foundation) - Project-ID 434434223 – SFB 1461 and the Carl-Zeiss-Stiftung in the project 'Memristive Materials for Neuromorphic Engineering (MemWerk)'.

REFERENCES

1. E. Kandel, J. Schwartz, T. Jessell, S. Siegelbaum, and A. Hudspeth, *Principles of Neural Science* (McGraw-Hill, 2012).
2. M. A. Ruggero, S. S. Narayan, A. N. Temchin, and A. Recio, "Mechanical bases of frequency tuning and neural excitation at the base of the cochlea: Comparison of basilar-membrane vibrations and auditory-nerve-fiber responses in chinchilla," Proceedings of the National Academy of Sciences **97**, 11744–11750 (2000).
3. A. Hudspeth, "Integrating the active process of hair cells with cochlear function," Nature Review Neuroscience **15**, 600 (2014).
4. A. J. Hudspeth, F. Jülicher, and P. Martin, "A critique of the critical cochlea: Hopf—a bifurcation—is better than none," Journal of Neurophysiology **104**, 1219–1229 (2010).
5. T. A. J. Duke and F. Jülicher, "Critical oscillators as active elements in hearing," in *Active Processes and Otoacoustic Emissions in Hearing* (2008) p. 63.
6. D. T. Kemp, "Stimulated acoustic emissions from within the human auditory system," The Journal of the Acoustical Society of America **64**, 1386–1391 (1978).
7. F. Abreu Araujo, M. Riou, J. Torrejon, S. Tsunegi, D. Querlioz, K. Yakushiji, A. Fukushima, H. Kubota, S. Yuasa, M. D. Stiles, and J. Grollier, "Role of non-linear data processing on speech recognition task in the framework of reservoir computing," Scientific Reports **10**, 1 (2020).
8. J. L. Flanagan, "Speech synthesis," in *Speech Analysis Synthesis and Perception* (Springer, 1972) pp. 204–276.
9. J. Lazzaro and C. Mead, "Circuit models of sensory transduction in the cochlea," in *Analog VLSI Implementation of Neural Systems* (1989) p. 85.
10. T. J. Hamilton, C. Jin, A. van Schaik, and J. Tapson, "An active 2-D silicon cochlea," IEEE Transactions on Biomedical Circuits and Systems **2**, 30–43 (2008).
11. S.-C. L. (Editor), T. D. (Co-Editor), G. I. (Co-Editor), A. Whatley, and R. Douglas, *Event-based neuromorphic systems* (John Wiley & Sons, 2015).
12. A. Jiménez-Fernandez, E. Cerezuela-Escudero, L. Miro-Amarante, M. J. Domínguez-Morales, F. de Asís Gómez-Rodríguez, A. Linares-Barranco, and G. Jiménez-Moreno, "A binaural neuromorphic auditory sensor for FPGA: A spike signal processing approach," IEEE Transactions on Neural Networks and Learning Systems **28**, 804–818 (2017).
13. C. Lenk, K. Ved, S. Durstewitz, T. Ivanov, M. Ziegler, and P. Hövel, "Bio-inspired (neuromorphic) acoustic sensing," in *Springer Series in Bio- and Neurosciences* (Springer, 2022).
14. H. S. Wang, S. K. Hong, J. H. Han, Y. H. Jung, H. K. Jeong, T. H. Im, C. K. Jeong, B.-Y. Lee, G. Kim, C. D. Yoo, and K. J. Lee, "Biomimetic and flexible piezoelectric mobile acoustic sensors with multiresonant ultrathin structures for machine learning biometrics," Science Advances **7**, eabe5683 (2021).
15. C. Lenk, S. Gutschmidt, and I. W. Rangelow, "Vorrichtung und Verfahren zur Detektion von Schall in Gasen und Flüssigkeiten," Patent **DE 102018117481B8** (2019).
16. C. Lenk, L. Seeber, M. Ziegler, P. Hövel, and S. Gutschmidt, "Enabling adaptive and enhanced acoustic sensing using nonlinear dynamics," in *2020 IEEE International Symposium on Circuits and Systems (ISCAS)* (IEEE, 2020) pp. 1–4.
17. C. Lenk, T. Ivanov, V. Gubbi, K. Ved, M. Ziegler, T. Fritsch, J. Küller, and D. Beer, "Bio-inspired, nonlinear and adaptive acoustic sensing - study of sensor design," in *Fortschritte der Akustik - DAGA2022* (2022) pp. 1–4.
18. N. Lam, S. Hayashi, and S. Gutschmidt, "A novel mems sensor concept to improve signal-to-noise ratios," International Journal of Non-Linear Mechanics **139**, 103863 (2022).
19. Z. Jackson, C. Souza, J. Flaks, Y. Pan, H. Nicolas, and A. Thite, "Jakobovski/free-spoken-digit-dataset: v1.0.8," (2018).
20. I. W. Rangelow, T. Ivanov, A. Ahmad, M. Kaestner, C. Lenk, I. S. Bozchalooi, F. Xia, K. Youcef-Toumi, M. Holz, and A. Reum, "Review article: Active scanning probes: A versatile toolkit for fast imaging and emerging nanofabrication," Journal of Vacuum Science & Technology B **35**, 06G101 (2017).
21. T. Ivanov, "Piezoresistive cantilevers with an integrated bimorph actuator," Thesis (2003), urn:nbn:de:hebis:34-1153.
22. S. Durstewitz, C. Lenk, and M. Ziegler, "Bio-inspired acoustic sensor with gain adaptation enhancing dynamic range and onset detection," in *2022 IEEE International Symposium on Circuits and Systems (ISCAS)* (2022).
23. C. Lenk, K. Ved, S. Gutschmidt, T. Ivanov, P. Hövel, T. Meurer, and M. Ziegler, "Dynamically adaptable acoustic sensor with nonlinear filtering functionality," in *47th international conference on Micro and Nano Engineering* (2021).
24. A. B. Elgoyhen, "Cochlear efferent innervation: function, development and plasticity," Current Opinion in Physiology **18**, 42–48 (2020).
25. P. Dallos, A. Popper, and R. Fay, *The cochlea* (Springer, 1996).
26. J. J. Guinan, "Olivocochlear efferents: Their action, effects, measurement and uses, and the impact of the new conception of cochlear mechanical responses," Hearing Research **362**, 38–47 (2018), annual Reviews 2018.
27. C. Lenk, L. Seeber, and M. Ziegler, "Tuning acoustic sensing properties of mems cantilever by nonlinear operation," in *Mikro-Nano-Integration; 8th GMM-Workshop* (2020) pp. 85–87.

What Do Recent Discoveries in Cochlear Mechanics Tell Us About Otoacoustic Emissions?

Jonathan H. Siegel

Knowles Hearing Center and Roxelyn & Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, Evanston, Illinois, USA

Corresponding author: j-siegel@northwestern.edu

Abstract. Recent optical measurements of vibrations of the organ of Corti contrast with those of the basilar membrane in their larger amplitude and much wider bandwidth of active gain and nonlinearity [6-9, 14, 15, 18, 19]. A report by Ren and He (15) demonstrates a considerably richer spectrum of intermodulation distortion in the motion of the reticular lamina than in the basilar membrane of the Mongolian gerbil cochlea. Data from this species from my lab also show a broad distortion spectrum much more similar to the reticular lamina than the basilar membrane in the Ren and He data. This result, along with other evidence of broad-band nonlinearity in ear canal acoustics strongly suggests the need to reevaluate otoacoustic emission (OAE) phenomena. In particular, these experimental data do not support reverse slow basilar membrane traveling waves [22] as the dominant mode of conducting OAE signals to the ear canal, but they are instead conducted primarily by fluid pressure in scala vestibuli, at least for this species in the basal region of the cochlea. Several important unresolved issues in OAEs may be resolved by simultaneous measurements of vibrations in the organ of Corti and basilar membrane, together with comparable measurements in ear canal acoustics.

INTRODUCTION

Recent discoveries using advanced optical techniques [6-9, 14, 15, 18, 19] have allowed great progress in understanding how the hair cells in the organ of Corti cause amplification of vibrations of the basilar membrane, particularly in the peak region of the basilar membrane response. It was surprising that the organ of Corti vibrations in living animals were consistently larger than those of the basilar membrane, but also show a broader range of frequencies where apparently active gain and nonlinearity was observed. It is particularly striking that gain well below the local characteristic frequency (CF) does not couple into enhancing basilar membrane traveling waves [7, 9]. There is currently no explanation for where energy generated well below-CF flows and what purpose it has, yet conservation of energy demands that it must be dissipated somewhere, if not by the basilar membrane. It is a common assumption that efficiency must have guided the evolution of the cochlear amplifier, but there appears to be a limit. One possibility is that this excess power is funneled into the ear canal and contributes to otoacoustic emissions.

The new intracochlear mechanical discoveries lead me to propose several questions about the origins of otoacoustic emissions and the route(s) by which the emission signals leave the cochlea and enter the middle ear. The most commonly proposed route of transmission of OAE signals has been via slow-wave retrograde transmission along the basilar membrane, where transmission to the middle was envisioned to result from the pressure in scala vestibuli at the basal end of the cochlea [22]. However, Ren and colleagues [10, 11] measured only forward-propagating distortion waves (or responses to localized current injection [12]) and reasoned that the coupling from

hair cells to the middle ear was carried primarily by fast pressure waves, with secondary slow-wave forward propagation. Furthermore, the size of the active and/or nonlinear region giving rise to OAEs has been largely based on speculation, not on systematic experimental measurement. This speculation has, in turn, led to confusion about how to interpret some experimental measurement of OAEs and complicates interpretation of OAEs in human diagnostics. For example, only a few studies have reported intermodulation distortion as a function of the frequency separation of a pair of stimulus tones [5, 21], which show a steep reduction in distortion amplitudes with increasing f_2/f_1, suggesting that nonlinear intermodulation distortion is spatially restricted on the basilar membrane.

If the basilar membrane is the only route of propagation back to the middle ear, then standing waves should be detected basal to the place where the signal is launched. In fact, standing waves are a common feature of one-dimensional cochlear models (e.g., [13, 17]). However, amplitude and phase *vs* frequency plots measured experimentally at basal locations are typically remarkably smooth, indicating a dominance of forward propagation along the basilar membrane [10-12] and in reticular lamina motion as well (e.g., [6, 7, 9, 19]). There are a few exceptions where spectral ripples are seen below CF [8], but these do not appear to represent the large fluctuations that might be expected were reverse slow-wave propagation the main route of coupling to the middle ear. The thrust should be to extract estimates of the relative contributions of the two proposed reverse propagation mechanisms to determine relative dominance.

These uncertainties have also contributed to interpreting other OAE phenomena, including observations that interference tones well above f_2 affect the levels of DPOAEs [16] and that the pressure response to low-level low-frequency probe tones shows nonlinear interaction with suppressor tones much higher in frequency [2, 3]. In the latter case, the primary site of nonlinear interaction between the probe and suppressor tones is almost certainly near the basal location of the peak of excitation by the suppressor tone, but in this situation the basilar membrane shows no comparable nonlinearity.

The recent studies of intracochlear vibrations may resolve the paradox, but new sets of measurements, combining vibrations in the organ of Corti and basilar membrane and ear canal acoustics are badly needed (Figure 1). None of the connections between any two of these measurement locations are well-characterized at present. Data presented in this paper will hopefully demonstrate the validity of this proposal.

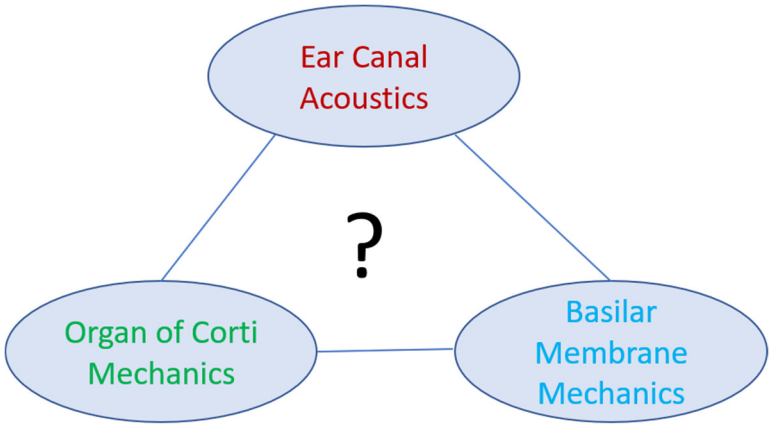


FIGURE 1. Diagram showing the three measurement sites needed to resolve the relation between cochlear mechanics and otoacoustic emission phenomena measured in ear canal acoustics. None of the three connections between the phenomena are well understood.

METHODS

Otoacoustic emissions were measured in deeply-anesthetized, tracheotomized Mongolian gerbils with body temperature maintained by a heating pad as described in detail elsewhere [2]. The cartilaginous part of the ear canal and pinna are removed to allow optimal coupling of an Etymotic ER-10B+ otoacoustic emission probe using a custom speculum. The transfer function of the speculum/emission probe microphone was measured and used to compensate the emission signals. Stimulus levels were calibrated in cavity approximating the external ear of the gerbil, using a Bruel and Kjaer 4138 1/8" microphone. Distortion product otoacoustic emissions were measured using synchronous averaging and the EMAV software. All animal procedures were approved by the Animal Care and Use Committee of Northwestern University.

RESULTS

The data in Figure 2A were measured in 11 anesthetized Mongolian gerbils with no evidence of inner ear damage. The level of the $2f_1$-f_2 distortion product peaked broadly near f_2=20 kHz and the $3f_1$-$2f_2$ and f_2-f_1 products were also at least 10 dB above the noise floor throughout the tested frequency range (f_2 < 40 kHz). The $4f_1$-$3f_2$ product was also detected above the noise floor, especially for f_2 below ~8 kHz. In contrast, the $2f_2$-f_1 distortion product was only detected consistently above the noise floor for f_2 below ~8 kHz. The dotted vertical line in Figure 2A marks the f_2 = 24 kHz condition illustrated in Figure 2B adapted from the Ren and He study of intermodulation distortion measured in reticular lamina and basilar membrane vibrations [18]. The Gerbil DPOAE spectral data in Figure 1A much more closely matches that of Ren and He than does their basilar membrane measurements (Figure 1B, right). The only measured intermodulation distortion product in the latter is $2f_2$-f_1, which is not detected in the Gerbil DPOAE measurements for f_2 > 8 kHz (Figure 1A). But the absence of distortion products at frequencies below f_1 in the basilar membrane measurements of Ren and He is striking.

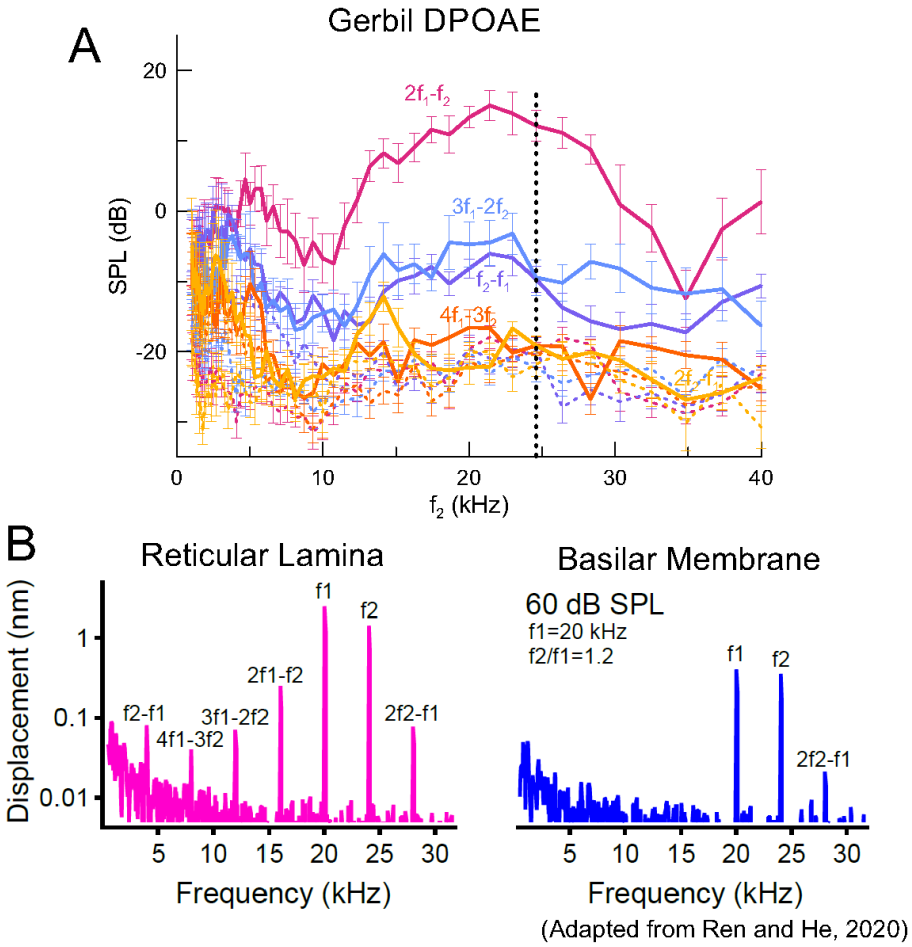


FIGURE 2. Intermodulation distortion products measured in Mongolian gerbils in ear canal DPOAEs in our unpublished data (A) and reticular lamina and basilar membrane vibrations by Ren and He [18] (B). The broad spectrum of intermodulation distortion is apparent both in the DPOAEs (A) and in reticular lamina vibrations (B), but not in the basilar membrane at the same location. The dotted vertical line in (A) shows the f_2=24 kHz condition depicted in (B). The relative prominence of different orders of distortion product varies widely with the frequencies of stimuli.

DISCUSSION/SUMMARY

The remarkable contrast between the internal mechanics of the organ of Corti and the basilar membrane is primarily in the bandwidth of active and nonlinear behavior. For those interested in the otoacoustic emission phenomenon, it is important to learn the degree to which OAEs correspond to organ of Corti or basilar membrane mechanics, which in turn may correspond to the relative importance of the two purported reverse coupling pathways. The data presented in Figure 2 strongly suggest strong direct coupling between the organ of Corti and ear canal acoustics, implying that coupling is primarily through the fluid pressure in scala vestibuli. If so, then the wide bandwidth of two-tone suppression and interference tone effects on OAEs may finally have an explanation.

However, while this result presented in this paper obviously applies to the basal turn of the Gerbil cochlea, it remains to be seen as a general principle. Even though there has been no consistent experimental evidence for reverse slow-wave propagation of OAE signals in direct measurements of basilar membrane motion, the relative dominance of this propagation mechanism vs coupling through fast pressure waves may depend on both cochlear location and species. It would seem best to make measurements in the three sites depicted in Figure 1 in the same experimental animal, and at different locations and in various species before generalizing these results.

Computational models play an important role in making predictions that can be tested in experiments. Three-dimensional models that approximate both forward and reverse propagation on the basilar membrane as well as coupling through fluid pressure, such as that by Bowling and Meaud [1], make explicit testable predictions. Even simpler one-dimensional models, such as that of Moleti and Sisto [17] have been helpful in demonstrating that they fail to produce broad-band ear canal phenomena demonstrated experimentally. This model exhibits broad spatial nonlinear interaction in the reticular lamina but basilar membrane nonlinearity restricted to the peak regions of the stimuli due to loose coupling between the two structures. This result indirectly argues for a second propagation mechanism, since the basilar membrane can not transmit signals that are not evident in its motion. Significantly, the Moleti and Sisto model is consistent with the steep dependence of intermodulation distortion on f_2/f_1 reported by Cooper and Rhode [5].

But do otoacoustic emissions provide any experimental information about inner ear mechanics that cannot be better studied in direct intracochlear measurements? One advantage is the ability to detect small acoustic signals and nonlinear interactions originating from a broad region of the cochlea, where direct measurements at single intracochlear sites may be obscured by system noise. This is similar to the cochlear microphonic, which can reveal a signal-to-noise ratio (S/N) even greater than small acoustic signals in the ear canal [4]. This might guide intracochlear measurements to focus on interesting phenomena that might not be readily detected using rapid scans that do not optimize S/N. OAEs provide sensitive tests of system linearity, both as measuring small deviations from linearity in the presence of intermodulation distortion (i.e., DPOAEs) and in two-tone suppression, typically measured as a change from a control condition. What appears to be linear growth in amplitude with increasing level of below CF in reticular lamina vibrations may appear linear to the eye, but this is a very insensitive measure of linearity.

OAEs are currently also of great importance in evaluating animal models with genetic manipulations that affect hair cells and have become part of batteries of tests of peripheral human hearing function. How a better understanding of OAE mechanisms might improve these applications is not clear, but perhaps we've been missing something important.

ACKNOWLEDGMENTS

Supported by the Knowles Hearing Center and Northwestern University.

REFERENCES

[1] Bowling T, Meaud J (2018) Forward and reverse waves: Modeling distortion products in the intracochlear fluid pressure Biophys J 114:747–757

[2] Charaziak KK, Siegel JH (2015a) Tuning of SFOAEs evoked by low-frequency tones is not compatible with localized emission generation. J Assoc Res Otolaryngol 16:317-329

[3] Charaziak KK, Siegel JH (2015b) Low-frequency Tone-pip-evoked Otoacoustic Emissions Originate over a Broad Cochlear Region in Chinchillas in: *Mechanics of Hearing: Protein to Perception* AIP Conference

Proceedings 1703, edited by D.P. Corey and K.D. Karavitaki. (American Institute of Physics, Melville, NY, 2015), pp. 090016-1–090016-6
[4] Charaziak KK, Shera CA, Siegel JH (2017) Using cochlear microphonic potentials to localize peripheral hearing loss. Frontiers in Neuroscience. doi: 10.3389/fnins.2017.00169
[5] Cooper NP, Rhode WS (1997) Mechanical responses to two-tone distortion products in the apical and basal turns of the mammalian cochlea J Neurophysiol 78: 261–270
[6] Cooper NP, Vavacou A, Van der Heijden, M (2018) Vibration hotspots reveal longitudinal funneling of sound-evoked motion in the mammalian cochlea. Nat Commun 9, 3054 doi.org/10.1038/s41467-018-05483-z
[7] Dewey JB, Applegate BE, Oghalai JS (2019) Amplification and suppression of traveling waves along the mouse organ of Corti: evidence for spatial variation in the longitudinal coupling of outer hair cell-generated forces. J Neurosci 16:2608-2618
[8] Dong W, Xia A, Raphael PD, Puria S, Applegate B, Oghalai JS (2018) Organ of Corti vibration within the intact gerbil cochlea measured by volumetric optical coherence tomography and vibrometry J Neurophysiol 120: 2847–2857
[9] Fallah E, Strimbu CE, Olson ES (2019) Nonlinearity and amplification in cochlear responses to single and multi-tone stimuli. Hear Res 377:271–281
[10] He W, Fridberger A, Porsov E, Grosh K, Ren T (2008) Reverse wave propagation in the cochlea. Proc Natl Acad Sci USA 105: 2729–2733
[11] He W, Fridberger A, Porsov E, Ren T (2010) Fast reverse propagation of sound in the living cochlea, Biophys J 98: 2497–2505
[12] He W, Ren T (2013) Basilar membrane vibration is not involved in the reverse propagation of otoacoustic emissions Sci Rep 3 : 1874 DOI: 10.1038/srep01874
[13] Kim DO, Molnar CC, Matthews JW (1980) Cochlear mechanics: nonlinear behavior in two-tone responses as reflected in cochlear-nerve-fiber responses and in ear-canal sound pressure. J Acoust Soc Am 67:1704–1721
[14] Lee HY, Raphael PD, Park J, Applegate BE, Ellerbee AK, Oghalai JS (2015) Noninvasive in vivo imaging reveals differences between tectorial membrane and basilar membrane traveling waves in the mouse cochlea Proc Natl Acad Sci USA 112:3128–3133
[15] Lee HY, Raphael PD, Xia A, Kim J, Grillet N, Applegate BE, Ellerbee Bowden AK, Oghalai JS (2016) Two-dimensional cochlear micromechanics measured in vivo demonstrate radial tuning within the mouse organ of Corti J Neurosci 36:8160-8173
[16] Martin GK, Stagner BB, Lonsbury-Martin BL (2010) Evidence for basal distortion-product otoacoustic emission components. J Acoust Soc Am 127:2955–2972
[17] Moleti A, Sisto R (2020) Does the “reticular lamina nonlinearity” contribute to the basal DPOAE source? J Assoc Res Otolaryngol 21:463-473
[18] Ren T, He, W (2020) Two-tone distortion in reticular lamina vibration of the living cochlea Commun Biol 3:35; https://doi.org/10.1038/s42003-020-0762-2
[19] Ren T, He, W, Kemp DT (2016) Reticular lamina and basilar membrane vibrations in living mouse cochleae Proc Natl Acad Sci USA 113:9910–9915
[20] Rhode WS (2007) Mutual suppression in the 6 kHz region of sensitive chinchilla cochleae J Acoust Soc Am 121: 2805–2818
[21] Robles L, Ruggero MA, Rich NC. (1997) Two-tone distortion on the basilar membrane of the chinchilla cochlea. J Neurophysiol 77:2385–2399
[22] Shera CA, Guinan JJ (1999) Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian otoacoustic emissions J Acoust Soc Am 105:782–798
[23] Siegel JH (2018) Does reticular lamina active gain explain broad suppression tuning of SFOAEs? In: To the Ear and Back Again - Advances in Auditory Biophysics: Proceedings of the 13th Mechanics of Hearing Workshop Bergevin C, Puria S (eds) American Institute of Physics Inc., Vol. 1965 140003

Stimulus Frequency Otoacoustic Emissions Extracted by Pharmacologic Blocking of Outer Hair Cells Without Sound Suppression or Subtractive Scaling

Shawn S. Goodman[1,a)], Daniel Tay[1,b)], Jeffery T Lichtenhan[2,3,c)], John J Guinan Jr[4,5,d)]

[1]*Communication Sciences and Disorders, University of Iowa, Iowa City, Iowa, USA*
[2]*Department of Otolaryngology, Washington University School of Medicine in St. Louis, St. Louis, Missouri, USA*
[3]*Department of Otolaryngology, University of South Florida, Tampa, Florida, USA*
[4]*Department of Otolaryngology, Harvard Medical School, Boston, Massachusetts, USA*
[5]*Eaton-Peabody Laboratories, Massachusetts Eye and Ear, Boston, Massachusetts, USA*

[a)] *Corresponding author: shawn.s.goodman@gmail.com*
[b]*daniel-tay@uiowa.edu*
[c]*jlichtenhan@gmail.com*
[d]*john_guinan@meei.harvard.edu*

Abstract. Stimulus-frequency otoacoustic emissions (SFOAEs) are important non-invasive probes of cochlear function. The SFOAE is usually separated from the probe sound by measuring a response that contains both the probe sound and the SFOAE, and then subtracting a separate measurement of the probe sound alone. The probe-alone measurement is obtained using (1) a high-level, near-frequency "suppressor" sound assumed to suppress the SFOAE, or (2) the probe sound at a high level where the SFOAE is small compared to the probe (i.e., is compressed) and is presumed to be negligible after the high-level measurement is scaled down to the probe level. To the extent that the high-level sound has a negligible SFOAE, the complete SFOAE should theoretically be extracted by the subtraction. However, a potential complication is that the high-level sounds may produce unintended SFOAE components, particularly from basal cochlear regions. Here we use a novel technique to address the extent to which the high-level suppressor sound and subtraction extracts the full SFOAE without producing unintended additional energy at the probe frequency. Measurements were made with a standard high-level suppressor-tone extraction paradigm, and the resulting SFOAEs were compared to SFOAEs obtained using a pharmacologic SFOAE reduction that does not require a high-level sound. We made SFOAE measurements in guinea pigs while sequentially abolishing outer-hair-cell (OHC) function and SFOAE amplification by slowly perfusing salicylate or KCl solutions into the cochlear apex. The solution fronts were driven uniformly along the cochlear length toward the cochlear aqueduct in the cochlear base. This approach removes cochlear amplification at a given cochlear place without simultaneously affecting amplification or responses that originate basal to this place. At the end of the perfusion when OHC function and SFOAE production were pharmacologically blocked throughout the entire cochlear length, the probe sound alone (i.e. with no accompanying SFOAE) could be measured directly without the need for a high-level suppressor. SFOAEs were calculated by subtracting probe-only ear canal measurements made after the perfusion from measurements made before and during the perfusion. SFOAEs extracted using the apical perfusion technique, i.e. pharmacologic blocking, differed little from the SFOAEs extracted using a near-frequency high-level suppressor sound but were, on average, a few dB higher in amplitude. This suggests that the near-frequency high-level tone suppressed most, but not all, of the SFOAE.

INTRODUCTION

Measurements of auditory and vestibular function are used in laboratories, clinics, and operating rooms worldwide each day. The overarching goal of our work is to improve the interpretation of functional measurements from ear-brain system by improving our understanding of their cellular and spatial origin. At present, we are addressing the origin of measurements in the auditory periphery, in particular: otoacoustic emissions (OAEs). One of our novel approaches includes the use of a cochlear perfusion technique in which solutions are delivered through a pipette sealed into the cochlear apex and travel along the length of scala tympani towards the cochlear aqueduct in the base. Solutions containing ototoxic drugs are used to sequentially abolish closely spaced regions along the cochlear length while changes in cochlear function are measured. The apical perfusion technique provides a method to remove cochlear amplification and responses from one place along the cochlea without simultaneously affecting amplification or responses that originate basal to this location. For example, apical perfusions overcome the limitations of acoustic

trauma from noise exposure that affect a given region of cochlear amplification and also affect locations basal to the location of interest. We have used the apical perfusion technique to identify the origin of otoacoustic and neural measurements from both normal and diseased ears (Lichtenhan et al. 2014, 2016; Lee et al. 2019; Goodman et al. 2020; Guinan et al. 2021).

Previously, Goodman et al. (2020) used the apical perfusion technique to address whether stimulus frequency otoacoustic emissions (SFOAEs) originate from the short-wave peak of the traveling wave or the long-wave tail region. Salicylate or KCl in artificial perilymph were apically perfused while SFOAEs and auditory nerve compound action potentials (CAP) measurements were made. The perfusion time when the CAP amplitudes reduced were used to identify the characteristic-frequency (CF) region of the probe tone (Lee et al. 2019), and the relative time when the amplitudes of SFOAEs reduced were used to understand the spatial origin of SFOAEs. It was found that SFOAEs originate ~1/2 octave basal to the CF region, not from the basal end of the cochlea, for high frequency (≥2 kHz) tones, and up to two octaves *apical* of CF for lower frequency tones. Here we use the apical perfusion technique to understand whether the suppression method for extracting SFOAEs works as envisioned, or if the suppressor tone creates additional SFOAE components, and whether the near-frequency suppressor method captures all the SFOAE.

METHODS

We evoked SFOAEs with a probe tone while perfusing the guinea pig cochlea from apex to base with ototoxic solutions (salicylate or KCl) that reduced or eliminated cochlear amplification sequentially from low frequency regions to high frequency regions. Salicylate affects cochlear amplification by blocking prestin-based OHC motility (Kakehata and Santos-Sacchi, 1996), while KCl affects cochlear amplification by depolarizing OHCs. Auditory nerve (AN) compound action potentials (CAPs) measurements were simultaneously recorded with SFOAE measurements. Reduction of CAP amplitudes by salicylate or KCl provided reference times for when the solutions reached the tone CF region (Lee et al. 2019). Both salicylate and KCl reduce CAPs when the fluid front reaches the CF region. In our previous work using apex-to-base cochlear perfusions with salicylate, KCl, and kainic acid, we showed that for low-level stimuli (~40 dB SPL), the perfusion time at which the solutions reduced the CAP amplitudes identified the time at which the solutions reached the region of the peak of the travelling wave (i.e., the CF region). Methodological details are as in Goodman et al. (2020). Experimental protocols were approved by the Institutional Animal Care and Use Committee at Washington University in St. Louis (20150091, 20180133).

Acoustic Extraction

SFOAEs were measured using the double-evoked suppression method (Keefe, 1998). Sets of three 250-ms stimuli with 10 ms rise-fall times were presented: (**1**) "*P*": a 40 dB SPL probe tone at frequency F_p, that evoked the SFOAE, (**2**) "*PS*": the probe tone plus a simultaneous 60 dB SPL suppressor tone at 50 Hz above F_p, and (**3**) "*S*": the suppressor tone alone. The simultaneous suppressor tone is assumed to prevent the generation of the SFOAE so that $P + S - PS$ = "residual" ≈ SFOAE. The three-stimulus set (*P*,*S*,*PS*) was presented in blocks of 20 repetitions and averaged. The noise floor was estimated as the standard error of the residuals within each block.

Pharmacologic Extraction

The cochlear perfusion provides a pharmacologic block of the production of SFOAEs, so that at the end of the perfusion there is no SFOAE and the ear-canal sound pressure from the probe tone, "P_{end}", is the probe sound alone. Ideally, the SFOAE produced during the perfusion can be obtained as SFOAE = P - P_{end}. Note that this does not require the presentation of a high-level sound for the measurement of P or P_{end}; only the probe-tone stimulus is used.

Unfortunately, the acoustic signal in the ear canal changes slowly in both magnitude and phase across the duration of the ~45-minute experiments. The changes are observable in the responses to a high-level stimulus that receive little contribution from SFOAEs and can be used to measure the changes in the stimulus itself. During the experiments, the measured high-level response magnitudes changed by up to 0.5 dB and also showed small phase changes with no consistent direction noted. Changes were also seen in data from control animals that had artificial perilymph perfused instead of an ototoxic solution. This suggests the changes were due to subtle changes in cochlear impedance, perhaps due to the fluid perfusion. Because of the slow changes in stimulus magnitude and phase, probe-tone responses from the post-perfusion pharmacologic SFOAE ablation, "P_{end}", cannot simply be subtracted from all earlier waveforms in the experiment. The P_{end} measurements must first be be corrected for the changes that occurred during the perfusion, then subtracted from P.

The changes that occurred during the experiment can be estimated from the magnitude and phase changes measured from a high-level reference stimulus across the duration of the experiment. This assumes that recordings from this high-level reference stimulus have relatively little SFOAE contribution and therefore provide a good estimate of changes in the stimulus itself (i.e., changes are not in the SFOAEs). For this reference we used the third stimulus,

the 60 dB SPL suppressor alone, "*S*". "*S''* is close in frequency to the probe stimulus (50 Hz above F_p) so changes in "*S*" should provide a good estimate of the changes that occurred for stimuli at the probe frequency.

Our objective in these correction calculations was to obtain an estimate of "P_{alone}" (the pressure produced by the sound source acting on the ear-canal impedance without the addition of an SFOAE) at each time during the measurement series, i.e., "$P_{alone}(t)$". To do this we first calculated the estimated change from time "*t*" to the measurement at the end of the perfusion: $\Delta(t) = S(t) / S_{end}$, where $S(t)$ is the value of "*S*" presented by itself (i.e., as the third element of the stimulus sequence) measured at time *t* in the perfusion, and S_{end} is the value of "*S*" at the end of the perfusion. Then $P_{alone}(t) = P_{end} / \Delta(t)$, and $SFOAE_{pharm}(t) = P(t) - P_{alone}(t)$, where $SFOAE_{pharm}$ is the SFOAE obtained by pharmacologic extraction. In these calculations both the amplitudes and phases were taken into account (the actual computations were done in the frequency domain).

Fitting procedure for SFOAE reductions

The amplitude-versus-perfusion-time measurements in the analysis window were fit with a mathematical function chosen because it had enough flexibility to capture the underlying shape of the reductions (including asymmetry) with only two parameters. The fitting procedure was described in Goodman et al. (2020).

RESULTS

SFOAE data were gathered on 12 guinea pigs that were perfused apex to base and had SFOAEs calculated by both methods. Figure 1 shows data from an example animal in which the probe tone was 3.5 kHz. This figure illustrates how we fit lines to the data so that statistics could be extracted and compared in a uniform way across animals. In this animal, SFOAEs from the pharmacologic extraction were a few dB larger than those from acoustic extraction (Fig. 1A). The fastest rate of SFOAE reduction occurred in the pharmacologic extraction before the acoustic extraction, but the difference was small (Fig. 1D). The statistics obtained from this animal are shown by the symbols and lines at the top of Figure 1D.

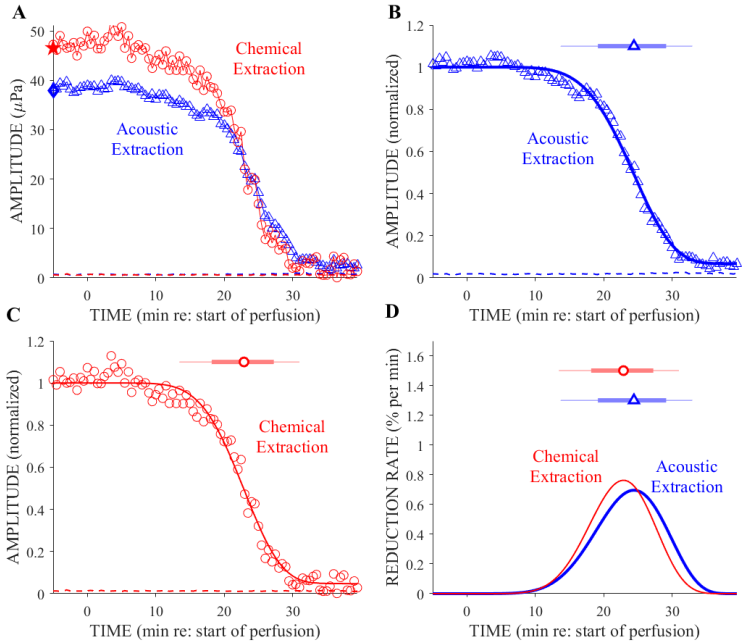


FIGURE 1. Example SFOAE amplitude reductions and curves fitted to these data (from animal DW165, probe frequency = 3.5 kHz, salicylate perfusion). SFOAEs obtained using acoustic extraction are shown using blue color and triangles; SFOAEs obtained using pharmacological extraction are shown in red color and circles. (A) SFOAE amplitudes versus perfusion time are shown with solid lines and markers. Dashed lines show noise floors. The blue diamond and red pentagram shapes located at -5 min on the x-axis show the amplitude values of the SFOAEs at the beginning of the experiment. (B and C) Amplitude versus perfusion time data and the curves fitted to these data as described in Methods. The amplitudes are linear measures normalized by their average values before the perfusion start. (D) The rate of reduction curves were calculated from the slopes of the curve fits in (B, C), and were normalized so the area under each curve is 100. SFOAEs obtained using acoustic extraction are shown with blue thick lines, and SFOAEs obtained using pharmacological extraction are shown with red thin lines. These curves show the percentage per minute reduction of the SFOAE amplification that originated from the cochlear place passed by the solution front

at that time. At the top of panels (B)–(D), triangle and circle symbols indicate the times of the reduction-rate peaks, and horizontal lines indicate perfusion times that encompass 68% reductions (thick lines) or 95% reductions (thin lines) around the peak.

Example data from four animals at a variety of probe frequencies are shown in Figure 2. Statistics extracted from all 12 animals are shown in Figure 3. Figure 3A shows the times at which the SFOAEs were reduced. SFOAEs from higher probe frequencies were generally reduced at later times because it took longer for the solution to flow from the apical injection point to the amplification region (the region from CF to ½ octave below CF) because higher CF regions are more basal. However, for a probe tone at a given frequency, there was relatively little difference in when the SFOAEs were reduced for pharmacologic vs. acoustic extraction (Fig. 3A, C). In most cases, the pharmacologically-extracted SFOAEs had higher amplitudes than the acoustic-extracted SFOAE (Fig. 3B, D).

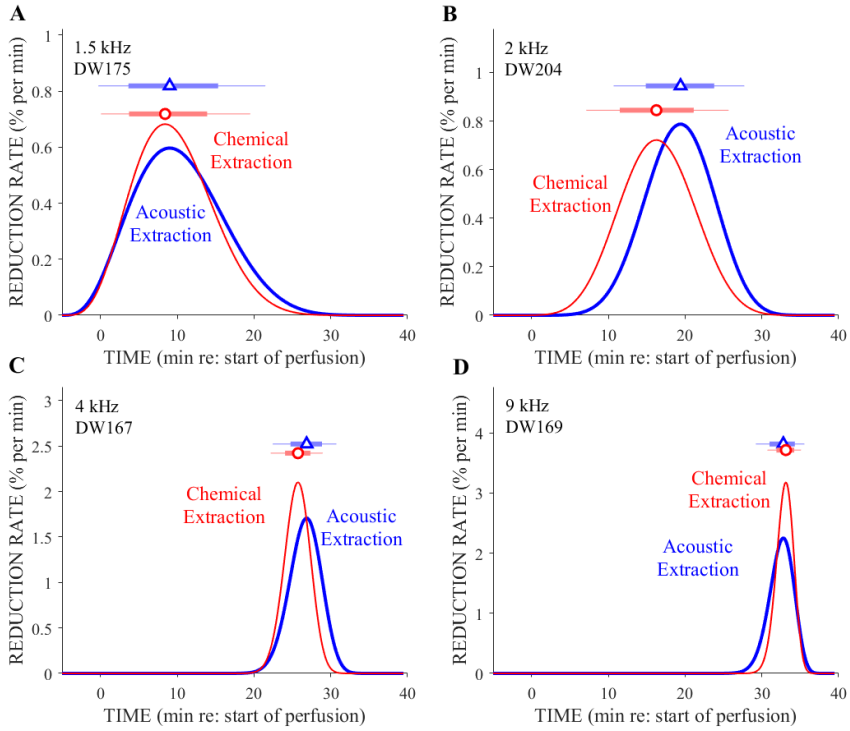


FIGURE 2. Example reduction-rate curves across a range of probe frequencies for SFOAEs obtained using acoustic extraction (blue thick lines) and SFOAEs obtained using pharmacological extraction (red thin lines). Above the curves, symbols (triangles and circles for acoustic and pharmacological extractions, respectively) show the time points at which the most rapid reductions occurred, and horizontal lines show the times over which 68 and 95% of the reduction occurred (thicker lines and thinner lines, respectively).

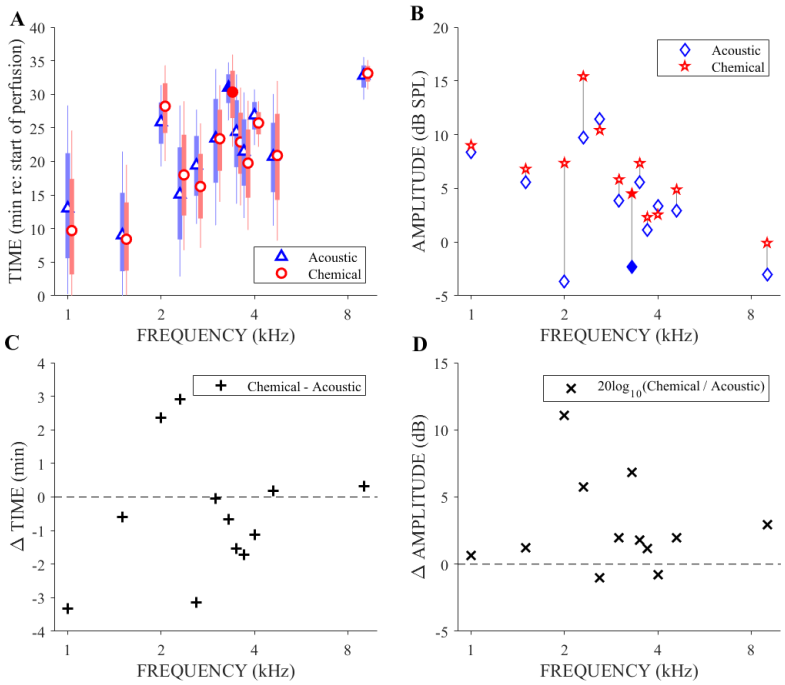


FIGURE 3. Differences between reduction times and amplitudes for SFOAEs obtained using acoustic extraction versus using pharmacologic extraction. (A) Reduction times as a function of probe frequency. Symbols show the peak reduction times; vertical lines show the duration times that encompass 68% (thick lines) and 95% (thin lines) rate reductions. Multiple measurements at 2, 3, and 3.5 kHz have been plotted with slight frequency offsets to avoid overlap. Open symbols show salicylate perfusions and filled symbols show KCl perfusions. (B) Amplitudes at the beginning of the experiment as a function of probe frequency (See Fig. 1A). Vertical lines connect the symbols to visually show the difference in acoustic versus pharmacological extraction for individual animals. (C) The arithmetic differences in peak reduction times are shown: Δ Time = peak reduction from pharmacological extraction minus peak reduction from acoustic extraction. (D) The ratios of amplitudes at the beginning of the experiment, expressed in decibels, are shown: Δ Amplitude = 20log10(amplitude from pharmacological extraction divided by amplitude from acoustic extraction).

Statistical Tests

Our data set constitute a paired set of matched samples. Such differences may be tested with the paired Student's t-test or with the Wilcoxon signed-rank test (which does not assume that the differences between paired samples are normally distributed). The Wilcoxon test can have greater statistical power than Student's t-test for non-normal distributions. We performed both tests on our data and obtained the same results with both tests; therefore, only the paired t-test results are reported.

The mean difference between peak reduction times obtained using acoustic extraction and pharmacologic extraction (Fig. 3C) was not significantly different from zero ($t(11) = -0.978$, $p = 0.35$). For the SFOAE amplitudes at the beginning of the experiment, the mean dB difference between acoustic extraction and pharmacologic extraction (Fig. 3D) was 2.79 dB (median = 1.87, S.D. = 3.475, min = -1.004, max = 11.05), and this value was significantly different from zero ($t(11) = 2.78$, $p = 0.018$).

DISCUSSION

SFOAEs obtained by pharmacologic extraction had significantly higher amplitudes than SFOAEs obtained by acoustical extraction, although by only a few dB. The peak reduction times, however, were not statistically different between the two extractions. The most straightforward explanation for the amplitude difference is that acoustic extraction did not reveal the full SFOAE. Acoustic-based SFOAE extractions rely on the amplification and nonlinear growth of emissions. Methods that use a suppressor, like the acoustical method in our experiment, are expected to extract the full emission only if the suppressor tone fully suppresses the emission. To the extent that there is an incomplete suppression, only a partial emission will be extracted. In this experiment, the suppressor was presented at 20 dB higher than the probe. Perhaps this level was not enough to extract the full emission.

It is possible that the presence of a high-level tone might produce SFOAE components that are not typically present and are produced by the high-level tone putting the amplification mechanisms into a nonlinear region that reflects energy and produces SFOAEs (Shera et al., 2004). Our data provide no evidence for such unintended SFOAE components for the near-frequency suppressor that we used. However, suppressors that are at much higher frequencies may, indeed, produce such unintended probe-frequency emissions (e.g. Charaziak and Siegel 2015; see Shera et al., 2004).

It has been suggested that a second tone (called a "suppressor") at a frequency two or more octaves above the SFOAE-evoking probe tone can suppress SFOAE components that originate near the CF region of the second tone, even though this region is far-basal of the probe-tone CF region (Charaziak, and Siegel, 2015). One reason this might happen is that there is amplification of reticular lamina motion in this region (Siegel, 2018) and the high-frequency "suppressor" tone would suppress this RL motion (Fallah et al., 2019; Dewey et al., 2019). In agreement with Goodman et al. (2020) the present data show that as the solution front continues basal past the probe-tone CF region there is no further reduction of the SFOAE, which indicates that no SFOAE is originating from this far-basal region. The present data extend the conclusions of Goodman et al. (2020) without needing any suppression by a high-level tone. These results are consistent with the cochlear amplification of RL motion at frequencies far below CF being local amplification or "non-propagating amplification" as described in another manuscript submitted to the 2022 Mechanics of Hearing Meeting (Guinan, 2022).

Acknowledgements

Research reported in this publication was supported by the National Institute of Deafness and Other Communication Disorders within the National Institutes of Health R01 DC014997 (JTL).

References

1. J. T. Lichtenhan, J. J. Hartsock, J.J., R. M. Gill, J. J. Guinan, Jr., A. N. Salt, The Auditory Nerve Overlapped Waveform (ANOW) Originates in the Cochlear Apex. J Assoc Res Otolaryngol **15**, 395-411 (2014).
2. J. T. Lichtenhan, J. J. Hartsock, J. R. Dornhoffer, K. M. Donovan, A. N. Salt, Drug Delivery Into the Cochlear Apex: Improved Control to Sequentially Affect Finely Spaced Regions Along the Entire Length of the Cochlear Spiral. J. Neurosci. Methods. **273**:201–209 (2016).
3. C. Lee, J. J. Guinan, Jr., M. A. Rutherford, W. A. Kaf, K. M. Kennedy, C. A. Buchman, A. N. Salt, J. T. Lichtenhan, Cochlear Compound Action Potentials from High-Level Tone Bursts Originate from Wide Cochlear Regions that are Offset Toward the Most Sensitive Cochlear Region. J. Neurophysiol. **121**, 1018-1033 (2019).
4. S. S. Goodman, C. Lee, J. J. Guinan, Jr., J. T. Lichtenhan, The Spatial Origins of Cochlear Amplification Assessed by Stimulus-Frequency Otoacoustic Emissions. Biophys. J. **118**, 1183-1195 (2020).
5. J. J. Guinan, Jr., S. M. Lefler, C. A. Buchman, S. S. Goodman, J. T. Lichtenhan, Altered Mapping of Sound Frequency to Cochlear Place in Ears with Endolymphatic Hydrops Provide Insight into the Pitch Anomaly of Diplacusis. Nature Sci Rep **11**, 10380 (2021).
6. S. Kakehata and J. Santos-Sacchi. Effects of Salicylate and Lanthanides on Outer Hair Cell Motility and Associated Gating Charge. J. Neurosci. **16**, 4881-4889 (1996).
7. D. H. Keefe, Double-Evoked Otoacoustic Emissions. I. Measurement Theory and Nonlinear Coherence. J. Acoust. Soc. Am. **103**, 3499-3508 (1998).
8. C. A. Shera, A. Tubis, C. L. Talmadge, J. J. Guinan, Jr., The Eual Effect of "Suppressor" Tones on Stimulus-Frequency Otoacoustic Emissions. Assoc. Res. Otolaryngol. Abstr. 27, Abs. 776 (2004).
9. K. K. Charaziak and J. H. Siegel, Tuning of SFOAEs Evoked by Low-Frequency Tones is Not Compatible with Localized Emission Generation. J. Assoc. Res. Otolaryngol. **16**, 317-329 (2015).
10. J. H. Siegel, (2018). "Does Reticular Lamina Active Gain Explain Broad Suppression Tuning of SFOAEs?" in *To the Ear and Back Again – Advances in Auditory Biophysics*, AIP Conference Proceedings 1965, 140003 (2018).
11. E. Fallah, C. E. Strimbu, E. S. Olson, Nonlinearity and Amplification in Cochlear Responses to Single and Multi-Tone Stimuli. Hear. Res. **377**, 271-281 (2019).
12. J. B. Dewey, B. E. Applegate, J. S. Oghalai, Amplification and Suppression of Traveling Waves along the Mouse Organ of Corti: Evidence for Spatial Variation in the Longitudinal Coupling of Outer Hair Cell-Generated Forces. J. Neurosci. **39**, 1805-1816 (2019).
13. Guinan, J.J., Jr., Traveling-Wave Amplification by Deiters-Cell Phasing of Reticular Lamina Motion So Outer Hair Cells Amplify Pressure. Submitted to the 2022 Mech. of Hearing meeting.

Suppression of Organ-of-Corti Vibrations and Otoacoustic Emissions in Mice

Karolina K. Charaziak[1a)]

[1]*Keck School of Medicine, Caruso Department of Otolaryngology, University of Southern California*
[a)] Corresponding author: charazia@usc.edu

Abstract. The inner ear response to a tone can be diminished in the presence of another tone through the process of suppression. Typically, the strongest suppression is expected when both tones, the probe and the suppressor, are of similar frequency and fall within the region of active amplification on the basilar membrane (BM). However, suppression can occur in otoacoustic emission (OAE) responses even when the suppressor tone is more than an octave above the evoking probe stimulus—a result that does not mimic BM behavior. The recent development of new imaging techniques (Optical Coherence Tomography, OCT) has brought an unexpected twist to these results: Whereas two-tone suppression is limited to the tonotopic peak region when measured at the BM, the same does not hold true when vibrations are measured in the organ of Corti (OoC). For instance, vibrations at the top of OoC are actively amplified and suppressed over a broad region corresponding to the tail of the traveling wave. The purpose of this study is to compare suppression patterns as measured in ear-canal pressure and organ of Corti vibrations. We measured cochlear vibrations and ear-canal pressure in mice in response to tones swept in frequency at narrow (1.1) and wide (2.1) suppressor-to-probe ratios. The OAE response (residual) is extracted by vector subtraction of ear-canal responses to probe-alone *vs.* probe swept together with the suppressor. Preliminary data reveal that OAE residuals can be detected for both narrow (1.1) and wide (2.1) ratios in mice. The OAE residuals obtained with wide ratios grow linearly with the stimulus. In contrast, the narrow-ratio OAE residuals tend to have more compressive growth. When residual responses are extracted from the vibrations of the organ of Corti, a similar pattern emerges: the wide-ratio residuals tend to have more linear growth compared to the narrow-ratio residuals. These patterns suggest that wide-ratio OAE residuals may correlate more closely with vibrations within the organ-of-Corti vibrations than with the BM.

INTRODUCTION

Otoacoustic emissions (OAEs) are sounds generated by the cochlea of the inner ear either spontaneously or in a response to a probe stimulus. Despite decades of research on OAEs, the mechanisms behind their generation and propagation out of the cochlea are still not fully understood. Intuitively, emissions evoked with a relatively simple stimulus—such as a pure tone—should be the easiest to interpret. The potential advantage of a single-tone (stimulus-frequency, SF) OAEs over emissions evoked with more complex stimuli, is also the source of their major drawback: Because SFOAE overlaps in time and frequency with the evoking tone, the two signals cannot be easily separated at the recording microphone. To overcome this difficulty, SFOAEs are typically measured indirectly, where the stimulus pressure is estimated *in situ* and subtracted from the mixture of the probe and OAE signals [1]. Here we use a secondary high-level tone to suppress the OAE signal and provide an *in situ* estimate of the stimulus pressure. The vector difference of the ear-canal pressure measured in response to probe-alone and probe presented concomitantly with the suppressor represents the SFOAE residual. Only when the suppression is complete, the SFOAE residual would be representative of the total SFOAE evoked by the probe tone [2].

Typically, SFOAE is extracted with a suppressor tone (f_s) at a frequency close to the probe frequency (i.e., narrow f_s/f_p ratio). In such a case, the SFOAE magnitude grows compressively with increasing probe level, and SFOAE phase-gradient group delays are longer than the cochlear travel time to the probe's best frequency (BF) place in the cochlea. These characteristics agree well with the predictions of coherent reflection theory of OAE generation and propagation in the cochlea, where the active basilar membrane (BM) vibrations in the peak-region of traveling wave are the primary source of the emission signal [3, 4]. However, SFOAEs can be also measured in a

condition where suppressor frequency is much higher than the probe (i.e., wide f_s/f_p ratio), so that suppressor excitation expands basally to the region of active amplification of the probe response on the BM (i.e., the passive "tail region" [5]). These wide ratio SFOAEs tend to display linear growth with increasing the probe intensity, and have phase-gradient group delays that are shorter than the cochlear travel time to the probe BF place [6]. These findings are unexpected because there is typically little suppression seen in the BM vibrations for suppressor frequencies over an octave above the probe frequency [7]. Several hypotheses have emerged regarding the generation mechanism of the wide-ratio OAEs based on our knowledge of the BM vibrations [8–10]. However, recent advances in imaging technology have revealed that the *in vivo* organ of Corti (OoC) does not simply vibrate as a rigid body sitting atop of the BM [11]. Most importantly, the motions of OoC, as measured near the location of the outer hair cells (OHCs), are actively amplified in both the tail and peak regions of the traveling wave. While the OoC amplification in the tail region is of a smaller magnitude, and depends less on the stimulus intensity (i.e., linear growth) than amplification near peak region (the BF place of the probe), it has been shown that these vibrations are suppressible by a secondary tone near its own tonotopic place [12]. In the current study we compare two-tone suppression of OoC vibration with SFOAE measured using wide and narrow f_s/f_p in mice.

METHODS

The vibratory and OAE responses were measured in 4–8 weeks old mice (CBA/CaJ) anesthetized with ketamine (80–100 mg/kg) and xylazine (5–10 mg/kg). The left ear-canal was resected and a tip of an acoustic probe (containing a microphone and speakers; ER10X, Etymotic Research) was placed a few millimeters from the eardrum. At the end of the experiment the animals were euthanized, and some tests were repeated *postmortem*. All procedures were approved by the Institutional Animal Care and Use Committee at USC.

We utilized custom-built volumetric optical coherence tomography and vibrometry (VOCTV) system to image the apical turn of an intact cochlea directly through the otic capsule (Fig. 1a) and to record intracochlear motions in response to acoustical stimulation, as described in detail elsewhere [12–14]. First, a cross-sectional image of the apical turns was obtained (B-scan) and voxels with high reflectivity at the BM and near center of the OoC mass (in the vicinity of the OHCs) were chosen for vibrometry (Fig. 1a). Typically, the angle between the light source and the BM was ~50–65º meaning that the measured vibrations capture both transverse and radial motions. All measurements were carried out in an electrically shielded sound-attenuating booth with the animal placed on the vibration isolation table.

The acoustical stimuli were designed digitally (100–kHz sampling rate) and delivered through one of the sound sources of the ER10X probe. The stimuli were calibrated *in situ* to produce desired SPL and zero phase at the probe microphone. A swept frequency stimulus (100-ms duration, plus 2-ms, cosine-shaped onset/offset ramps, and 5-ms zero padding at the offset) was presented in three configurations: probe alone (P_p), probe plus suppressor (P_{ps+}), and probe plus suppressor with inverted polarity (P_{ps-}). The probe frequency was swept exponentially from 0.5 to 20 kHz. The suppressor frequency was set at constant ratio to the probe frequency of either 1.1 (narrow) or 2.1 (wide). For brevity, we will refer to any measurements collected with either narrow or wide f_s/f_p ratio simply as 1.1 or 2.1 data. The suppressor level was fixed at 75 dB SPL, while the probe level was varied from 40 to 70 dB SPL in 10 dB steps. The intracochlear responses from BM and OoC regions to the three-block stimulus were synchronously averaged 200–400 times. The OAE responses were calculated as a global average of the ear-canal pressure measured in a response to the same stimulus condition to improve the signal to noise ratio. For all responses, raw data were grouped into 10 subaverages, that were used to calculate the Rayleigh statistic at each frequency [15]. Only data points that displayed significant phase coherence ($p < 0.01$) were included in the analyses.

Each response buffer was convolved with a time-reversed and scaled version of the probe stimulus, which effectively time-compresses the response [16, 17]. The first order impulse response was extracted with a time window (100^{th} order recursive exponential) with a half-width cutoff of 0 and 1.2 ms re zero time (i.e., the centroid of the time-compressed probe stimulus). The impulse responses were then band-pass filtered (2–16 kHz Kaiser window) and converted to a frequency domain (i.e., transfer function) via FFT. Thus, all data are displayed as normalized to the stimulus pressure (gain).

The residual response was calculated as a vector difference between the cochlear impulse response to a probe alone condition and in the presence of the suppressor [the response to the suppressor sweep was canceled out by averaging the two suppressor "on" buffers, i.e., $(P_{ps+} + P_{ps-})/2$]. The suppression residual represents the vector change in the probe-alone response due to the presence of the suppressor tone.

RESULTS

First, we discuss the intracochlear vibrations measured in response to the probe-alone stimulus (Fig. 1), followed by the analyses of the vibratory residual responses (Fig. 2), and finally we compare the SFOAE residual characteristics with the vibratory residuals (Fig. 3).

Examples of transfer functions measured at the level of OoC and BM are shown in Fig. 1b and c, respectively. Nonlinear amplification was seen in both BM and OoC vibrations for frequencies in the vicinity of the characteristic frequency (CF; i.e., BF at 40 dB SPL). At lower frequencies (<1 octave re CF) the BM responses (c) were typically linear and passive (i.e., there was no difference between gain measured before and after death). In the OoC region (b), however, the low-frequency vibrations were actively amplified as compared to the *post mortem* recordings (colored *vs*. gray), but the amplification did not vary with the probe level (linear gain). These results are consistent with the view that both the peak (~CF) and the tail regions (<< CF) of the traveling wave are actively amplified at the level of the OHCs, but the OHC-driven amplification is coupled to the BM only in the peak region [12].

The average gain at the BF (solid) and at the BF/2.1 (dotted) across stimulus levels is shown in Fig. 1d for the BM (blue) and OoC (red) vibrations. For both the BM and the OoC regions the gain and the phase-gradient group delays at the BF decreased with increasing stimulus level (d, f solid). The slope of the gain at the BF was significantly different from 0 dB/dB (e; solid lines), consistent with a compressive nonlinearity. At the BF/2.1, cochlear gain (d; dotted lines) varied relatively little with the stimulus level (slope of gain ~0 dB/dB, e; dotted lines).

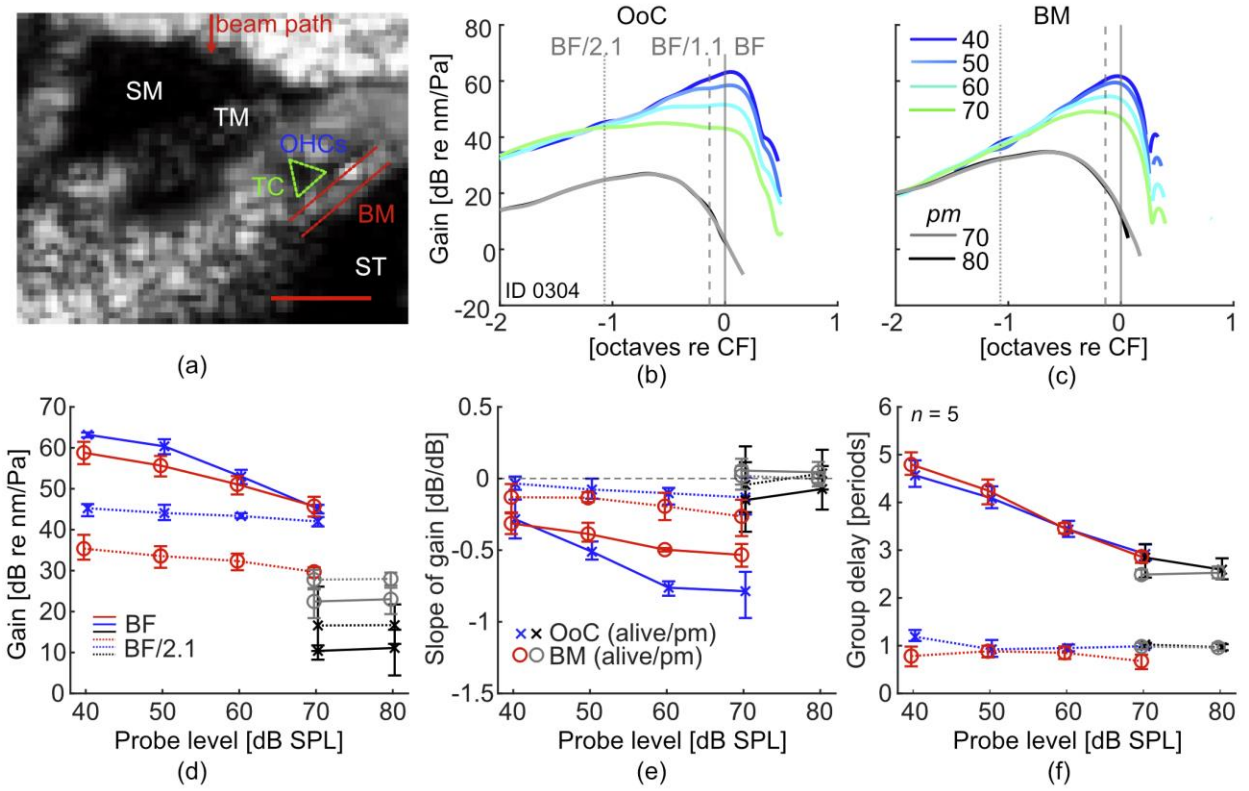


FIGURE 1. Cochlear responses to the probe-alone sweeps of varying levels. B-scan through an apical turn of a mouse cochlea (a; the horizontal scale bar: 100 μm). For vibratory measurements we chose brightest points in the center of OHCs and points directly below it (at the BM level). Panels b and c depict examples of transfer functions measured in before (color) and after (gray scale) death at varying probe levels [see legend in (c)] at the level of OoC and BM, respectively. The CF (9.3 kHz) represent the average BF for 40-dB SPL responses at the level of BM, OoC and tectorial membrane (TM, data not shown). The vertical lines mark probe frequency equal to an average BF, BF/1.1 and BF/2.1 for a 40-dB SPL probe (recall that BF = CF at 40 dB SPL; BF unlike CF, changes with the probe level). Panels d–f show the average (n = 5) transfer function parameters (gain and its slope *vs.* probe level, as well as phase-gradient group delay) extracted at either BF (solid) or at BF/2.1 (dotted) for OoC (blue crosses) and BM (red circles) vibrations at varying probe levels (x-axis). Corresponding responses collected *post mortem* are shown in black/gray. The error bars represent 95% confidence intervals estimated through a block (by animal ID) bootstrapping. SM: scala media; TC: tunnel of Corti; ST: scala tympani.

At the OoC level, but not at the BM, the gain at BF/2.1 was significantly above the gain measured *post mortem* (d, blue dotted vs. black dotted). At BF/2.1 the group delays for either BM or OoC vibrations were short as compared to the delays at the BF (f; dotted vs. solid) and they did not vary with the probe level.

When the probe stimulus was presented together with a suppressor at a narrow f_s/f_p ratio, both the OoC and BM vibrations were reduced, particularly so in the peak region (Fig. 2a, b, solid vs. dashed) [7]. In contrast, very little change in the gain at the peak was seen when the suppressor was swept at the wide f_s/f_p ratio (c, d). However, when the suppressor frequency was at/below the BF of the measurement location (dotted vertical line in c, d), a large reduction in the OoC, but not BM, vibrations was observed [12].

Typically, suppression of cochlear vibrations is quantified separately for the magnitude and phase of the response. Here we adapt an approach used for SFOAE measurements, where a residual response is calculated as a vector difference between the probe-alone and probe-with-suppressor responses [18]. The residual is a vector, that reflects changes in both magnitude and phase of the probe response due to the presence of the suppressor tone. The residual magnitude spectra calculated from data showed in Fig. 2a–d, are shown in panels e and f for the OoC and BM regions, respectively. The 1.1 ratio residuals had a band-pass shape centered at the BF (e, f, red). The 2.1 residuals were of smaller magnitude at the BF (blue). However, at the BF/2.1 the wide ratio residuals were larger than narrow ratio residuals particularly so for the OoC vibrations (e, see dotted vertical line).

The average parameters of the residual transfer functions as extracted at the BF (for the 1.1 ratio condition) or at BF/2.1 (for the 2.1 condition) are shown in Fig. 2g–l. For the BM data, the wide ratio residuals were often below the noise criteria at the BF/2.1 frequency and thus we excluded them from averaging. For comparison, the mean probe-alone responses are shown in solid lines (as in Fig. 1d–f). Overall, the residual parameters reflected well the probe-alone response parameters, as long as they were both evaluated at/near the suppressor frequency. This finding indicates that the suppression method for extracting SFOAEs does not limit the OAE residual's ability to estimate intracochlear vibrations in response to the probe-alone condition.

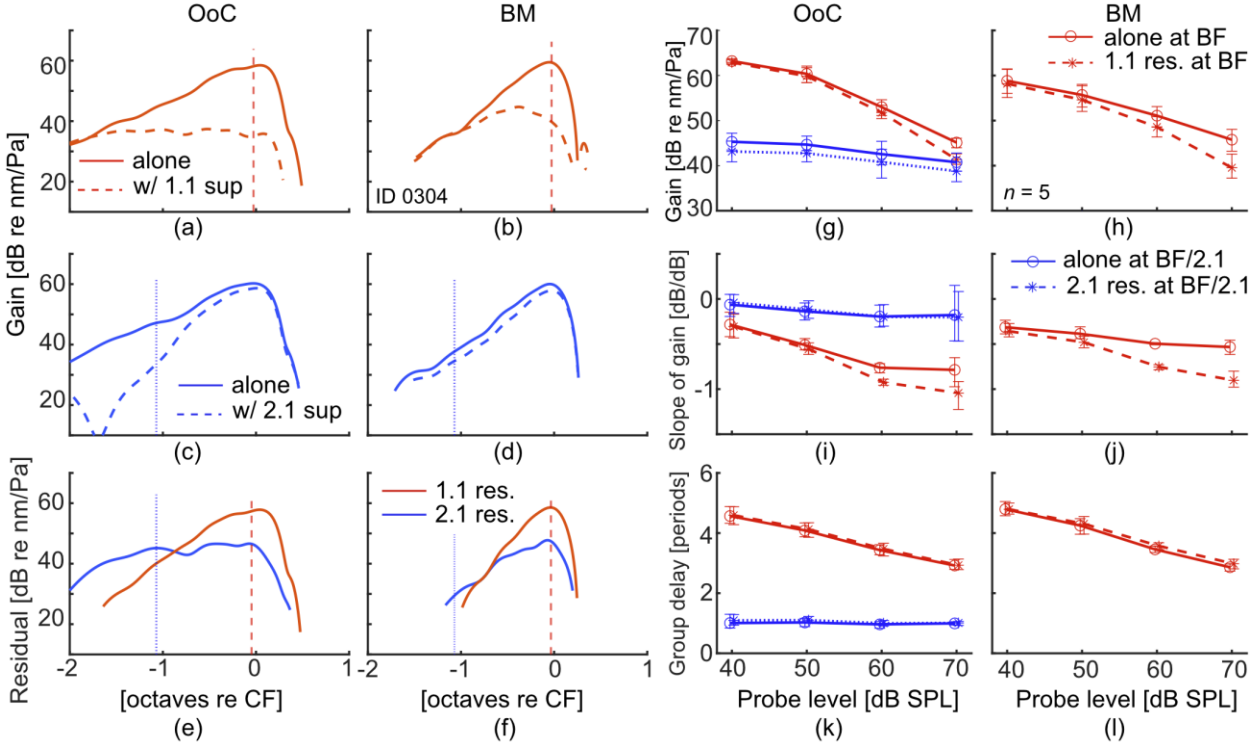


FIGURE 2. Suppression of cochlear vibrations. The OoC and BM transfer functions measured in response to a 50-dB SPL probe alone (solid) and in presence of the 75-dB SPL suppressor (dashed) are shown in panels (a-b) and (c-d) for the 1.1 (red) and 2.1 (blue) suppressor ratios, respectively. The corresponding residual transfer functions are shown in panels e and f. The vertical red and blue lines mark the condition at which the suppressor frequency matched the BF and the BF/2.1, respectively. The mean 1.1 (red dashed) and 2.1 (blue dotted) residual transfer function parameters are shown in panels g–l together with the parameters extracted directly from probe alone responses (solid; as in Fig. 1). For the BM vibrations, the residual magnitude was not measurable in all animals at the frequency corresponding to BF/2.1, thus these data are not shown.

The SFOAE resiudal measured for the same animal as in Figs. 1, 2 are displayed in Fig. 3a–f. The 1.1 ratio SFOAE gain decreased with increasing stimulus level (c; dark to light colors) indicative of compressive nonlinearity. The phase-gradient group delays were long and increased with increasing probe frequency (e). In contrast, the 2.1 ratio SFOAEs had more linear growth (d; overlapping lines) and relatively short delays that did not change considerably with the probe frequency and intensity (f). In most of the animals, the strongest SFOAE residuals were present at low to mid frequencies (~5–10 kHz) for either ratio condition. Because cochlear amplification and delays vary with frequency, it is most meaningful to compare SFOAE at a frequency that is related to the tonotopic place from which the vibratory responses were acquired (CF ranged from 7.8 to 9.3 kHz; mean 8.5 kHz). Thus, for the 1.1 ratio condition, we extracted SFOAE parameters at peak frequencies close to average CF (from 8.2 to 10.3; mean 9.4 kHz). The 2.1 SFOAEs residual parameters were evaluated at peak frequencies close to the average BF/2.1 (from 5 to 5.5; mean 5.3 kHz). Because at the lower frequencies the SFOAE residuals were generally stronger (Fig. 3c, d), the average 2.1 SFOAEs (near 5.3 kHz) were of much higher magnitude than the average 1.1 SFOAEs (near 9.4 kHz; Fig. 3g).

The average 2.1 SFOAE gain did not vary with the stimulus level (Fig. 2g; blue), indicative of a linear growth of SFOAE magnitude with level. Recall that linear growth was also observed for the 2.1. OoC residuals at BF/2.1 frequency (i.e., ~0 dB/dB slope; Fig. 3h solid vs. dotted blue lines). Similarly, the 2.1 SFOAE group delays were invariant across stimulus levels (i, solid blue), and were on average 1.9 times longer than the group delays of the cochlear vibrations measured at the BF/2.1 (dotted blue). (Recall that there was no difference between BM and OoC delays at BF/2.1; Fig 1f, dotted blue vs. red). When compared to cochlear vibrations delays at the BF (Fig. 1f, solid),

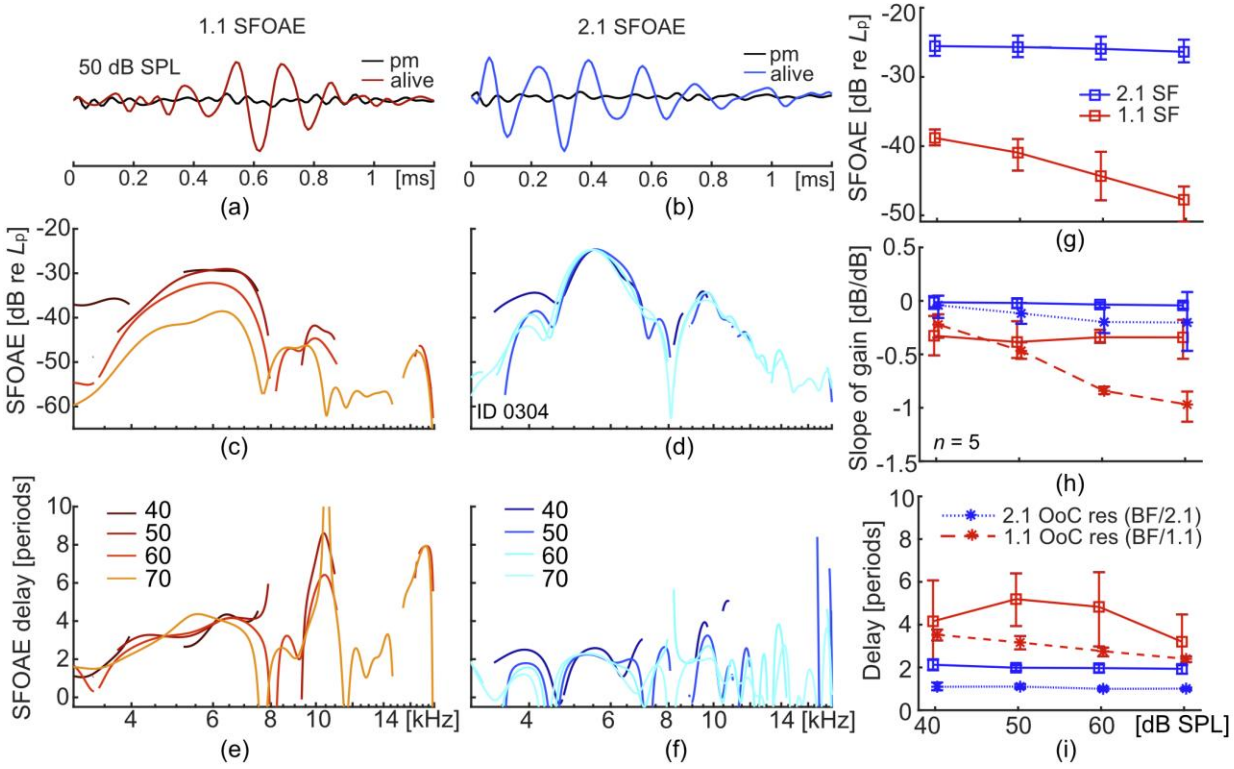


FIGURE 3. Residual SFOAE transfer functions. Panels a and b show the time-domain representations of 1.1 and 2.1 SFOAE residuals measured in response to a 50-dB SPL probe before (color) and after death (black) in one animal. The SFOAE magnitudes (re stimulus level) and phase-gradient group delays measured for 40 to 70 dB SPL probe levels are shown in c and e for the 1.1 ratio and in d and f for the 2.1 ratio condition for the same animal. Only data points meeting Rayleigh criterium are shown. In panel g the average SFOAE gain for the 1.1 ratio (at high frequencies; ~9.4 kHz) and for the 2.1 ratio (at low frequencies; ~5.3 kHz) are shown in blue and red solid lines, respectively. In panels h and i the corresponding slope of gain and group delays are shown together with average parameters extracted from OoC residual spectra at either BF/1.1 for the 1.1 ratio condition (red dashed) or at BF/2.1 for 2.1 ratio condition (dotted blue). For narrow ratio data at BF/1.1 instead of at BF (Fig. 1) are shown. While there was no significant difference between slope of gain as extracted at BF and BF/1.1, the phase-gradients group delays evaluated at BF/1.1 were shorter than group delays extracted at BF (with the mean ratio of between the two of 0.8).

the 2.1 SFOAE delays were shorter, with the average ratio between the two of ~0.6. Thus, the wide ratio SFOAEs have delays shorter than the cochlear travel time to the probe tonotopic place (BF), but longer than the cochlear travel time to the suppressor tonotopic place (BF/2.1).

The average 1.1 SFOAE gain decreased with increasing the stimulus level (Fig. 3g, red), but at a slower rate than changes in the intracochlear gain near the BF (h, solid vs. dashed red). The 1.1 SFOAE group delays were typically longer than the cochlear delays at the BF (Fig. 1f, solid lines) with the mean ratio of 1.2. When 1.1 SFOAE delays were compared to the cochlear delays extracted at the suppressor tonotopic place (BF/1.1; Fig. 3i, red dashed), the ratio increased to 1.5. The fact that the 1.1 SFOAE delays are longer than the forward travel time to either probe or suppressor tonotopic place, agrees with the theory that OAEs escape the cochlea via slow pressure waves in cochlear fluids [3, 4].

DISCUSSION

The aim of this study was to evaluate the suppression of intracochlear vibrations in context of SFOAE measurements. In this preliminary report we measured vibrations from the OHC-region of OoC and from the BM (Fig. 1a) in response to a swept-tone stimulus presented in isolation or together with a suppressor tone. The ratio of the probe tone to suppressor tone frequency was kept constant at either 1.1 (narrow) or 2.1 (wide) condition. The 1.1 suppressor depressed cochlear nonlinear gain in the "peak region" (Fig. 2a, b). In contrast, an octave higher suppressor had relatively little effect on the cochlear gain at the peak (Fig. 2c, d), but it considerably affected the tail of the OoC vibrations, that unlike the BM motions, are actively amplified (Fig. 1b, c) [12].

To compare the suppression of cochlear vibrations to OAE responses, we calculated the residual response, which represents any change (i.e., both in amplitude and/or phase) in the probe tone response due to presence of the suppressor (Fig. 2e, f). The residual approach is used for SFOAE measurement, as unlike in case of the vibratory responses, the stimulus and OAE pressures cannot be measured separately. It is generally assumed that the residual SFOAE response approximates well the total emission evoked by the probe tone. We evaluate that claim indirectly by comparing the residual intracochlear gain to gain measured in response to the probe presented by itself. The spectral characteristic of the residual transfer function correlate well with the features of the probe-alone response, if both are evaluated near the frequency of the suppressor tone (i.e., at BF or BF/1.1 for the narrow ratio, and at BF/2.1 for the wide ratio; Fig. 2g–l). Small differences between residual and probe-alone gains were observed at higher stimulus levels (e.g., 60–70 dB SPL; g, h), likely due to incomplete suppression by a 75-dB SPL suppressor tone. As a result, the slope of the gain (i.e., the strength of the compressive nonlinearity) tended to be higher for the residual vs. probe alone data (i, j). Thus, the residual method on its own does not appear to be a limiting factor when trying to elucidate a cochlear response to a probe tone, at least for most of the conditions used in this study.

The SFOAE gain extracted for 1.1 suppressor ratio decreased with increasing stimulus level at mid to high intensities, as expected if these OAEs originated in the peak region of the traveling wave [3]. However, the strength of the compressive nonlinearity of the 1.1 SFOAEs at higher probe levels was less than measured directly in the cochlear vibrations near the BF (Fig. 3h, red solid vs. dashed). Perhaps it is not entirely correct to compare growth of OAE response that likely originates over a region of the cochlea, with a growth of cochlear vibrations measured at a single location. Because of the spread of cochlear excitation at higher probe levels, the collective SFOAE may reflect the average gain in the excited region rather than at a single location. The delays of 1.1 SFOAEs at BF were quite variable (likely due to relatively low signal to noise ratio), but on average they tended to decrease with increasing the probe level (Fig. 3i, solid red). The 1.1 SFOAE delays were also longer than the delays at either BF or BF/1.1 by a factor of 1.2 and 1.5, respectively. These results altogether fit the theory of SFOAEs being generated (at least in part) in the peak region of the traveling wave [3].

The 2.1 SFOAEs displayed linear growth, and short group delays that varied little with either probe frequency or level (Fig. 2d, e, i). These features of the 2.1 SFOAE residuals agree well with vibrations measured in the tail region of the traveling wave, where amplification at the level of OoC is linear, and delays are short. The delay of the 2.1 SFOAE residual is also longer than OoC delays at the suppressor tonotopic place (BF/2.1), as expected if most of the OAE signal originated in that region and escaped the cochlea via a slow pressure wave. Although not the only interpretation of these results, the simplest is that the SFOAE residual originates near the tonotopic place of the suppressor tone and might therefore be used to gauge cochlear mechanics at that location, at least at low to mid probe intensities. While the characteristics of the 2.1 SFOAE residuals match well the characteristics of the OoC vibrations at BF/2.1, it remains to be understood how these motions stimulate the cochlear fluids to give rise to a backward OAE wave.

ACKNOWLEDGMENTS

Supported by grant R00 DC016906 from the NIH and by USC. We thank Drs. Christopher Shera and Alessandro Altoè for help with this ongoing project.

REFERENCES

1. Kalluri R, Shera CA (2007) Comparing stimulus-frequency otoacoustic emissions measured by compression, suppression, and spectral smoothing. J Acoust Soc Am 122:3562–3575
2. Siegel JH (2006) The biophysical origin of otoacoustic emissions. Auditory Mechanisms: Processes and Models 361–368
3. Zweig G, Shera CA (1995) The origin of periodicity in the spectrum of evoked otoacoustic emissions. J Acoust Soc Am 98:2018–2047
4. Shera CA, Guinan JJ (2003) Stimulus-frequency-emission group delay: A test of coherent reflection filtering and a window on cochlear tuning. J Acoust Soc Am 113:2762–2772
5. Guinan JJ (1990) Changes in stimulus frequency otoacoustic emissions produced by two-tone suppression and efferent stimulation in cats. In: Dallos P, Geisler CD, Matthews JW, Ruggero MA, Steele CR (eds) The Mechanics and Biophysics of Hearing. Springer-Verlag, Madison, pp 170–177
6. Siegel JH (2006) Local and Propagating Components of Stimulus Frequency Otoacoustic Emissions. Assoc. Res. Otolaryngol, Abstr.: 976?
7. Ruggero MA, Robles L, Rich NC (1992) Two-tone suppression in the basilar membrane of the cochlea: mechanical basis of auditory-nerve rate suppression. J Neurophysiol 68:1087–1099
8. Shera C, Tubis A, Talmadge C, Guinan JJ (2004) The dual effect of "suppressor" tones on stimulus-frequency otoacoustic emissions. In: Assoc. Res. Otolaryngol., Abstr. 27, 538. p 3687
9. Siegel JH, Cerka AJ, Recio-Spinoso A, Temchin AN, van Dijk P, Ruggero MA (2005) Delays of stimulus-frequency otoacoustic emissions and cochlear vibrations contradict the theory of coherent reflection filtering. J Acoust Soc Am 118:2434–2443
10. Charaziak KK, Siegel JH (2015) Tuning of SFOAEs evoked by low-frequency tones is not compatible with localized emission generation. J Assoc Res Otolaryngol 16:317–329
11. Gao SS, Wang R, Raphael PD, Moayedi Y, Groves AK, Zuo J, Applegate BE, Oghalai JS (2014) Vibration of the organ of Corti within the cochlear apex in mice. Journal of Neurophysiology 112:1192–1204
12. Dewey JB, Applegate BE, Oghalai JS (2019) Amplification and suppression of traveling waves along the mouse organ of Corti: evidence for spatial variation in the longitudinal coupling of outer hair cell-generated forces. J Neurosci 2608–2618
13. Gao SS, Wang R, Raphael PD, Moayedi Y, Groves AK, Zuo J, Applegate BE, Oghalai JS (2014) Vibration of the organ of Corti within the cochlear apex in mice. J Neurophys 112:1192–1204
14. Lee HY, Raphael PD, Xia A, Kim J, Grillet N, Applegate BE, Ellerbee Bowden AK, Oghalai JS (2016) Two-dimensional cochlear micromechanics measured in vivo demonstrate radial tuning within the mouse organ of Corti. J Neurosci 36:8160–8173
15. Versteegh CPC, van der Heijden M (2012) Basilar membrane responses to tones and tone complexes: Nonlinear effects of stimulus intensity. J Assoc Res Otolaryngol 13:785–798
16. Farina A (2000) Simultaneous measurement of impulse response and distortion with a swept-sine technique. 108th AES Convention
17. Bennett CL, Özdamar Ö (2010) Swept-tone transient-evoked otoacoustic emissions. J Acoust Soc Am 128:1833–1844
18. Brass D, Kemp DT (1993) Suppression of stimulus frequency otoacoustic emissions. J Acoust Soc Am 93:920–39

Imaging the ratio dependence of distortion products at apical regions of the gerbil cochlea

Wei Dong[1, 2, a)] and Sebastiaan W.F. Meenderink[1]

[1)] *VA Loma Linda Healthcare System,*
Loma Linda,
CA,
92374
[2)] *Department of Otolaryngology – Head and Neck Surgery,*
Loma Linda University Health,
Loma Linda,
CA 92350

(Dated: 29 April 2022)

Abstract. Distortion product otoacoustic emissions (DPOAEs) exhibit features at low frequencies that are distinct from high-frequency ones in that they are less ratio-dependent: their characteristics (amplitude and phase) vary little with the frequency separation of the primary tones. The underlying mechanism may arise from the unique apical cochlear mechanics that are poorly understood. Intracochlear observations provide straightforward evidence for the generation mechanism of DPOAEs. In the current study, two-tone induced distortion products (DPs) were recorded from the apically located organ of Corti complex [OCC includes the organ of Corti, the basilar membrane, and the tectorial membrane] of alive gerbil cochleae using optical coherence tomography (OCT) together with simultaneously measured ear canal DPOAEs. Within the OCC, DPs were robust and recorded from all key structures. The ratio-dependency of DPs appeared to link to the characteristics of the primary tones that were broadly tuned and exhibited nonlinearity in a broad region. The upper sideband, $2f_2 - f_1$ DP was found ratio independent due to the invariant overlapping region of the primary tones. The lower sideband, $2f_1 - f_2$ DP differed, and varied with f_2/f_1 ratio. It included both locally and non-locally (i.e. traveling) DP components. Thus, their correlation to related DPOAEs was not straightforward. Our observation provides insights into the generation of low-frequency DPOAEs and related cochlear mechanics, which helps a more precise DPOAE applications in detecting sensory auditory damage in the clinic.

INTRODUCTION

Normal cochlear operation relies on active nonlinear outer hair cell (OHC) processes that compress mechanical responses into a manageable dynamic range, but which also introduces distortions in the cochlea's response [**?**]. As a consequence, the response to a two-tone stimulus with frequencies f_1 and f_2 (ratio $f_2/f_1 > 1$) contains—in addition to these primary frequencies— a family of intermodulation tones at frequencies $f_1 - n(f_2 - f_1)$ and $f_2 + n(f_2 - f_1)$, with $n \in \mathbb{N}$. The intracochlear locations that generate these distortion products (DPs) depend on: (1) the overlap in the responses to the two stimulus tones [**?**], and (2) the extent of the nonlinearity within the cochlea. After generation, DPs go through similar signal processing as externally presented tones (i.e. they may propagate, get amplified and/or filtered). As a consequence, DPs measured at a certain intracochlear location are a mixture of local and non-local (i.e., traveling) components [**? ?**].

Based on the tuned responses observed at the basilar membrane (BM) [**?**], varying f_2/f_1 ratio is believed to modify the spatially distributed, region of DP generation such that larger f_2/f_1 ratios result in less overlap of the primaries' responses and smaller DPs. Observations of BM responses support this notion in that larger DPs are generated with (narrow) f_2/f_1 ratios close to unity [**? ?**]. However, recent intracochlear observations in the outer hair cell/reticular laminar (OHC/RL) region, both at high-frequency basal [**?**] and low-frequency apical regions [**?**], have demonstrated broader tuning and a more basally extended nonlinear region compared to the BM response. With this, DPs originate from a broader region within the cochlea, as confirmed by robust DPs measured at the RL region [**?**], and argues against a strong f_2/f_1 ratio dependence of DP generation by the OHC.

Once generated, DPs may propagate back to the ear canal, where they can be recorded as distortion product otoacoustic emissions (DPOAEs). These sounds are the sum of DP wavelets that arise from a spatially distributed generation region, which includes the regions of nonlinear distortion generation and linear, coherent reflection [**?**].

[a)] Electronic mail: wei.dong@va.gov

DPOAEs vary with f_2/f_1 ratio differently depending on their frequency in that low-frequency DPOAEs are less sensitive to f_2/f_1 ratio than high-frequency ones [**?**]. This presumably relates to differences in the micromechanics along the length of the cochlear duct, where apical, low-frequency responses differ from basal, high-frequency ones.

To address the (extent of) low-frequency DP generation and their contributions to DPOAEs, we measured two-tone-evoked vibratory responses in the organ of Corti complex (OCC, which includes the organ of Corti, the BM, and the tectorial membrane) from the second turn of the gerbil cochlea using optical coherence tomography (OCT). We used the OCC vibration patterns at both DP and stimulus frequencies to illustrate how the OHC motion affects the vibrations of surrounding structures in the OCC. Furthermore, we use measured DP phase data and predicted DP amplitudes (calculated from the Taylor series expansion of the measured primary responses) to show that the measured DPs consist of both local and non-local components, depending on the stimulus frequencies relative to the measurement location's best frequency (BF).

MATERIALS AND METHODS

Animal preparation

Cochlear vibrations were measured from the left ear in adult, female Mongolian gerbils (*M. unguiculatus*, $n = 12$). Gerbils were anesthetized with a ketamine/xylazine cocktail (80 and 10 mg/kg, respectively, intraperitoneal injection) with supplemental doses given as needed to maintain areflexia. Core temperature was kept at $\sim 38^{\circ}$C. Animals were tracheotomized, but not actively ventilated. The pinna and the cartilaginous ear canal were resected, and a probe containing a microphone and transducer assembly (ER-10X, Etymotic Research) was placed within a few mm of the tympanic membrane. A ventrolateral surgical approach was used to expose the cochlea. Animals were not allowed to recover from anesthesia and were euthanized by anesthetic overdose at the end of the experiment. The care and use of animals were in accordance with guidelines of, and approved by, the Institutional Animal Care and Use Committee (IACUC) of the VA Loma Linda Healthcare System.

Optical coherence tomography and vibrometry

A spectral domain OCT system (Thorlabs Telesto III TEL321C1 with an LSM04 objective) that operated with a central wavelength of 1310 nm (bandwidth: 170 nm) was used to noninvasively image through the cochlear bone and record the vibratory responses in the middle turn of the exposed, but otherwise structurally uncompromised, cochlea. Acquisition of optical spectra was triggered by externally generated TTL pulses (27.9 pulses/ms) that were synchronized to the stimulus generation and microphone acquisition system (RX6: Tucker Davies Technologies).These spectra were converted into depth-resolved, axial information (A-line) using Fourier analysis. Cross-sectional intensity images (B-scans) of the cochlea were constructed from multiple A-lines by scanning the OCT beam perpendicular to the tonotopic axis (e.g., Fig. 2A). To improve the quality of these images, the scan path was traversed multiple ($>$10) times and the intensity profiles for each A-line averaged. For vibration measurements (M-scans), the axial phase information of the A-lines was used. Here, the OCT beam was kept at a fixed position, and a series of time-stamped A-lines was recorded while an acoustic stimulus was presented to the ear.

Acoustic stimulation and analysis

Acoustic stimuli consisted of two equal-intensity (L_1=L_2) pure tones with frequencies f_1 and f_2 ($f_2/f_1 > 1$). The f_2 varied from 0.1 to 5 kHz in 0.1 kHz steps, while keeping the f_2/f_1 ratio fixed ($f_2/f_1 = 1.05 \ldots 1.6$). Each stimulus was repeated multiple times, and the average microphone and OCT response were used to extract the amplitude and phase at the stimulus frequencies and the DP frequencies ($2f_1 - f_2$ and $2f_2 - f_1$) using Fourier analysis. No artifact rejection was employed. The noise floor was estimated by first averaging the responses in two sub-buffers, and then analyzing the difference between these two buffers. The DP (and DPOAE) phases were normalized to the primary phases. That is, for DP at $f_1 + n(f_2 - f_1)$, the reported phase is: $\Phi_{dp} = \phi_{dp} - \phi_1 + n(\phi_2 - \phi_1)$, where $\phi \ldots$ represent the phase of a frequency component as extracted from the recorded signal. A phase of 180° (i.e., 0.5 cycle) suggests that the measured DP was generated locally (at the measurement site) by a simple, compressive nonlinearity. Any phase

departure from 180° then suggests that the measured DP was dominated by a non-local component that propagated to the recording site. Besides this phase characteristic to determine local vs non-local DP generation, we also predicted the DP amplitude based on the Taylor series expansion of the measured primary responses. For a locally generated DP, we expect that this prediction matches the measured DP. Custom software, written in MATLAB, was used for stimulus generation, hardware synchronization, and all signal analysis; control of the OCT system was implemented in C♯.

RESULTS

We demonstrate the f_2/f_1 ratio dependence of low-frequency DPOAEs and their intracochlear contributors (DPs) that were obtained simultaneously from intact cochleae of living gerbils. We observed robust DPs within the OCC and DPOAEs in the ear canal over a range of different f_2/f_1 ratios. In the current report, the findings are illustrated using the two largest sideband DPs at frequencies $2f_1 - f_2$ and $2f_2 - f_1$, respectively. Responses from one representative animal are shown since other animals showed similar results. We show the f_2/f_1 ratio dependence of DPs both using 2D vibrometry maps for an A-line that typically transected the lateral compartment (LC), the OHC, and the outer pillar cell (OPC) regions within the OCC (red line in Figs. 2 and 4), and for (1D) responses for each of these three locations (Figs. 3 and 5).

Low-frequency DPOAEs show little dependence on f_2/f_1

Figure 1 illustrates how DPOAEs vary with the separation of the primary tones, the f_2/f_1 ratio, in alive gerbil. The amplitude and phase of the $2f_1 - f_2$ DPOAEs at frequencies below 1 kHz were invariant with f_2/f_1 ratio, and only systematically varied with f_2/f_1 at higher DPOAE frequencies (Figs 1A, C). A clear optimal f_2/f_1 ratio (i.e. maximizing DPOAE amplitude) could not be determined. The $2f_2 - f_1$ DPOAEs showed even less of a dependence on f_2/f_1 ratio over the frequency range tested (Figs. 1B, D). These characteristics are consistent with observations in humans [**?**] and raise questions regarding the physiological site of their origin.

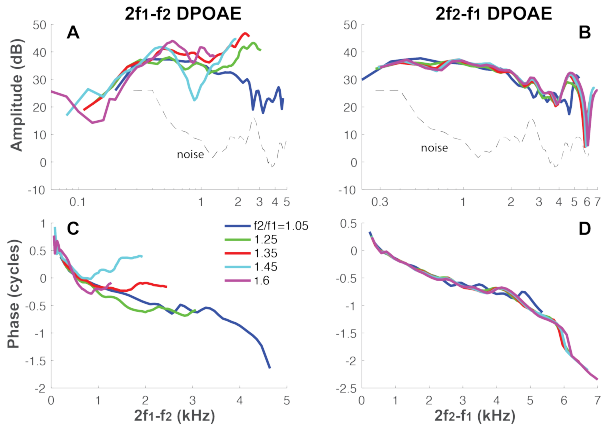


FIGURE 1. The f_2/f_1 ratio dependence of $2f_1 - f_2$ and $2f_2 - f_1$ DPOAEs. (A) Amplitude and **(C)** corresponding phase for the lower sideband DPOAE at $2f_1 - f_2$ for different f_2/f_1 ratios. Panels **(B)** and **(D)** are for the upper sideband DPOAEs at $2f_2 - f_1$. Parameters: L_1=L_2=70 dB, f_2/f_1=1.05 to 1.6, data from animal LL_08132020.

The high-frequency sideband DP @ $2f_2 - f_1$ is ratio independent

When viewing the second cochlear turn (where the BF is around 2.5 kHz) through the dense cochlear bony wall, the reflectance and spatial resolution of the OCT B-scan images are sufficient to provide high-resolution anatomical (Fig. 2A) and functional evaluations in living gerbil cochleae [**?**]. We identified several key structural components of the

OCC as indicated in the OCT image: the tunnel of Corti (ToC), the OHC, the BM, and the LC containing Hensen's cells and tectal cells. The latter cells are connected to the apical surface of the OHC (the RL) via their finger-like extensions.

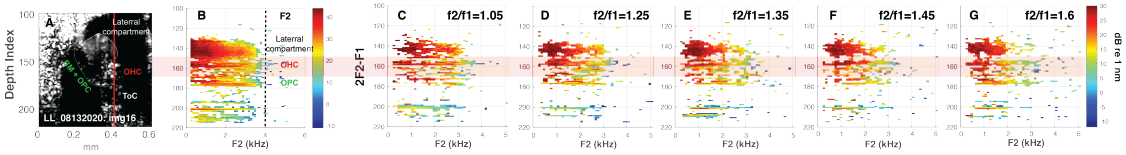


FIGURE 2. Dependence of $2f_2 - f_1$ DP amplitude on ratio f_2/f_1. **(A)** OCT B-scan showing a cross-section of the gerbil cochlea. The vertical red line indicates the A-line for which responses are shown in (B–G).**(B)** Two-dimensional amplitude map of f_2 responses, both as a function of frequency (abscissa) and depth into the tissue (ordinate). **(C–G)** Similar amplitude maps for $2f_2 - f_1$ responses using different f_2/f_1 ratios, as indicated in the title of each panel. Corresponding DPOAEs are shown in Figs. 1B,D.

The 2D vibrometry maps show that different structures along this A-line vibrated similarly for responses at f_2 (Fig. 2B) and $2f_2 - f_1$ (Fig. 2C–G). Varying f_2/f_1 ratio from 1.05 to 1.6 did not appreciably change these maps, indicating that the $2f_2 - f_1$ DP generation was not sensitive to the f_2/f_1 ratio. Responses in the OHC and LC regions had a low-pass character, while vibrations in the OPC region were tuned. This suggests that DPs generated in the active OHC region (where cochlear nonlinearity resides) drive the surrounding structures in a manner that is similar at the stimulus and newly generated $2f_2 - f_1$ DP frequencies. The vibration patterns also suggest that the responses were directly linked to the physical properties of the measured structure, consistent with multi-tone responses of the cross section of the OCC recorded from the same location [**?**].

The ratio dependence of $2f_2 - f_1$ DPs is further described by responses at three selected depths along this A-line, representing the LC, OHC, and OPC region, respectively (Fig. 3). First, we use responses at f_1 to illustrate the variation of the stimulus (Figs. 3A–C), which varied with f_2/f_1 ratios (the f_2 was fixed). Due to their low-pass nature, f_1 responses changed little for different f_2/f_1 with f_2 below the local BF of 2.3 kHz. At higher f_2's ,the effect of cochlear filtering of the f_1 response is observed. In other words, because of the low-pass filter nature of the primary responses, varying f_2/f_1 did not change the overlap region, but increased the (local) level of f_1 for higher f_2s. Determined by the responses at both f_1 and f_2,the $2f_2 - f_1$ DP showed little dependence on f_2/f_1 ratio for f_2 frequencies below the local BF (BF was determined from low-intensity single tone BM-responses from the same OCC cross-section, *data not shown*), but its amplitude decreased for increasing f_2/f_1 at higher f_2s (Figs. 3A'–C').

Only in the OHC region (Fig. 3B"), the $2f_2 - f_1$ DP phases induced by narrow f_2/f_1 ratio of 1.05 were close to -0.5 cycle, consistent with local DP generation (see Materials and Methods). At higher f_2/f_1 ratios and at other locations (i.e., LC and OPC, Figs. 3A",C"), the phase responses depart from -0.5 cycle and showed a systematic 'V shape' pattern. Further, the amplitude of the 3rd term in the Taylor series expansion of the primary responses $\left(\text{i.e., } \vec{L_2^2} \times \vec{L_1}/6\right)$ was much larger than the measured DPs (*not shown*). Both observations suggest that $2f_2 - f_1$ DP could not be considered as the direct output from a simple, local compressive nonlinearity, rather additional processes were involved in shaping them post-generation.

The low-frequency sideband DP @ $2f_1 - f_2$ shows complex vibration patterns and is f_2/f_1-dependent

From the same A-line, the vibration patterns for the $2f_1 - f_2$ low-frequency sideband DP, were frequency-, location-, and ratio-dependent (Fig. 4), resulting in a more complex response pattern than observed for f_2 or the $2f_2 - f_1$ DP (Fig. 2). The OHC region (horizontal red band in Figs. 4B–F) vibrated at $2f_1 - f_2$ at all testing f_2's, even when f_2 was above the local cutoff of 4 kHz. Responses in the LC and OPC regions followed the OHC motions at DP frequency and OPC-DP appeared to be tuned.

Representative $2f_1 - f_2$ DP responses at the three OCC locations are given in Fig. 5, which shows how the amplitude and phase varied with f_2/f_1 ratio. Different from the smooth responses to the stimulus tones (see Figs. 3A–C), the $2f_1 - f_2$ DPs in the OHC region showed several local minima and maxima (Fig. 5B). At f_2 frequencies around the BF, amplitude of DP increased with f_2/f_1 ratio. On the other hand, the frequency range over which DP phase was -0.5 cycle decreased with increasing f_2/f_1 ratio (Fig. 5E). This suggests that the non-smoothed pattern in

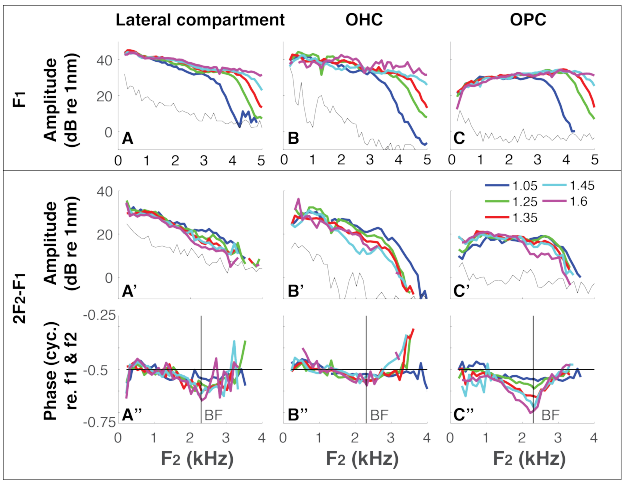


FIGURE 3. Response curves of $2f_2 - f_1$ at representative locations in the OCC. Amplitude for responses at f_1 for **(A)** LC, **(B)** OHC, and **(C)** OPC. **(A'–C')** Amplitude and **(A"–C")** corresponding phase for responses at $2f_2 - f_1$. Phase around -0.5 cycle suggests that a locally generated DP dominated the recorded response.

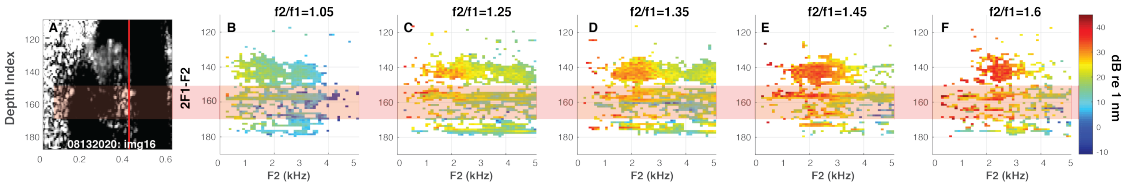


FIGURE 4. Dependence of $2f_1 - f_2$ amplitude on ratio f_2/f_1. (A) OCT B-scan showing a cross-section of the gerbil cochlea. The vertical red line indicates the A-line for which responses are shown in (B–G).**(B–G)** Two-dimensional amplitude maps as a function of stimulus frequency and depth into the tissue for $2f_2 - f_1$ responses using different f_2/f_1 ratios, as indicated in the title of each panel. Vertical spacing: 2.6 μm/pixel. Corresponding DPOAEs are shown in Figs. 1A,C.

amplitude was because of interference of local and non-local components, the non-local DP component affected the local measurements more at wider ratios. The DP amplitudes in the LC and OPC resembled the OHC responses, but appeared to be additionally shaped by the structural physical properties. The robust DPs generated by f_2s above the local cutoff are evidence for DP components that are traveling forward to their own BFs, and are consistent with observations from the high-frequency basal region [**?** **?**]. The existence of this non-local DP component is also confirmed by the phase curves at these frequencies that departure from -0.5 cycle. At f_2 below 1.5 kHz, the $2f_1 - f_2$ phase corresponded to a negative slope, which was considered as dominated also by a non-local, reverse traveling component [**?**].

As for $2f_2 - f_1$, we explored the correlation between DPs and the primary responses by calculating the Taylor series expansion for $2f_1 - f_2$. If the DPs were direct outputs from a simple compressive nonlinearity, the calculated amplitude maps (Fig. 6A-E) should be comparable to the recorded DP maps (Fig. 4) for f_2 below the cutoff frequency. The amplitude of the vector difference between these DP estimates and the recorded DPs are shown in Fig. [6. The difference was close to zero at regions with phase of -0.5 cycle (see arrows pointing the gap regions in Figs. 6A'-E'). This confirmed that over this limited frequency range the recorded DPs were dominated by a local interaction of the two primaries. At other frequencies, the DPs were different from the prediction by the primary responses, and other (non-local) components, or additional signal processing needs to be considered.

DISCUSSION

Intracochlear DPs were recorded from various structures in the OCC, located in the second turn of gerbil cochleae. They were evoked by two-tone stimuli using various f_2/f_1 ratios. We simultaneously recorded DPOAEs, which are

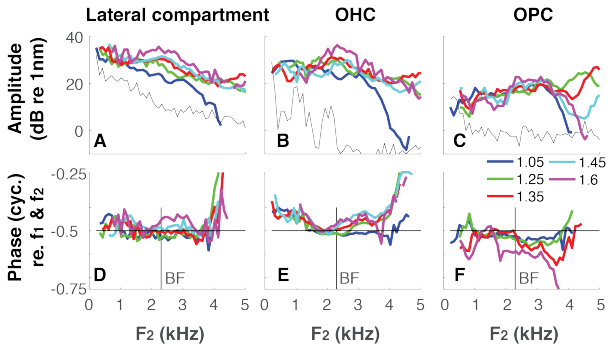


FIGURE 5. Response curves of $2f_1 - f_2$ at representative locations. **(A–C)** Amplitude and **(A'–C')** corresponding phase for responses at $2f_1 - f_2$ for the three selection OCC regions. Phase around -0.5 cycle suggests that a locally generated DP dominates the response.

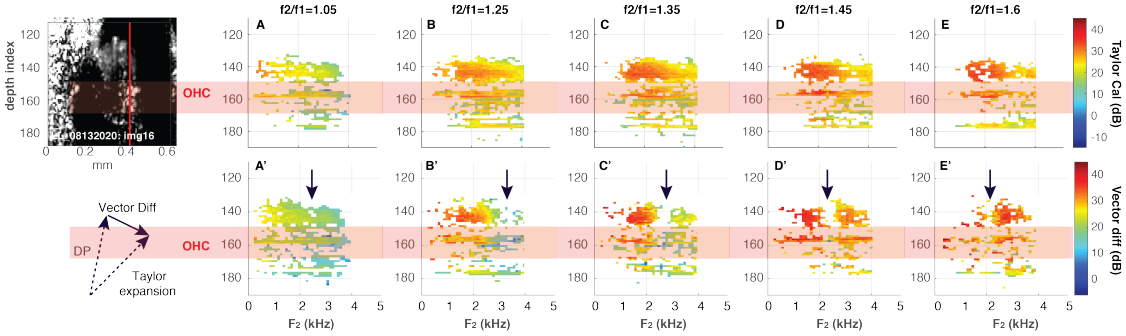


FIGURE 6. Comparison between measured and calculated $2f_1 - f_2$ DP amplitude. **(A–E)** Two-dimensional amplitude maps as a function of stimulus frequency and depth into the tissue for calculated $2f_2 - f_1$ responses. The calculations involve the Taylor series expansion of a compressive nonlinearity in response to the primaries. **(A'–E')** Amplitude maps showing the vector difference between the measured (Fig. 4) and the calculated DP responses. Vertical arrows indicate frequency regions where the measured and calculated response were equal, which is interpreted as evidence for local DP generation. Vertical spacing: 2.6 μm/pixel.

the summation of all DP wavelets reaching the ear canal. Seemingly, the DPs originate in the OHC region, where the cochlea's nonlinearity resides. These DPs then drove vibrations in other OCC structures.

We documented how the properties of primary tones determine the DP characteristics. The tuning of the upper-sideband DP ($2f_2 - f_1$) was similar to that of the primaries, with similar cutoff frequency of 4 kHz, suggesting that they all were filtered similarly by the cochlea (Figs. 2, 3), i.e., the f_2. Varying the f_2/f_1 ratio only had an effect for f_2 above BF. In addition, their phases departure from -0.5 cycle and they could not be precisely predicted from the primaries using Taylor series expansion. These observations suggest that $2f2 - f_1$ DPs primarily originate from a location basal to the f_2 response, presumably their own BF location.

On the other hand, the responses of the lower-sideband DPs ($2f_1 - f_2$) were more complicated and included both local and non-local components. Only for f_2 frequencies around the BF of the recording site we found that the DPs were -0.5 cycle out of phase with the primaries, suggesting local DP generation. This was confirmed by calculations using the Taylor series expansion and the primary responses, which only accurately "predicted" the DPs over this near-f_2 frequency range (Fig. 6). This local-generation region narrowed with increasing f_2/f_1 ratio. This is consistent with the scenery of the stimulus that f_1 moved to lower frequencies, being more non-local, at wider ratios. In addition, $2f_1 - f_2$ DPs appeared to increase with f_2/f_1 ratio at the BF, suggesting the two-tone suppression might also play a role in determining the size of the primaries to generate DPs. Our observation emphasize the properties of the primaries, i.e., local or non-local, determine the properties of measured $2f_1 - f_2$ DP.

In summary, our results further illustrate how primary tones generate/shape the DPs. Our data suggested that the less f_2/f_1 ratio dependent of low-frequency DPOAEs originate form the little variation of the overlap region of the primaries. On the other hand, the current data could not determine the generation region, the spatial extent, where DP wavelets contributed to DPOAEs. Further designed experiments are needed to localize the region and establish the

relationship between the DPs and DPOAEs.

ACKNOWLEDGMENTS

This work was supported by NIH/NIDCD R21DC019998, and VA Merit Award C2296-R to W.D. These contents do not represent the views of the U.S. Department of Veterans Affairs or the United States Government.

REFERENCES

Similar tuning of distortion-product otoacoustic emission ratio functions and cochlear vibrations in mice

James B. Dewey[1, a)] and Christopher A. Shera[1, 2]

[1]*Caruso Department of Otolaryngology – Head and Neck Surgery, University of Southern California, Los Angeles, CA 90033, USA*

[2] *Department of Physics and Astronomy, University of Southern California, Los Angeles, CA 90089, USA*

[a)]Corresponding author: jamesdew@usc.edu

Abstract. When elicited by two stimulus tones (at frequencies f_1 and f_2, $f_2 > f_1$), the amplitudes of specific distortion-product otoacoustic emission (DPOAE) components exhibit a characteristic bandpass shape as the ratio between f_2 and f_1 is varied. This bandpass shape has been attributed to various mechanisms including intracochlear resonance, suppression, and wave interference, and has been proposed to be related to cochlear frequency tuning. While human studies suggest modest correlations between psychophysical tuning and the tuning of DPOAE amplitude vs. f_2/f_1 ratio functions, a relationship between the latter and the tuning of cochlear mechanical responses has yet to be established. This was addressed here through direct comparisons of DPOAEs and cochlear vibrations in wild-type CBA/CaJ mice. DPOAEs were elicited using a fixed-f_2, swept-f_1 paradigm, and optical coherence tomography was used to measure displacements from cochlear locations with characteristic frequencies near f_2. The tuning sharpness of $2f_1–f_2$ DPOAE ratio functions was found to be remarkably similar to that of basilar membrane and/or tectorial membrane responses to single tones, with the tuning sharpness of all responses increasing similarly with decreasing stimulus level. This relationship was observed for f_2 frequencies ranging from ~8 to 22 kHz. Intracochlear distortion products did not exhibit a bandpass shape as the f_2/f_1 ratio was varied, indicating that interference between distortion products traveling to the stapes may be responsible for the tuning of the DPOAE ratio function. While these findings suggest that DPOAE ratio functions could be used to noninvasively infer cochlear tuning, it remains to be determined whether this relationship holds in other species and for lower frequency regions.

INTRODUCTION

Due to nonlinear amplification of cochlear vibrations by the outer hair cells (OHCs), stimulation of the ear with two tones (at frequencies f_1 and f_2, $f_2 > f_1$) results in the generation of mechanical distortion products (DPs)[1]. Some of this DP energy is transmitted to the ear canal as DP otoacoustic emissions (DPOAEs), with the $2f_1–f_2$ DPOAE being the most readily measurable component[2]. While DPOAEs undoubtedly provide a window onto OHC function and cochlear mechanics, their interpretation is complicated by the fact that they represent the sum of DP waves generated over a potentially broad cochlear region. DP waves arriving at the stapes from different locations can therefore interfere constructively and destructively, obscuring the relationship between the measured DPOAE and the DP generated at any particular cochlear location. Thus, beyond identifying whether OHCs are present and functional, it remains uncertain whether DPOAEs can accurately convey more specific information about the macro- and micromechanical processes that generate them.

A particularly intriguing example of the complexity in DPOAEs is the bandpass shape observed in the amplitudes of certain components, including $2f_1–f_2$, as the ratio between f_2 and f_1 is varied[3-6]. In fact, DPs at frequencies with the form $(n+1)f_1 – nf_2$ (n = 1, 2, 3, …) all peak at the same frequency, typically falling ~0.5–0.7 octaves below f_2[4,6]. While a decrease in DPOAE amplitude at large f_2/f_1 ratios can be explained by the reduced interaction between the vibrations elicited by the two stimulus tones, the decrease in amplitude at smaller ratios is

more puzzling. This is because, as f_1 approaches f_2, both the spatial extent of the DP generation region and the magnitude of the DPs generated at each location presumably increase[7]. While the bandpass shape has previously been attributed to mechanisms that shape DPs near where they are generated, such as a tectorial membrane resonance[8-10] or local suppression[11], it has been increasingly understood as the natural consequence of a more global, wave-interference phenomenon[12-14].

Regardless of the mechanism(s) underlying the bandpass shape, the tuning of DPOAE amplitude vs. f_2/f_1 ratio functions has often been proposed to relate in some way to the sharpness of cochlear frequency tuning[5-6,9,14]. For instance, the sharpness of cochlear tuning dictates the degree of nonlinear interaction between the stimulus-driven waves, with more sharply tuned responses requiring smaller f_2/f_1 ratios to elicit measurable DPs. More sharply tuned responses are also theoretically associated with more rapid spatial variation in the phases of cochlear traveling waves, which controls the degree of wave interference experienced by the DPs traveling to the stapes. Recent modeling of DPOAE generation in the human cochlea suggests that the tuning of DPOAE ratio functions is not only correlated with, but quantitatively similar to, the tuning of cochlear vibrations[14]. Indeed, DPOAE ratio functions tend to become more sharply tuned at lower stimulus levels and higher f_2 frequencies[5,14-15], mirroring the dependence of cochlear mechanical tuning on stimulus level and characteristic frequency (CF). Nevertheless, the tuning of DPOAE ratio functions has been found to be only modestly correlated with psychophysical tuning estimates[15-16], and their relationship with the actual tuning of cochlear vibrations remains untested. Here we directly assess this possible association in mice.

METHODS

Cochlear vibrations and DPOAEs were compared in adult (3–7 week old) wild-type CBA/CaJ mice of either sex. Mice were anesthetized (80–100 mg/kg ketamine, 5–10 mg/kg xylazine) and placed on a heating pad to maintain body temperature at 38° C, with supplemental anesthetics given to maintain areflexia. After fixing the skull to a head holder with dental cement, a ventrolateral surgical approach was used to expose the left bulla, which was then widely opened to allow optical access to the cochlear bone and ossicular chain. The pinna was then removed and the ear canal resected so than an acoustic probe (ER10X, Etymotic Research, Elk Grove, IL) could be sealed in place over the eardrum. Sound stimuli were calibrated using the pressure measured by the probe microphone, after correcting for its frequency-dependent sensitivity. Following completion of the desired measurements, mice were euthanized via anesthetic overdose. All procedures were approved by the Institutional Animal Care and Use Committee at the University of Southern California.

Optical coherence tomography (OCT) was used to image and measure vibratory responses from apical and basal turns of the mouse cochlea, where the CFs were ~7.5–11 and ~20–22 kHz, respectively (Fig. 1). The custom-built system and relevant methodology have been described previously[17-18]. For apical locations, vibrations were measured from the basilar membrane (BM), if not also the tectorial membrane (TM) and the bottom of the OHC region, near the Deiters' cells. For basal locations, the image quality was not sufficient to resolve the anatomy with any certainty. However, single-tone responses with adequate signal-to-noise ratios could be measured from either the bottom or top of the cochlear partition, likely near the BM or the reticular lamina/TM boundary, respectively.

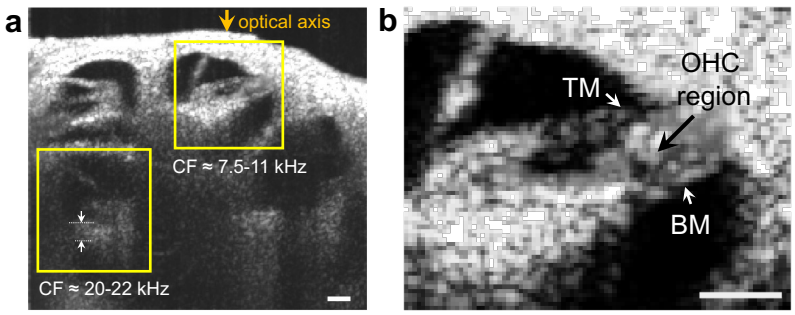


FIGURE 1. (a) Cross-sectional OCT image of a representative mouse cochlea. Vibratory measurements were obtained from the highlighted apical and basal regions. For the basal turn, vibrations were measured either from the bottom or top of the partition (arrows/dotted lines). (b) Close-up image of the apical turn. Vibrations were measured from approximate locations indicated on the BM, OHC region, and TM. Scale bars in a-b = 100 μm.

Single-tone responses were obtained from 1–15 kHz in 0.5 kHz steps (for apical sites) or 1–30 kHz in 1 kHz steps (for basal sites). Two-tone responses were measured from the ear canal and cochlea, when feasible, with f_2 fixed at or near the CF of the measurement site and f_1 varied to cover f_2/f_1 ratios of ~1.02–1.8. The two stimulus tones were presented at the same sound pressure level (SPL) and typically varied from 30–70 dB SPL in 10 dB steps. Stimuli were 102 ms tones with 1 ms ramps applied at the beginning and end of the waveforms, and responses were averaged over 8–32 stimulus presentations. A fast Fourier transform was applied to the steady-state portion of the average response to obtain the magnitude and phase at the stimulus frequencies and any relevant DP frequencies. For a given frequency component, the mean and standard deviation (SD) of response magnitudes at nearby frequencies were used to estimate the measurement noise. Only responses with magnitudes at least 3 SDs above the mean noise were considered for further analysis. All response phases were referenced to the phase of the stimulus measured in the ear canal, with DPOAE and intracochlear DP phases referenced to $2\varphi_1-\varphi_2$, where φ_1 and φ_2 were the phases of the f_1 and f_2 stimuli.

To compare the tuning sharpness of cochlear vibrations and DPOAE ratio functions, the frequency where the maximum response occurred was divided by the bandwidth 10 dB below the maximum, yielding the Q_{10dB}. For DPOAE responses, Q_{10dB} values were calculated with respect to the DP frequency.

RESULTS

The tuning sharpness of DPOAE ratio functions is similar to that of cochlear vibrations

Representative BM responses to single tones measured from the 9 kHz region are compared with DPOAE ratio functions from the same mouse in Figure 2. At low stimulus levels, BM responses were sharply tuned to the CF, while at higher levels, responses peaked at slightly lower frequencies and became more broadly tuned (Fig. 2a). DPOAE ratio functions obtained with f_2 = 9 kHz and f_1 varied from ~4.8 to 8.8 kHz were also tuned, exhibiting a bandpass shape that peaked when the f_2/f_1 ratio was roughly 1.3, or when the DPOAE frequency was ~5 kHz (Fig. 2b). While the f_2/f_1 ratio resulting in the largest DPOAE varied little with stimulus level, ratio functions did become more broadly tuned as the stimulus level was increased. In fact, the tuning sharpness of the DPOAE ratio functions was quantitatively similar to that of single-tone BM responses across the range of stimulus levels explored, decreasing by a factor of ~2 as stimulus levels increased from 40 to 70 dB SPL (Fig. 2c). Q_{10dB} values for DPOAE ratio functions were also similar to those of TM responses but were generally higher than those of OHC region responses. The broader tuning of OHC region responses can be attributed to their relatively low-pass characteristic[18-19].

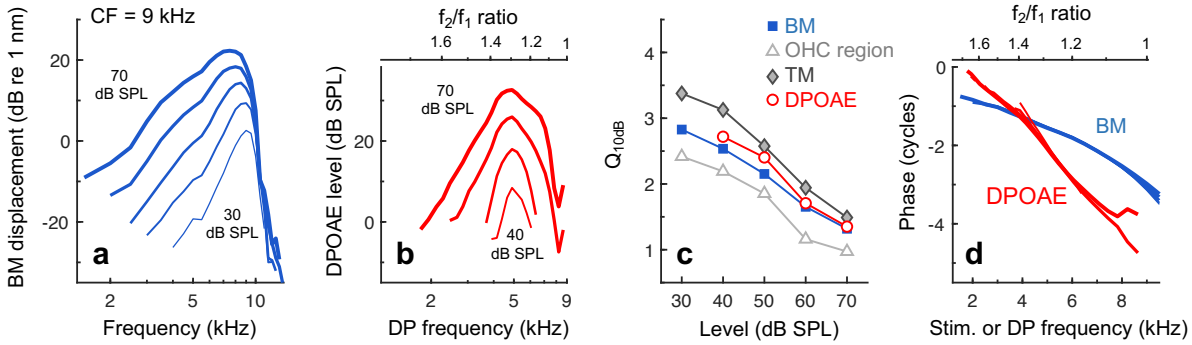


FIGURE 2. (a) BM displacement magnitudes as a function of stimulus frequency and level from the 9 kHz region in a representative mouse. (b) DPOAE amplitudes from the same mouse with f_2 = 9 kHz and f_1 varied. Both stimulus tones were presented at the same level (30–70 dB SPL in 10 dB steps, though DPOAEs were typically not measurable below 40 dB SPL). DPOAEs are plotted vs. the DP frequency, with the corresponding f_2/f_1 ratio shown at the top of the panel. (c) Q_{10dB} values calculated from the BM and DPOAE responses in a-b, plotted vs. stimulus level. Also shown are Q_{10dB} values of OHC region and TM displacements from the same mouse. (d) Phases of BM and DPOAE responses plotted vs. the stimulus or DP frequency, respectively, using a linear frequency axis. Corresponding f_2/f_1 ratios for DPOAEs are shown at the top of the panel.

While the physiological interpretation of DPOAE phases obtained with a fixed-f_2, swept-f_1 paradigm remains unclear[20], we also compare the phases of BM responses and DPOAEs (plotted vs. the DP frequency) in Figure 2d. DPOAE phases rotated rapidly with frequency, with the rate of phase accumulation being ~2–4 times that observed for BM responses over the same frequency range. When considering the phase of BM responses over the range of f_1 frequencies used in the DPOAE measurement, DPOAE phases rotated ~1.5–2.5 times as rapidly with frequency. Compared to the BM phase near the CF (i.e., f_2), DPOAE phases rotated ~1–1.5 times as rapidly with frequency. Delays computed from the negative slope of the DPOAE phase vs. frequency curves were largest near the peaks of the ratio functions, where they were ~0.85 ms, while near-CF BM phase-gradient delays were ~0.5–0.6 ms.

Similarity in tuning sharpness holds across CFs and f_2 frequencies

Measurements obtained at other CF regions and f_2 frequencies revealed a similar quantitative agreement between the tuning of cochlear vibrations and DPOAE ratio functions (Fig. 3). For f_2 frequencies just below or above 9 kHz, DPOAEs tended to peak at f_2/f_1 ratios near 1.3, though the optimal f_2/f_1 ratio and DP frequency could sometimes decrease or increase with stimulus level (Fig. 3b). For f_2 frequencies of ~20–22 kHz, DPOAEs peaked at f_2/f_1 ratios closer to ~1.1–1.2, where the DP frequency was 15–16 kHz (Fig. 3f). At these higher f_2 frequencies, DPOAE ratio functions often exhibited multiple smaller peaks, though the f_2/f_1 ratio or DP frequency at the largest peak was relatively stable across stimulus levels. Despite these differences in response morphology across f_2 frequencies, Q_{10dB} values for DPOAE ratio functions still generally decreased monotonically with increasing stimulus level and resembled those of BM (and TM) responses (Fig. 3c,g). The relative phase rotation vs. frequency for DPOAE and BM responses was also similar to that observed when f_2 was 9 kHz (Fig. 3d,h), though near-CF BM phase-gradient delays and peak DPOAE phase-gradient delays became shorter at higher CFs and f_2 frequencies.

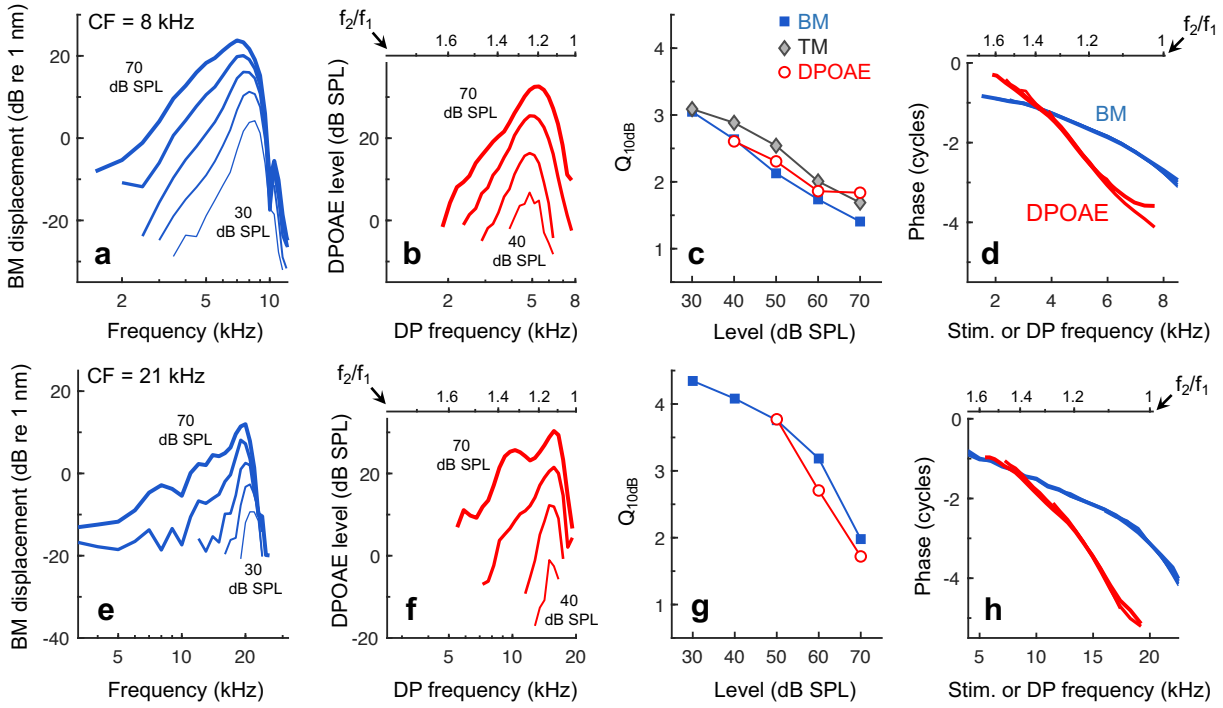


FIGURE 3. (a-d) As in Fig. 2 but for a different mouse in which the CF and f_2 frequency were 8 kHz. Single-tone OHC region responses were not obtained. (e-h) As above but for a mouse in which the CF was 21 kHz and f_2 was 20 kHz. Single-tone responses were only obtained from the bottom of the partition near the BM. Q_{10dB} values for DPOAEs could only be calculated for stimulus levels of 50–70 dB SPL.

Figure 4a summarizes variations in Q_{10dB} for responses of the BM (or top of the organ of Corti, for one basal measurement) to 50 dB SPL tones at eleven distinct CFs in eight mice, along with the accompanying DPOAE Q_{10dB} for an f_2 frequency fixed near the CF. Vibratory Q_{10dB} values steadily increased with CF, while the variation in DPOAE Q_{10dB} values with f_2 was not as smooth. However, there was still a clear separation between DPOAE Q_{10dB} values for the lower and higher ranges of f_2 frequencies, and the average DPOAE Q_{10dB} values for these two ranges (2.22 and 3.80, respectively) agreed well with the average vibratory Q_{10dB} (2.24 and 3.83).

Figure 4b compares phase-gradient delays computed for the vibratory and DPOAE responses, demonstrating the shorter absolute delays with increasing CF and f_2. For vibratory responses, phase-gradient delays were calculated at the CF, while DPOAE phase-gradient delays were taken at the peaks of the DPOAE ratio functions. When expressed in periods of the stimulus frequency, vibratory delays increased with frequency in a manner similar to the changes in Q_{10dB}, increasing ~1.5-fold from low to high CFs. The covariance in tuning sharpness and delay (in periods) is consistent with what is observed for simple harmonic oscillators[21]. Since spatial gradients in the phases of the stimulus-evoked traveling waves determine the influence of wave interference on DPOAEs, the more rapid phase variation at higher CF regions may explain why DPOAEs peak at smaller f_2/f_1 ratios with increasing f_2 frequency.

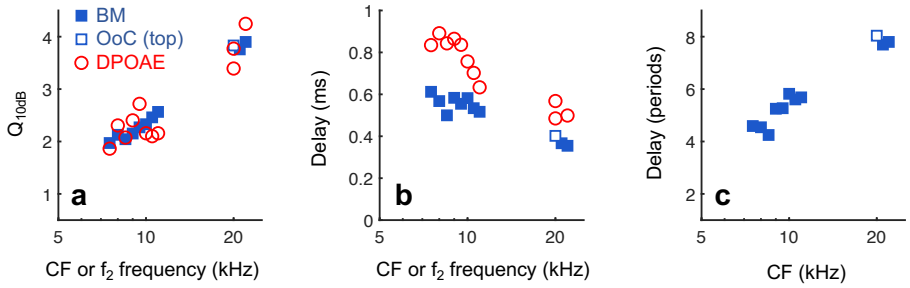


FIGURE 4. (a) Q_{10dB} for cochlear vibrations and DPOAE ratio functions across different CFs and f_2 frequencies for 50 dB SPL stimuli. Vibratory Q_{10dB} are plotted for the BM, except for one measurement from the top of the organ of Corti (near the reticular lamina and TM) in the 20 kHz region. (b) Corresponding phase-gradient delays for cochlear vibrations (calculated at the CF) and DPOAE ratio functions (calculated at the peak of the function). (c) Cochlear delays in stimulus periods follow the same trend as the Q_{10dB}.

Bandpass ratio functions are not due to filtering of DPs where they are generated

Intracochlear DPs were also often monitored on the BM, OHC region, and TM during the measurement of DPOAE ratio functions (Fig. 5). Intracochlear DPs grew in magnitude as the f_2/f_1 ratio approached 1 and were almost always largest at the smallest ratio examined (Fig. 5a). Though nonmonotonic growth was sometimes observed at intermediate ratios, intracochlear DPs never exhibited the bandpass shape observed in DPOAE amplitudes. This confirms that the bandpass tuning of DPOAE ratio functions does not originate from local filtering of DPs near their generation site due to mechanical resonance or suppression. Instead, the decline in DPOAE amplitudes at small f_2/f_1 ratios is presumably due to wave interference.

For reference, DPOAE and intracochlear DP phases are compared in Figure 5b. The phase rotation vs. frequency was similar for DPOAEs and intracochlear DPs except near the peak of the DPOAE ratio function, where the phase rotation was more rapid for DPOAEs. Interpreting these differences in phase behavior is not straightforward, as intracochlear DPs are a mixture of locally generated and propagating components, and DPOAEs are shaped by interference between sources with widely varying phases.

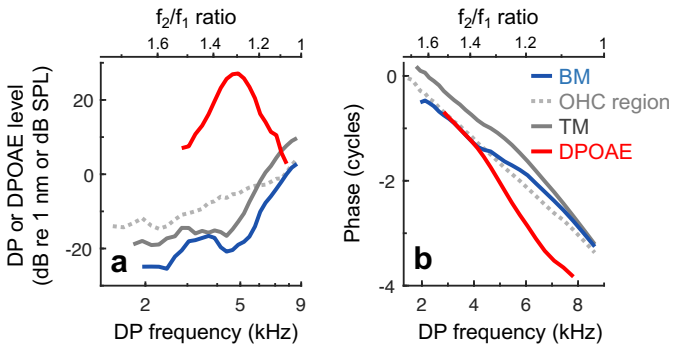


FIGURE 5. (**a**) DPOAE and intracochlear DP amplitude ratio functions from a representative mouse, with f_2 = 9 kHz and stimuli presented at 60 dB SPL. Intracochlear DPs were measured from the 9 kHz region. (**b**) Response phases shown on a linear frequency axis.

DISCUSSION

Here we demonstrated a rough quantitative agreement between the tuning sharpness of $2f_1–f_2$ DPOAE ratio functions and single-tone vibratory responses measured at the location tuned to f_2. The agreement holds across stimulus level, with tuning sharpness decreasing with increasing level, as well as across frequency, with tuning sharpness increasing at higher CFs and f_2 frequencies. Nevertheless, whether DPOAE ratio functions are actually useful for predicting cochlear tuning in individual ears remains to be tested. The diverse morphology of the DPOAE ratio functions, which sometimes exhibited multiple peaks and level-dependent notches, may render their Q_{10dB} values more variable than the vibratory Q_{10dB}, as suggested in Fig. 4a. However, average DPOAE and vibratory Q_{10dB} values were very similar, indicating that DPOAE ratio functions could at least provide a noninvasive window onto the tuning of cochlear vibrations at the group level.

Further work is needed to examine if the agreement between DPOAE and cochlear tuning is peculiar to mice or if it is observed in other species. It would be particularly valuable to assess the relationship between cochlear and DPOAE tuning for species in which the apex responds to much lower frequencies. The responses of these low-frequency regions are much more broadly tuned, and the mechanics may differ qualitatively from those in basal regions[22-23], such that the processes underlying DPOAE generation may be different as well. Whether an association between DPOAE and cochlear tuning is found in ears with OHC dysfunction and more passive mechanical responses also requires evaluation. Previous observations of tuned DPOAE ratio functions in hearing-impaired mice indicate that the relationship may not be so simple[24], though the mechanics were never directly assessed.

At least in sensitive ears, a relationship between DPOAE and cochlear tuning does not appear to result from a mechanism that shapes intracochlear DPs near where they are generated (Fig. 5). Instead, as the f_2/f_1 ratio is varied from large to small, DPOAE amplitudes initially grow, owing to the increased nonlinear interactions between the stimulus-driven waves, and then start to decline, owing to destructive interference between DP waves with increasingly variable phases[12-14]. The spatial variation in DP phase is partly inherited from the phases of the stimulus-driven traveling waves, which rotate more rapidly with distance when the responses are more sharply tuned (Fig. 4c). Thus, the tuning of DPOAE ratio functions is anticipated to at least be correlated with cochlear tuning. While the quantitative agreement we have observed in mice has also been demonstrated in a model of human DPOAE generation[14], the precise origin of such an agreement is unclear. Preliminary modeling of DPOAEs in mice suggests that the quantitative similarity between cochlear and DPOAE tuning sharpness depends on assumptions regarding how DPs are generated and propagate toward the stapes. Exploration of this relationship in other species would therefore be useful not only for further validating a DPOAE-based assay of cochlear tuning, but for identifying universal principles of cochlear mechanics and OAE generation.

ACKNOWLEDGMENTS

This work was supported by a Hearing Health Foundation Emerging Research grant (J.B.D.) and NIH/NIDCD grants R21 DC019209 (J.B.D.) and R01 DC003687 (C.A.S.). We thank John Oghalai and Brian Applegate for providing the OCT system used in these experiments.

REFERENCES

1. P. Avan, B. Büki, and C. Petit. *Phys. Rev.* ***93***, 1563–1619 (2013).
2. D. T. Kemp, *Arch. Oto-Rhino-Laryngol.* ***224***, 37–45 (1979).
3. D. O. Kim, *Hear. Res.* ***2***, 297–317 (1980).
4. P. F. Fahey and J. B. Allen, "Characterization of cubic intermodulation distortion products in the cat external auditory meatus," in *Peripheral Auditory Mechanisms*, edited by J. B. Allen, J. L. Hall, A. Hubbard, S. T. Neely, and A. Tubis (Springer-Verlag, New York, 1986), pp. 314–321.
5. F. P. Harris, B. L. Lonsbury-Martin, B. B. Stagner, A. C. Coats, and G. K. Martin, *J. Acoust. Soc. Am.* ***85***, 220–229 (1989).
6. A. M. Brown and S. A. Gaskill, "Can basilar membrane tuning be inferred from distortion measurement?" in *The Mechanics and Biophysics of Hearing*, edited by P. J. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele (Springer-Verlag, New York, 1990), pp. 164–169.
7. W. S. Rhode, *J. Acoust. Soc. Am.* ***122***, 2725–2737 (2007).
8. A. M. Brown, S. A. Gaskill, and D. M. Williams, *Proc. R. Soc. London, Ser. B* ***250***, 29–34 (1992).
9. J. B. Allen and P. F. Fahey, *J. Acoust. Soc. Am.* ***94***, 809–816 (1993).
10. A. N. Lukashkin, V. A. Lukashkina, P. K. Legan, G. P. Richardson, and I. J. Russell, *J. Neurophys.* ***91***, 163–171 (2004).
11. L. J. Kanis and E. de Boer, *J. Acoust. Soc. Am.* ***101***, 1527–1531 (1997).
12. C. L. Talmadge, A. Tubis, G. R. Long, and P. Piskorski, *J. Acoust. Soc. Am.* ***104***, 1517–1543 (1998).
13. C. A. Shera, "Wave interference in the generation of reflection- and distortion-source emissions," in *Biophysics of the Cochlea: From Molecules to Models*, edited by A. W. Gummer (World Scientific, Singapore, 2003), pp. 439–453.
14. R. Sisto, U. S. Wilson, S. Dhar, and A. Moleti, *J. Assoc. Res. Otolaryngol.* ***19***, 511–522 (2018).
15. U. S. Wilson, J. Browning-Kamins, A. S. Durante, S. Boothalingam, A. Moleti, R. Sisto, and S. Dhar, *Int. J. of Audiol.* ***60***, 890–899 (2021).
16. A. M. Brown, S. A. Gaskill, R. P. Carlyon, and D. M. Williams, *J. Acoust. Soc. Am.* ***93***, 3291–3297 (1993).
17. H. Y. Lee, P. D. Raphael, J. Park, A. K. Ellerbee, B. E. Applegate, and J. S. Oghalai, *Proc. Natl. Acad. Sci. U.S.A.* ***112***, 3128–3133 (2015).
18. J. B. Dewey, A. Altoè, C. A. Shera, B. E. Applegate, and J. S. Oghalai (2021). *Proc. Natl. Acad. Sci. U.S.A.* ***118***, e2025206118 (2021).
19. A. Vavakou, N. P. Cooper, and M. van der Heijden, *eLife* ***8***, e47667 (2019).
20. C. A. Shera, C. L. Talmadge, and A. Tubis, *J. Acoust. Soc. Am.* ***108***, 2933–2948 (2000).
21. C. A. Shera and K. K. Charaziak, *Cold Spring Harb. Perspect. Med.* ***9***, a033498 (2019).
22. N. P. Cooper, and W. S. Rhode, *Hear. Res.* ***82***, 225–243 (1995).
23. A. Recio-Spinoso and J. S. Oghalai, *J. Physiol.* ***595***, 4549–4561 (2017).
24. D. J. Konrad-Martin, S. Norton, K. E. Mascher, and B. L. Tempel, *Hear. Res.* ***151***, 205–220 (2001).

A Component of Stimulus-Frequency Otoacoustic Emissions Evoked due to Perturbation of Nonlinear Force in a Cochlear Model

Václav Vencovský,[1, a)] Ondřej Klimeš,[1, b)] and Aleš Vetešník[2, c)]

[1)] *Department of Radioelectronics, Czech Technical University in Prague, Czech Republic.*

[2)] *Department of Nuclear Chemistry, Czech Technical University in Prague, Czech Republic*

[a)] *Corresponding author: vaclav.vencovsky@gmail.com*

[b)] *Electronic mail: klimeon5@fel.cvut.cz*

[c)] *Electronic mail: ales.vetesnik@fjfi.cvut.cz*

Abstract. Stimulus-frequency otoacoustic emissions (SFOAEs) are evoked by a single tone presented into the ear. The emissions are generated by reflection of a forward traveling wave (TW) on localized irregularities in the impedance along the basilar membrane. The strongest wavelets are reflected from the place where the TW reaches its largest amplitude. Since the irregularities are localized in fixed places along the basilar membrane, latency of the SFOAE phase can be used to estimate delay of cochlear filters. We used an iterative approach of Shera et al. [JASA (2005) 118:287-313] to obtain an analytical solution for SFOAEs in a nonlinear two-dimensional cochlear model. The solution allowed for decomposition of a reflection component and a component due to perturbation of the nonlinear force. The nonlinear force which in the smooth cochlea may reflect the forward TW backwards [Talmadge et al., JASA (2000) 108, 2911-2931], is perturbed due to irregularities. The perturbed nonlinear force in the model generates an SFOAE component with latency comparable with the latency of the component due to reflection. This means that although the component is evoked due to nonlinear force, it has a comparable (long) latency to the SFOAE component due to reflection. Amplitude of this component due to perturbation of the nonlinear force grows as the amount and amplitude of irregularities increases. In addition, the component amplitude also grows with the tone intensity. The growth is determined by the amplitude of the nonlinear force. It is almost cubic at the lowest intensities and quickly saturates at levels above about 40 dB SPL (the intensity is determined by the position of the compressive nonlinearity in the input/output function of the basilar membrane displacement). An interesting result is that the nonlinear component partly cancels the reflection component because its phase is roughly by 0.5 cycles shifted invariantly to the stimulus intensity and frequency. As the level increases, the component due to perturbation reaches comparable amplitude to the component due to coherent reflection, especially in the portions of SFOAEs with the longest latencies. The destructive interference between the components due to perturbation of the nonlinear force and coherent reflection emphasizes shorter latency wavelets in the overall SFOAEs as stimulus intensity increases and contributes to the saturation of SFOAE amplitude reported, for example, in human experimental data of Abdala and Kalluri [JASA (2017) 142, 812-824].

INTRODUCTION

If a single tone is presented, amplitude of the pressure recorded in the external auditory canal usually quasiperiodically fluctuates as a function of the tone frequency; see Fig. 1 in Shera and Guinan [1]. The fluctuations are most apparent for low intensity (<40 dB SPL) tones and their cause is the interference between the evoking stimulus (presented tone) and emission from within the cochlea. The emission is generated due to coherent reflection of traveling wave [2, 3] from spatial randomly distributed mechanical irregularities along the organ of Corti [4]. For tones used as evoking stimuli, the emission is denoted stimulus-frequency otoacoustic emission (SFOAE) and it can be extracted by various techniques [5].

In this paper, we present the solution for SFOAEs in a two-dimensional, nonlinear cochlear model by using the cochlear "Born expansion" technique presented by Shera et al. [6]. The presented solution is composed of two SFOAE components: (1) a component due to coherent reflection and (2) a component which we denoted *a component due to perturbation of the nonlinear force*. This letter component would not be present in the solution if the model was linear. The paper presents SFOAEs derived from the cochlear model and the both SFOAE components. Similarly as in the solution for DPOAEs presented in [7], the component due to perturbation of the nonlinear force has approximately opposite phase than the component due to coherent reflection. Interaction of both components contributes to the saturation of SFOAE amplitude with increasing intensity of the evoking probe tone.

MODEL

We present the model equations in the frequency domain by assuming that the stationary BM displacement in the time domain can be approximated by the truncated Fourier series

$$\xi(x,t) \approx \sum_{n=-N}^{N} \hat{\xi}_n(x) e^{in\omega_0 t}, \tag{1}$$

where ω_0 is the fundamental frequency which is "small enough that the input frequencies may be chosen as close as is desired to any finite set of real values" [8]; and n are integers. Steady-state transversal displacement of the BM, $\hat{\xi}_n(x)$, at frequency $\omega = n\omega_0$ corresponding to the n*th* stationary input tone is in the two-dimensional (2D), nonlinear cochlear model given by

$$\int_0^L \hat{K}_n^0(x,\bar{x};\partial_x)\hat{\xi}_n(\bar{x})d\bar{x} = -\hat{U}_n(x) + (n\omega_0)^2 G_S(x)\hat{\sigma}_n, \tag{2}$$

where the circumflex represents a phasor, x is the position along the BM length, $\hat{\sigma}_n$ is the stapes displacement, $G_S(x)$ is the Green's function accounting for the fluid coupling between the stapes and BM, $\hat{K}_n^0(x,\bar{x};\partial_x)$ is a linear integro-differential operator [8]

$$\begin{aligned} K_n^0(x,\bar{x};\partial_x) = [k(x) + in\omega_0 h(x) - (n\omega_0)^2 m(x)]\delta(x-\bar{x}) - in\omega_0\partial_x[s(x)\partial_x]\delta(x-\bar{x}) \\ -(n\omega_0)^2 G(x,\bar{x})\delta(x-\bar{x}), \end{aligned} \tag{3}$$

where $\delta(x-\bar{x})$ is Dirac's delta centered at $\bar{x}$, $k(x)$, $h(x)$, and $m(x)$, are the stiffness, damping and mass per unit BM length, respectively, $s(x)$ is the shearing viscosity per unit BM length, and $G(x,\bar{x})$ is the Green's function accounting for fluid coupling between the BM segments. $\hat{U}_n(x)$ is the feedback undamping force which in the model simulates the cochlear amplifier. Vetešník and Gummer [9] separated the undamping force into the linear part $u(x)\hat{T}_n^{\mathrm{TM}}(x)\hat{\xi}_n(x)$ and the nonlinear part $\hat{U}_n^{\mathrm{NL}}(x)$; namely,

$$\hat{U}_n^{\mathrm{NL}}(x) \equiv u(x)\frac{1}{T_0}\int_0^{T_0} S\left[\sum_{n=-N}^{N} \hat{T}_n^{\mathrm{TM}}(x)\hat{\xi}_n(x)e^{in\omega_0 t}\right] e^{-in\omega_0 t}dt - u(x)\hat{T}_n^{\mathrm{TM}}(x)\hat{\xi}_n(x). \tag{4}$$

where the transfer function $\hat{T}_n^{\mathrm{TM}} = \hat{\eta}_n(x)/\hat{\xi}_n(x)$, the term $u(x)$ is a function controlling the gain and $S(\cdot)$ is a nonlinear function proportional to the second order Boltzmann function [8, 10]. $\hat{\eta}_n$ represents a displacement of the second array of oscillators which in the model simulates micromechanics of the organ of Corti and tectorial membrane (TM). These oscillators are driven by the BM acceleration; namely,

$$(-n^2\omega_0^2)\hat{\eta}_n(x) + in\omega_0\gamma_{\mathrm{TM}}(x)\hat{\eta}_n(x) + \omega_{\mathrm{TM}}^2\hat{\eta}_n(x) = -(-n^2\omega_0^2)\hat{\xi}_n(x), \tag{5}$$

where $\gamma_{\mathrm{TM}}(x)$ is damping and $\omega_{\mathrm{TM}}^2(x)$ is their resonance frequency. In their resonance, the phase shift between the $\hat{\eta}_n(x)$ and the BM acceleration $-(n\omega_0)^2\hat{\xi}_n(x)$ is $\pi/2$, which assures that $\hat{U}_n(x)$ is in phase with the BM velocity; $\hat{U}_n(x)$ therefore amplifies the BM motion [8].

The cochlear model is coupled with a middle ear model adapted from [11]. $\hat{\sigma}_n$ is the stapes displacement which is driven by the pressure at the ear drum $\hat{P}_{\mathrm{dr}}$; namely,

$$m_{\mathrm{me}}(-n^2\omega_0^2)\hat{\sigma}_n + h_{\mathrm{me}}in\omega_0\hat{\sigma}_n + k_{me}\hat{\sigma}_n = S_{\mathrm{ow}}\left\{\frac{1}{W_{\mathrm{bm}}}\Big[-\int_0^L G(0,\bar{x})(-n^2\omega_0^2)\hat{\xi}_n(\bar{x})d\bar{x} - G_s(0)(-n^2\omega_0^2)\hat{\sigma}_n\Big] + G_{\mathrm{me}}\hat{P}_{\mathrm{dr}}\right\}, \tag{6}$$

where m_{me} and h_{me} are the total mass of the middle-ear ossicles, k_{me} is the stiffness of the middle-ear ossicles, ligaments, and the tympanic cavities, W_{bm} is the BM width, and S_{ow} is the surface of the oval window. The values of the mass, damping and stiffness were set according to the stapes reflectance [10]. Equation (6) can be combined with Eqs. (7) and (4) to

$$\int_0^L \hat{K}_n^A(x,\bar{x};\partial_x)\hat{\xi}_n(\bar{x})d\bar{x} = -\hat{U}_n^{\mathrm{NL}}(x) + (n\omega_0)^2 G_s(x)\hat{K}_{\mathrm{in}}\hat{P}_{\mathrm{dr}_n}, \tag{7}$$

where

$$K_n^A(x,\bar{x};\partial_x) = [k(x) + in\omega_0 h(x) - (n\omega_0)^2 m(x)]\delta(x-\bar{x}) - in\omega_0\partial_x[s(x)\partial_x]\delta(x-\bar{x}) \\ -(n\omega_0)^2 G(x,\bar{x}) + n\omega_0 \hat{G}_{\mathrm{ME}}(x,\bar{x}) + u(x)\hat{T}_n^{\mathrm{TM}}(x)\delta(x-\bar{x}), \tag{8}$$

where

$$\hat{G}_{\mathrm{ME}}(x,\bar{x}) = \frac{-(n\omega_0)^2 G_S(x)\frac{S_{\mathrm{ow}}}{W_{\mathrm{bm}}} G(0,\bar{x})(-n^2\omega_0^2)}{m_{\mathrm{me}}(-n^2\omega_0^2) + h_{\mathrm{me}} in\omega_0 + k_{\mathrm{me}} + G_S(0)\frac{S_{\mathrm{ow}}}{W_{\mathrm{bm}}}(-n^2\omega_0^2)}, \tag{9}$$

and

$$\hat{K}_{\mathrm{in}} = \frac{\frac{S_{\mathrm{ow}}}{W_{\mathrm{bm}}} G_{\mathrm{me}}\hat{P}_{\mathrm{dr}}}{m_{\mathrm{me}}(-n^2\omega_0^2) + h_{\mathrm{me}} in\omega_0 + k_{\mathrm{me}} + G_S(0)\frac{S_{\mathrm{ow}}}{W_{\mathrm{bm}}}(-n^2\omega_0^2)} \tag{10}$$

By using the Green's function $\hat{H}_n(x,s)$, a solution:

$$\hat{\xi}_n(x) = \int_0^L \hat{H}_n(x,s)\left[-\hat{U}_n^{\mathrm{NL}}(s) + (n\omega_0)^2 G_S(s)\hat{K}_{\mathrm{in}}\hat{P}_{\mathrm{dr}}\right] ds. \tag{11}$$

To generate OAEs due to coherent-reflection, we must add mechanical irregularities into at least one of the terms in Eq. (8). We place the irregularities into the function $u(x)$ affecting the amount of undamping force. The irregularities were added in the form of Gaussian distributed (white) noise; namely,

$$\tilde{u}(x) = u(x) + \Delta u(x), \tag{12}$$

where $\Delta u(x) = \varepsilon u(x)$ and $\varepsilon \approx \mathcal{N}(0,\sigma^2)$. The irregularities can be set up in the kernel

$$\Delta\hat{K}_n = \Delta u(x)\hat{T}_n^{\mathrm{TM}}(x)\delta(x-\bar{x}). \tag{13}$$

The cochlear model with irregularities can then be described by

$$\int_0^L \hat{K}_n^A(x,\bar{x};\partial_x)\hat{\xi}_n^P(\bar{x})d\bar{x} = -\left[\hat{U}_n^{\mathrm{NL}}(x) + \Delta\hat{U}_n^{\mathrm{NL}}(x)\right] + (n\omega_0)^2 G_s(x)\hat{K}_{\mathrm{in}}\hat{P}_{\mathrm{dr}_n} - \int_0^L \Delta\hat{K}_n(x,\bar{x})\hat{\xi}_n^P(\bar{x})d\bar{x}, \tag{14}$$

where the superscript "P" is used to distinguish BM displacement $\hat{\xi}_n^P(x) = \hat{\xi}_n(x) + \Delta\hat{\xi}_n^P(x)$ in the model with irregularities. Solution of Eq. (14) is

$$\hat{\xi}_n^P(x) = \int_0^L \hat{H}_n(x,s)\left[-\hat{U}_n^{\mathrm{NL}}(s) - \Delta\hat{U}_n^{\mathrm{NL}}(s) + (n\omega_0)^2 G_S(s)\hat{K}_{\mathrm{in}}\hat{P}_{\mathrm{dr}}\right] ds - \int_0^L\int_0^L \hat{H}_n(x,s)\Delta\hat{K}_n(s,\bar{x})\hat{\xi}_n^P(\bar{x})d\bar{x}ds. \tag{15}$$

Because the term $\hat{\xi}_n^P(x)$ appears on the both sides of Eq. (15) we employ an iterative approach presented by Shera et al. [6] to approximate the solution; namely,

$$\hat{\xi}_n^P(x) \approx \sum_{l=1}^N \hat{\xi}_n^{P,l}(x), \tag{16}$$

where

$$\hat{\xi}_n^{P,1}(x) \equiv \int_0^L \hat{H}_n(x,s)\left[-\hat{U}_n^{\mathrm{NL}}(s) - \Delta\hat{U}_n^{\mathrm{NL}}(s) + (n\omega_0)^2 G_S(s)\hat{K}_{\mathrm{in}}\hat{P}_{\mathrm{dr}}\right] ds, \tag{17}$$

$$\hat{\xi}_n^{P,l}(x) \equiv -\int_0^L\int_0^L \hat{H}_n(x,s)\Delta\hat{K}_n(s,\bar{x})\hat{\xi}_n^{P,l-1}(\bar{x})d\bar{x}ds. \tag{18}$$

SFOAEs are proportional to $\Delta\hat{\xi}_\omega^P(x)$ at $x=0$; namely,

$$\hat{P}_{\mathrm{SFOAE}} \propto \Delta\hat{\xi}_n^P(0) \approx -\int_0^L \hat{H}_n(0,s)\Delta\hat{U}_n^{\mathrm{NL}}(s)ds + \sum_{l=2}^N \hat{\xi}_n^{P,l}(0). \tag{19}$$

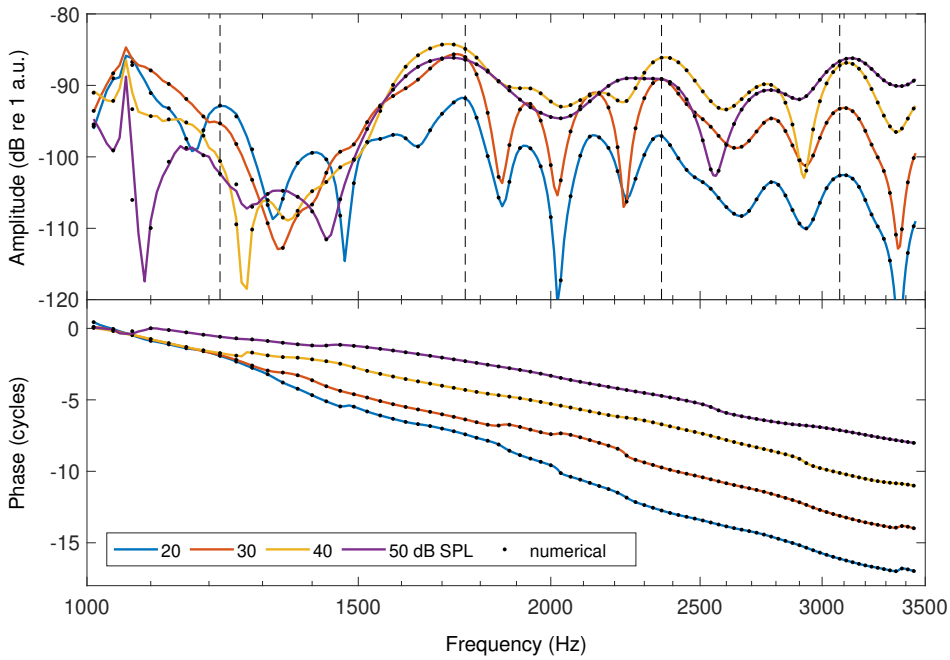


FIGURE 1. Amplitude and phase of SFOAEs derived from analytical solution [Eq. 19), the colored lines] and numerical solution (black dots) of the cochlear model. The equality between the analytical and numerical results proofs the validity of the analytical solution. Depicted data represent SFOAEs for probe tone levels L_p ranging from 20 to 50 dB SPL with a step of 10 dB.

This solution contains two components, the first component $-\int_0^L \hat{H}_n(0,s)\Delta\hat{U}_n^{\mathrm{NL}}(s)ds$ is due to perturbation of the nonlinear force $\Delta\hat{U}_n^{\mathrm{NL}}(s)ds$ and the second component $\sum_{l=2}^{N}\hat{\xi}_n^{P,l}(0)$ is due to coherent-reflection [Eq. (18)]. We distinguish here these two components, but it should be noted that it is the mechanism of coherent reflection which generates wavelets (reflected waves) which perturb the nonlinear force, i.e., which cause that the nonlinear force in the smooth model and in the model with irregularities differs by $\Delta\hat{U}_n^{\mathrm{NL}}(x)dx$. It should also be noted that the nonlinear force can reflect the traveling wave back but this mechanism should produce a low latency component of SFOAE [12, 13, 14]. It is shown below in the results that the component due to perturbation of the nonlinear force has long latency, which is comparable to the latency of the coherent reflection component.

RESULTS

To verify analytical solution [Eq. (19)], Fig. 1 depicts SFOAEs (displacement of the most basal BM segment $\Delta\hat{\xi}_\omega^P(0)$ expressed in dB relative to 1 a.u.; artifical units because the model is not calibrated to simulate BM displacement or acoustic pressure in the ear canal quantitatively) derived from the analytical solution for $N = 11$ with the BM displacement calculated from numerical solution of the cochlear model; the numerical BM displacement was obtained by subtraction of the most basal BM segment displacement for the model with irregularities and for the smooth model. Numerical and analytical solution well-nigh superimpose which puts a confidence in the analytical solution for $N = 11$ iterations. The irregularities were introduced into the model according to Eq. (12) for $\sigma = 0.05$.

It is then of interest to analyze how the component due to perturbation of the nonlinear force (NL denoted in the results) and the component due to coherent-reflection (CR denoted in the results) contribute to the final SFOAE and how they change with increasing probe tone level. Figure 2 depicts the CR and NL components for the same stimulus parameters as in Fig. 1. The grey area indicates the region with data from Fig. 1. It is shown here for a comparison between data in Fig. 1 and Fig. 2 because the y-axis scale is different in the both figures. Notice that the overall SFOAE has smaller dynamic range than the individual components. The reason is that both components have roughly opposite phase and hence destructively interfere and that the NL component grows faster with increasing level than the CR component. It seems that the destructive interference between the both components also contributes to the decrease of overall group delay with increasing probe tone level, which is visible in shallower phase slopes for the

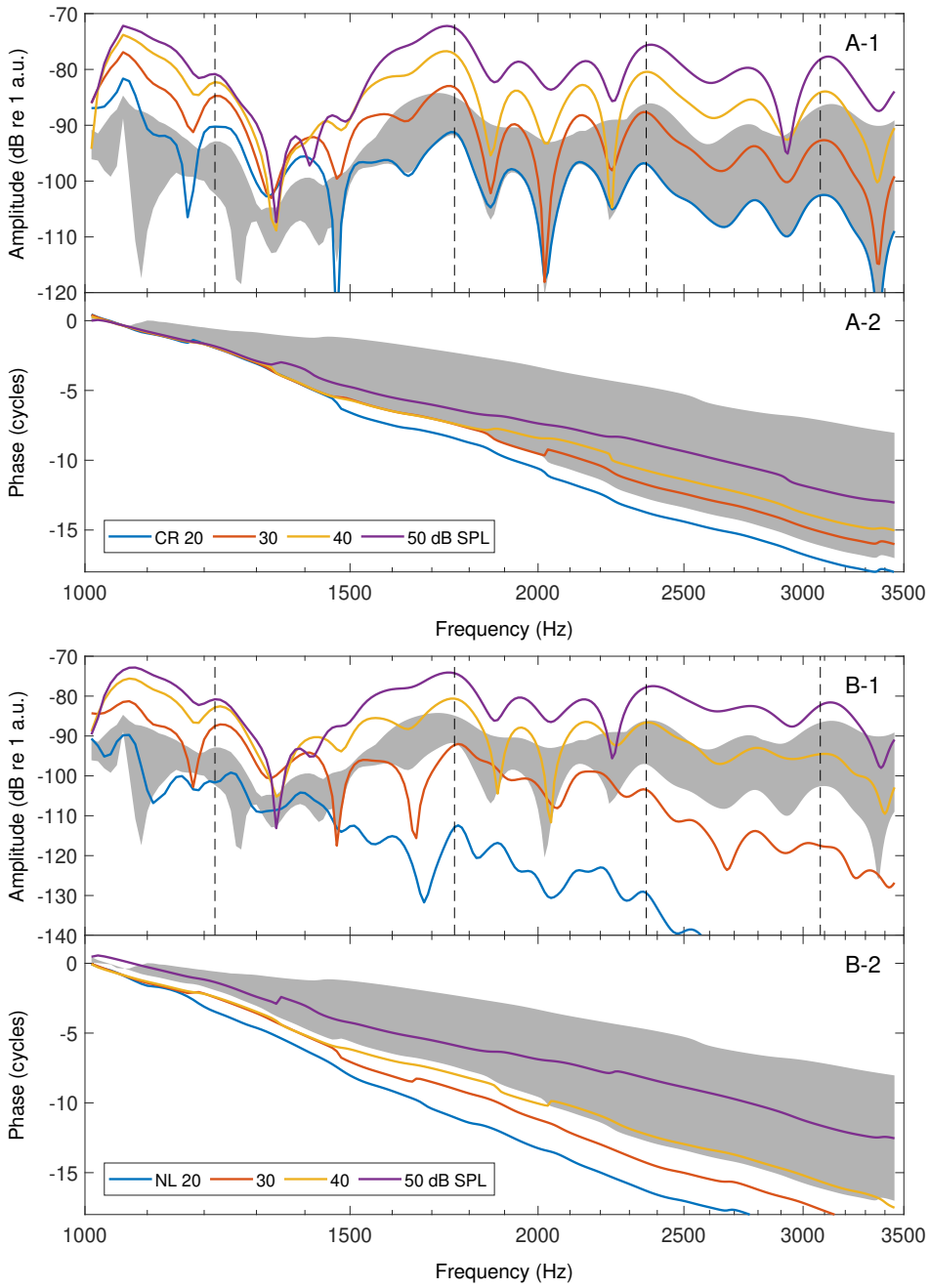


FIGURE 2. Amplitude and phase of the coherent reflection (CR) component for $N = 11$ (panels A-1,2) and component due to perturbation of the nonlinear force (NL) (panels B-1,2). The grey area indicates the range of SFOAEs depicted in Fig. 1. Notice that the dynamic range of CR and NL components is larger than the overall dynamic range of SFOAEs. The legends indicate the probe tone levels.

overall SFOAE and the CR and NL components.

The level growth of SFOAEs and CR and NL components is shown in Fig. 3 for several chosen probe tone frequencies (written in the graphs). The bottom graphs (E-H) depict the BM displacement as a function of a tone level, i.e., input/output (I/O) functions of the model. The I/O functions were simulated in the BM segments tuned to the chosen probe frequencies at 0 dB SPL. The grey bars indicate the level range in the panels A-D. Notice that the destructive interference for $f_p = 1220$ Hz between CR and NL components decreases the overall SFOAE amplitude. At lowest

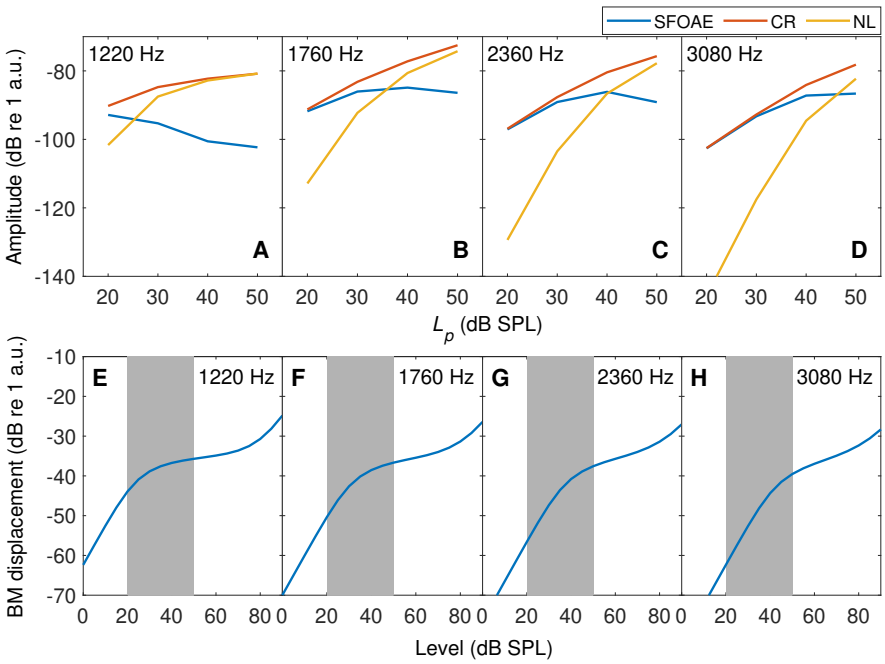


FIGURE 3. A-D: Amplitude of SFOAE and CR and NL components at four frequencies (written in the panels and indicated by the vertical dashed lines in Figs. 1 2) as a function of the probe level. E-H: I/O functions of the model showing the BM displacement as a function of the input tone instensity at four best frequencies (written in the panels) at 10 dB SPL.

probe intensities, the CR component grows approximately by 1 dB/dB and the NL component grows approximately by 2.8 dB/dB (the largest growth at 3080 Hz).

The both components (CR and NL) forming the SFOAE grows in amplitude even in the nonlinear region of the model I/O function. Destructive interference between the components causes that the overall SFOAE amplitude saturates with increasing level and even decreases. SFOAEs recorded in human subjects provide evidence for amplitude saturation, e.g. [15, 16], and in some cases evidence for amplitude decline or notches [16]. However, some studies indicate rather a compressive growth, not saturation [17] and average SFOAE amplitudes seem to compressively grow with increasing level even above 50 dB SPL [18].

CONCLUSION

The presented analytical solution for SFOAEs in a two-dimensional nonlinear cochlear model was validated by comparison with numerical results derived from the same model. The solution provided two components forming the final SFOAE: (1) coherent-reflection component and (2) a component due to perturbation of the nonlinear force. Interestingly, the phases of these two components were at the studied frequencies and levels approximately opposite. Destructive interference between the both components caused saturation of SFOAE amplitude with increasing level and flattened the SFOAE phase.

ACKNOWLEDGMENTS

Supported by an internal grant at the Czech Technical University in Prague SGS20/180/OHK3/3T/13, by the European Regional Development Fund-Project "Center for Advanced Applied Science" (Grant No. CZ.02.1.01/0.0/0.0/16_019/0000778). Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum provided under the programme "Projects of Large Research, Development, and Innovations Infrastructures" (CESNET LM2015042), is greatly appreciated.

REFERENCES

1. C. A. Shera and J. J. Guinan, "Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs," J. Acoust. Soc. Am. **105**, 782–798 (1999).
2. G. Zweig, "Finding the impedance of the organ of corti," J. Acoust. Soc. Am. **89**, 1229–1254 (1991), https://doi.org/10.1121/1.400653.
3. C. A. Shera and G. Zweig, "Noninvasive measurement of the cochlear traveling-wave ratio," J. Acoust. Soc. Am. **93**, 3333–3352 (1993), https://doi.org/10.1121/1.405717.
4. B. Lonsbury-Martin, G. Martin, R. Probst, and A. Coats, "Spontaneous otoacoustic emissions in a nonhuman primate. II. Cochlear anatomy," Hear. Res. **33**, 69–93 (1988).
5. R. Kalluri and C. A. Shera, "Comparing stimulus-frequency otoacoustic emissions measured by compression, suppression, and spectral smoothing," J. Acoust. Soc. Am. **122**, 3562–3575 (2007), https://doi.org/10.1121/1.2793604.
6. C. A. Shera, A. Tubis, and C. L. Talmadge, "Coherent reflection in a two-dimensional cochlea: Short-wave versus long-wave scattering in the generation of reflection-source otoacoustic emissions," J. Acoust. Soc. Am. **118**, 287–313 (2005), https://doi.org/10.1121/1.1895025.
7. A. Vetešník, V. Vencovský, and A. W. Gummer, "Interaction of basilar membrane nonlinearity and roughness in the generation of distortion-product otoacoustic emissions in a two-dimensional cochlear model," J. Acoust. Soc. Am. **submitted** (2022).
8. R. Nobili and F. Mammano, "Biophysics of the cochlea. II: Stationary nonlinear phenomenology," J. Acoust. Soc. Am. **99**, 2244–2255 (1996), http://dx.doi.org/10.1121/1.415412.
9. A. Vetešník and A. W. Gummer, "Transmission of cochlear distortion products as slow waves: A comparison of experimental and model data," J. Acoust. Soc. Am. **131**, 3914–3934 (2012), http://dx.doi.org/10.1121/1.3699207.
10. V. Vencovský, D. Zelle, E. Dalhoff, A. W. Gummer, and A. Vetešník, "The influence of distributed source regions in the formation of the nonlinear distortion component of cubic distortion-product otoacoustic emissions," J. Acoust. Soc. Am. **145**, 2909–2931 (2019a), https://doi.org/10.1121/1.5100611.
11. C. L. Talmadge, A. Tubis, G. R. Long, and P. Piskorski, "Modeling otoacoustic emission and hearing threshold fine structures," J. Acoust. Soc. Am. **104**, 1517–1543 (1998), http://dx.doi.org/10.1121/1.424364.
12. C. L. Talmadge, A. Tubis, G. R. Long, and C. Tong, "Modeling the combined effects of basilar membrane nonlinearity and roughness on stimulus frequency otoacoustic emission fine structure," J. Acoust. Soc. Am. **108**, 2911–2932 (2000), https://doi.org/10.1121/1.1321012.
13. Y.-W. Liu and T.-C. Liu, "Quasilinear reflection as a possible mechanism for suppressor-induced otoacoustic emission," J. Acoust. Soc. Am. **140**, 4193–4203 (2016), https://doi.org/10.1121/1.4968879.
14. V. Vencovský, A. Vetešník, and A. W. Gummer, "Nonlinear reflection as a cause of the short-latency component in stimulus-frequency otoacoustic emissions simulated by the methods of compression and suppression," J. Acoust. Soc. Am. **147**, 3992–4008 (2020), https://doi.org/10.1121/10.0001394.
15. C. Abdala and R. Kalluri, "Towards a joint reflection-distortion otoacoustic emission profile: Results in normal and impaired ears," J. Acoust. Soc. Am. **142**, 812–824 (2017), https://doi.org/10.1121/1.4996859.
16. K. S. Schairer, D. Fitzpatrick, and D. H. Keefe, "Input–output functions for stimulus-frequency otoacoustic emissions in normal-hearing adult ears," J. Acoust. Soc. Am. **114**, 944–966 (2003), https://doi.org/10.1121/1.1592799.
17. R. Sisto, F. Sanjust, and A. Moleti, "Input/output functions of different-latency components of transient-evoked and stimulus-frequency otoacoustic emissions," J. Acoust. Soc. Am. **133**, 2240–2253 (2013), https://doi.org/10.1121/1.4794382.
18. C. Abdala, Y. C. Guardia, and C. A. Shera, "Swept-tone stimulus-frequency otoacoustic emissions: Normative data and methodological considerations," J. Acoust. Soc. Am. **143**, 181–192 (2018), https://doi.org/10.1121/1.5020275.

Relationship between DPOAE and pure tone hearing levels: Numerical analysis with human cochlear finite element model

Takuji Koike[1, a)] and Sinyoung Lee[1, b)]

[1]*Department of Mechanical and Intelligent Systems Engineering*
Graduate School of Informatics and Engineering
The University of Electro-Communications
1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan

[a)] Corresponding author: koike@mce.uec.ac.jp
[b)] sylee@bio.mce.uec.ac.jp

Abstract. Dysfunction of the outer hair cells (OHCs) is thought to lead to hearing loss. However, direct diagnosis of the OHC dysfunction is difficult, and the activity of the OHC is indirectly estimated by measuring distortion product otoacoustic emissions (DPOAEs). In this study, a finite-element model of the normal cochlea considering the activity of the OHC and hearing impairment models in which the activity of the OHC was partially absent were constructed. A pure tone or a complex tone was applied to the models, and the relationship between DPOAE level and non-active portion on the BM was investigated. As a result, it was speculated that the dysfunction of the OHC at the basal side of the position of the characteristic frequency for the input frequency affects pure tone hearing and DPOAE levels.

INTRODUCTION

DPOAEs are considered to be derived from the active nonlinear behavior of the OHCs. However, the generation mechanisms of the DPOAEs have not been fully clarified, and the relationship between cochlear function and DPOAEs is still unclear. The level of the distortion component of $2f_1$-f_2 (DP level) is generally considered to be correlated with hearing level at f_2.[1] However, the relationship between the generated DP components and excitation force of the OHC at each part of the basilar membrane (BM) has not been clarified. In this study, to clarify the mechanism of the DPOAEs, we created four types of the cochlear models, i.e., a passive model, an active model, a model of high-tone sensorineural hearing loss and a model in which the activity of the OHCs is partially absent. Then, by comparing the vibration amplitude of the basilar membrane caused by a pure tone and the distortion component level caused by a complex tone while changing the part where the activity of the OHC is absent, the relationships among partial dysfunction of the OHC, hearing level, and levels of DPOAEs were examined.

METHODS

Geometries and Mechanical Properties of Cochlear Model

Figure 1 shows a FE model of the human cochlea used in this study. This model has been improved based on the previous model[2,3]. The structures and dimensions are the same as that of the passive model previously reported[3], the outline is hence explained briefly here. Although the cochlear takes a spiral form, the shape of the FE model was straightened out. The cochlear model consists of structural parts and liquid parts. The structural parts are the stapes, the stapedial annular ligament, the oval window (OW), the round window (RW), the BM, and the osseous spiral lamina (OSL). The liquid parts are the vestibule, the scala vestibuli (SV), the scala tympani (ST), and the cochlear aqueduct, which are all filled with lymph fluid. To avoid complication of the analysis, the structures in the cochlea were simplified as follows. Although the inner hair cells (IHCs), the OHCs, and the supporting cells exist on the BM, these cells were ignored, and the BM was assumed to be a thin flat membrane. The Reissner's membrane was also ignored

in this model, and the lymph fluid area was divided into two compartments by the BM and the OSL. These two liquid compartments, i.e., the SV and the ST, were connected at the helicotrema. The SV and the ST were considered rectangular ducts, and the outer shape of the cochlea was assumed to change linearly from the basal part to the apex. The mechanical properties of each part of the model were determined based on the previous reports and by conforming numerically obtained results to the values of the measurement results.[4, 5] The material properties of the BM have been modified from those in the previous paper[8] because the BM was assumed to be an orthotropic material in this study. The Young's modulus in the width direction of the BM was assumed to be two times greater than that in the other directions because the collagen fibers run in the direction of the width of the BM.

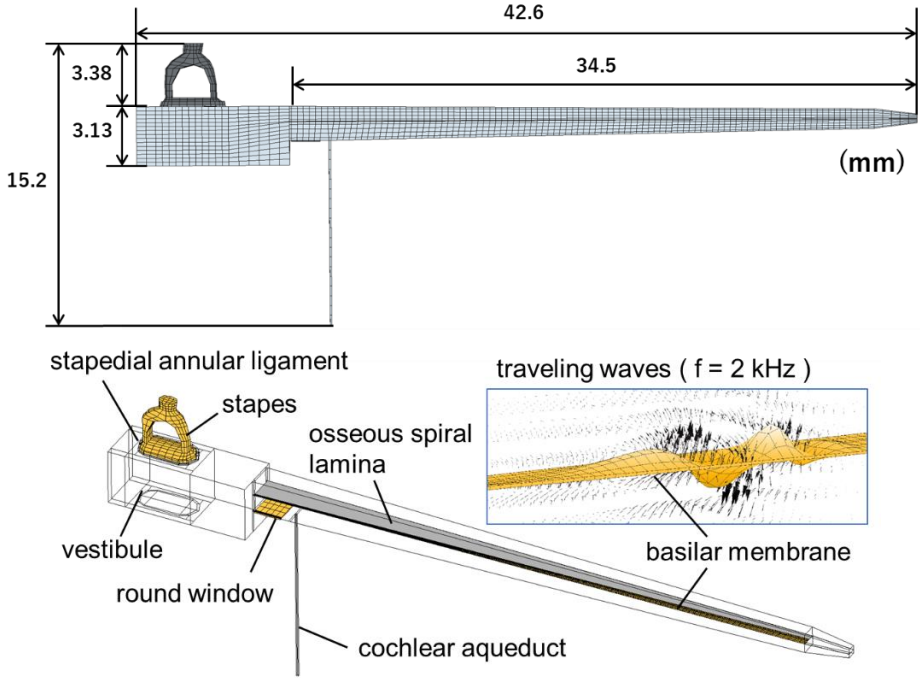


FIGURE 1. Finite-element model of the human cochlea. (a) Liquid parts. (b) Structural parts. The RW membrane and the BM were meshed by four-noded shell elements. Other parts were meshed by eight-noded hexahedral elements.

Representation of OHC Activities

The active cochlear mechanism of amplification related to the electromotility of the OHCs is included by applying additional pressure P_{OHC} on the BM according to its vibration velocity, V_{BM}, as same as the previous paper.[6] The P_{OHC} was expressed by the following equations.

$$\mathbf{P}_{OHC} = \mathbf{a}(z) \cdot \mathbf{b}(z) \cdot \tan^{-1}\{\mathbf{c}(z) \cdot \mathbf{V}_{BM}\} \quad (1)$$

where,

$$\mathbf{a}(z) = 13.5 - 180z, \quad (2)$$

$$\mathbf{b}(z) = \cos(0.436 + 10.9z), \text{ and} \quad (3)$$

$$\mathbf{c}(z) = 7200 - 20400z \quad (4)$$

z is the distance in the longitudinal direction of the BM from the basal end. a(z) shows that the generative force of the OHC at the basal side is smaller than apical side.[7, 8] b(z) shows that the angle between the BM and the OHC is 25° at the base of BM and 45° at the apex of BM[9], and c(z) shows the difference in the saturation curve of the generative force according to the position on the BM. Here, the generative force at the basal side is assumed to be easier to saturate than the apical side[10]. The additional pressure P_{OHC} was applied to the portion of the surface of the BM where the OHCs exist.

Active Areas on the Basilar Membrane

To simulate DPOAEs, four types of the cochlear models were created. One is a "passive model" in which the OHC activities are not included. The second is an "active model." The activity is introduced throughout the whole BM. The third is a model in which the BM is divided into two areas, i.e., active area and passive area, at an appropriate CF position (CFP). Here, the CFP was defined as the position where the traveling wave reached its maximum amplitude when a pure tone of a specific frequency (characteristic frequency, CF) was applied to the stapes head. The active and passive areas on the BM separated by an imaginary boundary line shown in Fig. 2(a). This model may simulate high-tone sensorineural hearing loss (h-HL). The fourth is a model in which the activity of the OHC is removed at a specific area on the BM. In other words, two imaginary boundary lines are defined, and the P_{OHC} is set to zero between the two lines (see Fig. 2(b)). This model may represent sensorineural hearing loss at specific frequencies.

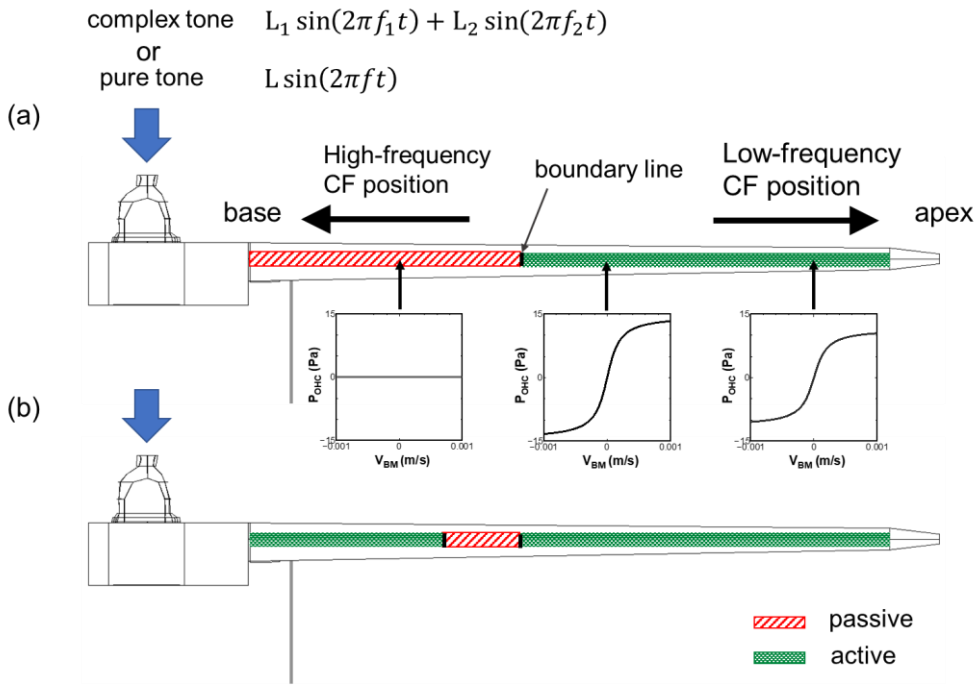


FIGURE 2. Example of partially active cochlear model. (a) Model of high-tone sensorineural hearing loss and functions of P_{OHC} at each point of the BM. (b) Model of sensorineural hearing loss at specific frequencies.

Stimuli Applied to the Model

A pure tone of frequency f or a complex tone consisting of two components with frequencies f_1 and f_2 (f_2 / f_1 =1.2) was applied to the stapes head of the cochlear models, and distortion components generated on the BM and at the stapes footplate were calculated. The sound pressure levels of f_1 and f_2 were 105 dB and 95 dB, respectively. These values correspond to approximately 65 dB and 55 dB at the ear canal when middle-ear pressure gain mainly obtained by the area ratio of the TM to the stapes foot plate is assumed to be 40 dB. These calculations were performed using CFD-ACE+ software, which is one of the commercially available multiphysics analysis software. In this manuscript, f_2 was set to 2 kHz.

RESULTS

Simulation of Pure-Tone Hearing Loss

Figure 3 shows the examples of time course of the BM vibration when pure tone (f = 2 kHz) is applied to the stapes head. The horizontal axis shows the normalized distance from the base to apex of the BM, and the vertical axis shows the amplitude of the BM vibration. In the case of the active (normal) cochlear model, the traveling wave on the BM showed maximum amplitude only at a specific area, and its envelope was spindle-shaped (Fig. 3(a)). In the case of the h-HL model with boundary line at the CFP for f_b = 4 kHz, the different vibration behavior was shown (Fig. 3(b)). Although the BM showed a spindle shape similar to that in the normal model, the maximum amplitude at the CF position was decreased. By contrast, its maximum amplitude position was the same as the normal one.

Figure 4 shows the relationship between the amount of change in the maximum amplitude in the h-HL model relative to that in the normal model and the position of the boundary line (shown by closed circle). The amplitude of the BM rapidly decreased as the position of the boundary line approached the CFP for the input frequency f, and when it moved to the apical side of the CFP for the input frequency, the amplitude became similar to that in the passive model in which the P_{OHC} is set to zero throughout the BM.

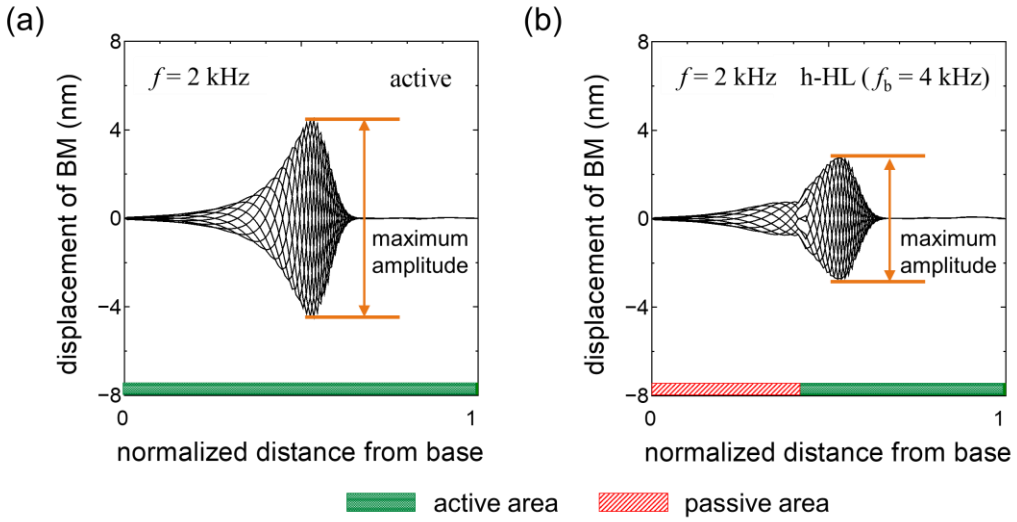


FIGURE 3. Time course of the BM vibration when a pure tone of 2 kHz was applied to the stapes head. (a) Active model. The activity of the OHC is introduced throughout the BM. (b) Model of high-tone sensorineural hearing loss. The boundary line is set at the CFP for 4 kHz.

Simulation of DPOAEs

When the complex tone was applied to the stapes head, distortion components were observed on the BM and at the stapes footplate in the active model. The amount of change in DP level at the stapes footplate of the h-HL model is shown overlaid on Fig. 4 by rhombus-shaped dots. The DP level decreased as the boundary line approaches the CFP for the input frequency f_2 = 2kHz, and the distortion component disappeared when the boundary line moved to the apical side of the CFP for the input frequency.

The amount of change in amplitude in the model of sensorineural hearing loss at specific frequency is shown in Fig. 5. The double-headed arrows indicate the areas in which the activity of the OHC was removed (P_{OHC} = 0). When the passive region exists on the basal side of the CFP for the input frequency f_2 (areas I and II), the DP level decreased. On the other hand, when the passive region exists on the apex side (area III), the DP level was slightly higher than that of the active model. In the cases of I and III, the amount of change in the DP level was less than 3 dB, which was

not significant for a hearing test. By contrast, in the case of II, where the passive range is close to the CFP for the input frequency f_2, the distortion component decreased by about 7 dB.

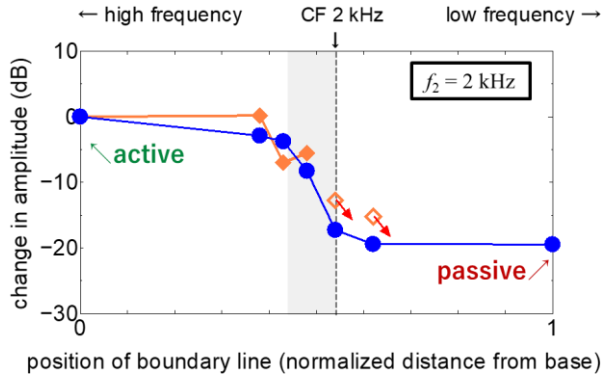


FIGURE 4. Relationship between the amount of change in amplitude in the h-HL model relative to that in the normal model and the position of the boundary line. The circles show the amplitude of the BM. Rhombus-shaped dots show DP level.

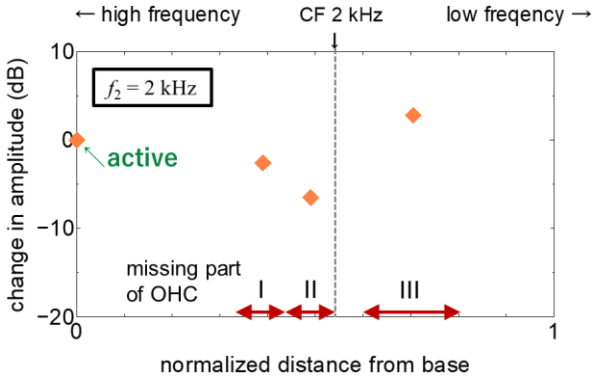


FIGURE 5. Amount of change in DP level in the model of sensorineural hearing loss at specific frequency relative to that in the normal model. The horizontal axis shows the median value of the distance between both end of the area where the activity of the OHC removed. The double-headed arrows indicate the areas in which the activity of the OHC was removed.

DISCUSSION

As shown in Fig. 3, the DP level and the amount of decrease in the BM vibration amplitude when pure tone is applied have similar tendency in the h-HL model. If the change in amplitude of the BM vibraton caused by a pure tone is regarded as the change in pure tone hearing level, the numerical results suggest that there is a relationship between pure tone hearing loss and DPOAE level. When the position of the boundaly line is in the area at the basal side of the CFP for the input frequency of f = 2 kHz (gray area in Fig. 3), both the BM amplitude caused by a pure tone input and the DP level decrease. McFadden et al. have reported in an experiment using mice that the OHC loss on the basal side was significant due to aging, and even if the OHC loss at the CFP for the input frequency was not significant, the hearing threshold of that frequency increased, and DPOAE level also decreased.[11] Our calculation results are consistent with this measurement result. Therefore, it is considered that the DP level and the pure tone hearing level are greatly reduced when the defect area of the OHC is located on the basal side near the CFP for the input frequency. Considerig these things, it is speculated that the amplification mechanism by the OHC in a certain region on the basal side of the CFP for the input frequency affects pure tone hearing and DPOAE level.

Although there are variouse theories concerning DPOAE generation, the following has been proposed. A large distortion component of $2f_1$-f_2 is generated near the CFP (first source) for the input frequency f_2, and the distortion component is divided into a wave traveling toward the basal side and a wave traveling toward the apical side from the first source. The wave traveling toward the apical side is amplified at the CFP for the distortion component of $2f_1$-f_2 (second source), and then, it is reflected toward the basal side. The combined wave of the two waves from the first source and the second source is measured as DPOAE.[19] In the h-HL model, the DP level decreased as the boundary line approached the CFP for the input frequency f_2 (gray area in Fig. 4). Therefore, it is considered that the OHC amplification mechanism in a certain region at the basal side of the CFP for f_2 is the first source. In the case of the model in which a partial OHC defect was created on the apical side of the CFP for the input frequency f_2 (area III in Fig. 5), the DP level is slightly greater than that of the active model even though the OHC function is lost. Since the area III includes the CFP for the distortion component of $2f_1$-f_2, it can be considered that the infuluence of the weve of the distortion component traveling toward the apical side, that is, the influence of the second source is reduced. This might cause a slight increase in PD level at the stapes, but the mechanism is not clear at present. Regarding the source of DPOAE, further studies are needed.

ACKNOWLEDGMENTS

This study was supported by JSPS KAKENHI (Grants No. 19J13684) and Grant-in-Aid from Hirose Foundation.

REFERENCES

1. Gaskill SA, Brown AM, J. Acoust Soc. Am., 88, 821-839 (1990)
2. Koike T, Sakamoto C, Sakashita T, et al., Hear. Res., 283, 117-125 (2012).
3. Lee S, Koike T, Procedia IUTAM, 24, 64-71 (2017).
4. Greenwood DD, J. Acoust. Soc. Am., 87, 2592-2605 (1990).
5. Nakajima HH, Dong W, Olson ES, et al., JARO, 10, 23-36 (2008).
6. Koike T, Lee S, AIP Conference Proceedings, 1965 (2018).
7. Chan E, Ulfendahl M, Acta. Physiol., 161, 533-539 (1997).
8. Santos-sacchi, Kakehata S, Kikuchi T, Katori Y, Takasaka Y, Neurosci., 256, 155-158 (1998).
9. Ramamoorthy S, Deo N. V., Grosh K, J. Acoust. Soc. Am., 121, 2758-2773 (2007).
10. Zinn C, Maier H, Zenner H. P., Gummer W. A., Hear. Res., 142, 159-183 (2000).
11. McFadden SL, Campo P, Quaranta N, et al., Hear. Res., 111, 114-126 (1997).

The Effect of Broadband Elicitor Duration on Transient-Evoked Otoacoustic Emissions and a Psychoacoustic Measure of Gain Reduction

William B. Salloom[1,a)], Kristen Wade[1], Hari Bharadwaj[1,2], and Elizabeth A. Strickland[1]

[1]*Department of Speech, Language, and Hearing Sciences, Purdue University, West Lafayette, Indiana 47907, USA.*
[2]*Weldon School of Biomedical Engineering, Purdue University, West Lafayette, Indiana 47907, USA.*

[a)]Corresponding author: wsalloom@purdue.edu

Abstract. Physiological and psychoacoustic studies of the medial olivocochlear reflex (MOCR) in humans have often relied on long elicitors (> 100-ms). This is largely due to previous research using otoacoustic emissions (OAEs) that found multiple MOCR time constants, including time constants in the 100s of milliseconds, when elicited by broadband noise. However, the effect of the duration of broadband noise elicitors on similar psychoacoustic tasks is currently unknown. Additionally, the relationship between gain reduction measured psychoacoustically and using OAEs as a function of elicitor duration are unknown. The current study measured the effects of ipsilateral broadband noise elicitor duration on transient-evoked OAEs (TEOAEs) and psychoacoustic gain reduction estimated from a forward-masking paradigm. Changes in the TEOAE were measured in terms of magnitude and phase. When phase was accounted for in the TEOAEs, the time constants were approximately equal to the psychoacoustic time constants, and were relatively short (~80 ms). When only changes in TEOAE magnitude were measured, and phase was omitted, the average time constant were longer (~172-ms). Overall, the psychoacoustic and physiological data in the study were consistent with the time course of gain reduction by the MOCR. Near-maximal MOCR effects can be measured with ipsilateral broadband elicitor durations of 200 ms or less in a forward-masking paradigm.

INTRODUCTION

The medial olivocochlear reflex (MOCR) is a bilateral reflex loop at the level of the brainstem that can adjust the gain of the cochlear amplifier in response to sound. The temporal dynamics of the MOCR have been well documented through the use of otoacoustic emissions (OAEs), which have found a combination of fast, medium, and slow time courses when elicited by broadband noise [1, 2]. Both physiological and psychoacoustic studies have typically used continuous or long-duration elicitors (> 100 ms) to maximally stimulate the MOCR. In many studies, the elicitor has been presented contralaterally to the probe or signal ear to avoid excitatory masking of the probe by the elicitor. Additionally, broadband elicitors are known to be strong MOCR elicitors when measuring OAEs. However, Roverud and Strickland (2014) [3], using a psychoacoustic measure of gain reduction, found differential effects of duration for on- and off-frequency tonal elicitors. For the on-frequency elicitor, thresholds increased with increasing duration up to about 50 ms, and then oscillated or plateaued. In contrast, thresholds with off-frequency elicitors continued to increase with elicitor duration. These results are consistent with cochlear gain reduction, possibly by the MOCR, in which the on-frequency elicitor is affected by gain reduction at the signal frequency place, but the off-frequency elicitor is not. DeRoy Milvae and Strickland (2021) [12] found that for a fixed 50-ms elicitor, the amount of gain reduction was only slightly higher for a pink broadband noise elicitor than for an on-frequency tone, when elicitor levels were calculated as the energy that would pass through a filter centered at the signal frequency. However, the effect of the duration of broadband noise elicitors on similar psychoacoustic tasks is currently unknown. Additionally, the relationship between gain reduction measured psychoacoustically and using

OAEs as a function of elicitor duration is unknown. The current study measured the effects of ipsilateral broadband noise elicitor duration on transient-evoked OAEs (TEOAEs) and psychoacoustic gain reduction estimated from a forward-masking paradigm in 19 normal-hearing human listeners.

METHODS

Subjects

Nineteen normal-hearing subjects (7 male and 12 female) completed the experiments in the current study. Their ages ranged from 19 to 35 years (median = 24 years), and all had thresholds ≤ than 15 dB HL at audiometric frequencies from 250-8000 Hz, and normal middle ear function (type A tympanograms). DPOAEs were present from 1500 to 8500 Hz (minimum criteria of -6 dB SPL distortion product and 6 dB SNR for 10 of 12 frequencies tested with no consecutive absent responses).

Psychoacoustic Stimuli and Methods

Psychoacoustic estimates of gain reduction were made using two forward masking techniques that rely on the timing of gain reduction via the MOCR. All stimulus durations include 5-ms $\cos^2$ onset/offset ramps. The signal was a 2-kHz sinusoid, with a duration of 10 ms. For the 'masker present' method, the masker was a 20-ms sinusoid, either at the signal frequency (on-frequency) or at 0.6 times the signal frequency (off-frequency). This masker duration should be too short to elicit the MOCR during the signal presentation [1, 2]. The elicitor was an ipsilateral 50-dB SPL pink broadband noise which was fixed at durations of 50, 65, 100, 200, 400, and 800 ms, which ended just before masker onset. Threshold was first measured for the signal alone. Then the signal was set at 5 dB SL, and on- and off-frequency masker levels determined that would just mask the signal. The maskers were then set at these levels, and signal threshold measured. Then signal threshold was measured with the elicitor preceding the fixed-level maskers. If the elicitor produces excitatory masking, an approximately equal shift in signal threshold should be produced regardless of masker frequency. If the elicitor reduces gain, the shift in signal threshold should be larger following the off-frequency masker than following the on-frequency masker. This is because the on-frequency masker is affected by gain reduction, but the off-frequency masker is not. The 'no-masker' condition was identical to the masker present condition, except that there was a 20-ms gap of silence between the offset of the elicitor and the onset of the signal. This condition was included because previous studies have shown that it gives an estimate of gain reduction that is nearly equal to that with an off-frequency masker [4], and because it is more similar to the paradigm used to measure the TEOAEs. An ipsilateral elicitor was used because it has been shown that gain reduction effects are larger than contralateral gain reduction in humans, at least measured behaviorally [4]. These established methods have been used previously to estimate gain reduction [4]. The equipment, general tracking procedure, and stimuli were generated with the same setup as [4].

Physiological Stimuli and Methods

TEOAE measures were made with an ER-10X Extended-Bandwidth Acoustic Probe System (Etymotic Research), utilizing integrated Forward Pressure Level (FPL) systems. All stimuli were generated with two TDT (Tucker-Davis Technologies) RZ6 programmable DSP processors, amplified by two separate TDT HB7 headphone drivers, and fed to the ER-10X probe. The ER-10X probe was calibrated by determining the Thévenin-equivalent characteristics of the sound sources [5], necessary to determine the acoustic impedance at a location in the ear canal. Thévenin-equivalent characteristics for the probe were estimated by measuring the acoustic response at the ER-10X microphone when the eartip was coupled to loads whose acoustic impedance values can be approximated using theoretical calculations. With the probe calibrated, the immittance properties of each subject's ear canal were checked for potential air-leaks or poor fitting of the probe. Air-leaks can cause changes of absorbance increasingly with the degree of the leak, especially at low-frequencies where air leaks reliably cause an increase of absorbance [6]. A criteria of low-frequency (0.2 kHz) absorbance was used to detect air leaks (less than 29%), or if the admittance phase was greater than 44° (see [6]).

TEOAEs were measured with 54 dB peSPL click probes with a flat incident power spectrum in the 0-10 kHz range. This click level produced emissions significantly higher than the noise floor, and was likely too low to elicit

the MOCR or middle-ear muscle reflex (MEMR). The elicitor was a 50-dB SPL pink broadband noise (0.25-8 kHz), including 5-ms $\cos^2$ onset/offset ramps, which varied in duration (50, 100, 200, and 400-ms) and always preceded the probe click by 5-ms. Both the click and the elicitors were generated with a sampling rate of 48,828.125 Hz. Five different stimulus configurations [click alone (baseline) or click + elicitor] were presented in a sequential block format, equating to a "set", and 16 total sets were collected and stored as a single "array" of data. Eight total arrays of data were collected from each subject, which is ~4-4.5 hours of data collection time, collected over 2-3 lab visits. After the discarding function in the program removed artifacts from each individual block (8 out of 40), 512 responses for each listening condition remained in an array. Thus, completing the experiment (8 arrays) equated to an overall estimate from 4096 responses for each listening configuration. The power spectral density of each TEOAE listening condition was estimated using the mulitaper method in MATLAB (MathWorks, Inc., Natick, MA). The total duration of a single stimulus presentation was always fixed at ~454-ms, which includes 405 msec from the onset of the stimulus window to the click and ~49-ms post-click. Therefore, the only stimuli adjustment between conditions was the onset of the elicitor relative to the onset of the stimulus presentation window, dictated by the elicitor duration.

MOCR effects from TEOAEs were estimated in two ways: 1) the change in the TEOAE magnitude alone ($\Delta TEOAE_m$), and 2) the change in the TEOAE magnitude and phase ($\Delta TEOAE_{m+p}$). This was done by measuring the ear canal sound pressure of the click probe in the click alone and click + elicitor the conditions and taking the vector difference. These metrics have been used in studies measuring stimulus-frequency OAEs (SFOAEs; [7]), and TEOAEs [8], which share a common reflection generation mechanism [9]. MOCR effects were recorded in a 21-ms window (5 to 26 ms post click). Because the psychoacoustic measurement is at a signal frequency of 2 kHz, the ΔTEOAE was calculated for a $1/3^{rd}$ octave band centered at 2 kHz. Lastly, to confirm that the MEMR didn't influence the experiments, an ipsilateral wideband acoustic immittance (WAI) procedure (see [10]) was used to estimate MEMR thresholds before TEOAE or psychoacoustic testing. WAI measures are considered a more sensitive measure of MEMR thresholds compared to typical clinical tympanometric measurements using a single probe tone [11]. Additionally, MEMR-induced absorbed-power changes were obtained concurrently in the TEOAE experiment by time-windowing the measurement to a few samples surrounding the input click. Significant changes to the input click level (0.1 dB) would indicate potential MEMR activation [8]. TEOAE and WAI measures were made using a PC with custom MATLAB software.

Time Constant Estimation

An inverse exponential function was fit to the psychoacoustic (off-frequency masked and masker absent) and TEOAE ($\Delta TEOAE_m$ and $\Delta TEOAE_{m+p}$) data to estimate a time constant. The formula for this function is, $Y(t) = Y_{max}(1 - e^{-t/\tau})$, which estimates ~63% of the maximal effect of the growth function, where 'Y' is the response of interest, 't' is elapsed time, and 'τ' is the time constant. For this study, Y is the signal threshold shift or ΔTEOAE with an elicitor, and 't' equates to the total duration from the onset of the precursor to the onset of the signal in ms (psychoacoustic data, precursor duration + 20; TEOAE data, elicitor duration + 5). In addition to the data points in the time constant estimation, a point was added at zero on the abscissa and ordinate to represent the condition without the precursor (i.e., no threshold shift) as well as to ground the model. Time constants have been estimated with this function in a previous study comparing behavioral and physiological measures of gain reduction [15]. In order for a subject's time constant to be included in the group averaged time constant, at least 60% of the variance had to be accounted for in the individual fit. Additionally, data from subjects with flat functions were excluded in the averaged time constant estimation, as the best fit would have been a time constant near 0 ms.

RESULTS

Ipsilateral Clinical, WAI, and TEOAE MEMR Thresholds

No subject in the study had a clinical ipsilateral MEMR threshold below 50 dB SPL in response to broadband noise, or an ipsilateral WAI MEMR threshold below 50 dB FPL in response to pink noise. The click-windowed responses from the TEOAE experiment indicated that the sound level of the clicks remained constant (< 0.1 dB change) across conditions with the presence of the elicitor. Results from these three separate measures indicate that is highly unlikely that the MEMR was active during the experiments.

Psychoacoustic Data

The average threshold shifts as a function of precursor duration are shown in Fig. 1A. A two-way 3x6 repeated measures analysis of variance (ANOVA) was conducted (IBM SPSS 28) to determine if the mean signal threshold shifts (dependent variable) significantly varied with the independent variables masking type and precursor duration. Aside from data from two subjects (S10 and S16), whose off-frequency masked threshold shifts were considered statistical outliers and therefore excluded from the analysis and overall averages, data from all other subjects met the assumption of normality (N = 17). A statistical outlier was defined as any threshold shift that lay outside 1.5 times the interquartile range. It should be noted that for both S10 and S16, the on-frequency and no-masker data met the assumption of normality, but these data were not included in the ANOVA as having unequal sample sizes would violate the repeated measures test.

Results indicated that there was a significant main effect of masker type [$F(2,32) = 56.27$, $p < 0.001$] on signal threshold shifts. Post hoc comparisons with Bonferroni corrections revealed that threshold shifts for the off-frequency masked conditions were significantly different from on-frequency masked conditions ($p < 0.001$), and similarly, there was a significant difference in threshold shifts between the no-masker and the on-frequency conditions ($p < 0.001$). Lastly, there was a small but significant difference between the off-frequency masked and no-masker conditions ($p = 0.019$). A significant main effect was found for precursor duration [$F(2.823,45.17) = 58.87$, $p < 0.001$], indicating that these shifts generally increased with precursor duration. For this finding, sphericity of the data could not be assumed, and the more conservative Greenhouse-Geisser critical F-value (and degrees of freedom) were reported to help correct for this violation. In summary, threshold shifts following a precursor were significantly larger with off-frequency masker than with the on-frequency masker, across precursor durations. If the precursor caused excitatory masking, the threshold shifts would be approximately equal. The larger effect with the off-frequency masker is consistent with gain reduction [4]. The approximately equal threshold shift with an off-frequency masker and with no masker has been found in previous studies [4], suggests that the no masker condition may also be used as an estimate of gain reduction.

Fig. 1B shows the fitted averaged data for the off-frequency masked conditions (N = 15) and no-masker (N = 19) conditions. The average time constant for the off-frequency masked conditions was 84.68 ± 8.52-ms ($R^2 = 0.91$), and the average time constant for the no-masker conditions was 82.76 ± 8.52-ms ($R^2 = 0.96$). The smaller sample size for

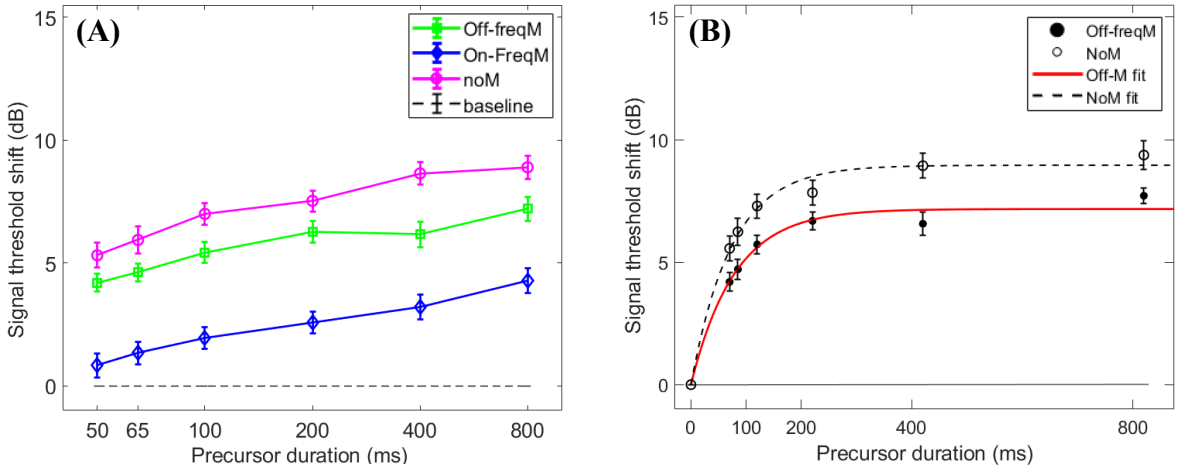


FIGURE 1. Average threshold shifts for the 2-kHz signal as a function of elicitor duration. Horizontal lines at zero in each panel represent the reference condition with no elicitor. Data in (A) indicate that the off-frequency (squares) and threshold shift were significantly larger than the on-frequency as a function of elicitor duration, consistent with gain reduction and not excitatory masking. No masker (circles) threshold shifts were slightly larger than off-frequency masker shifts. Both off-frequency masker and no masker conditions may be used as estimates of the magnitude of gain reduction. A log scale was used on the X axis for visual clarity. For this data, N = 17, as statistical outliers (S10 and S16) were excluded. Data in (B) shows the fitted function to the off-frequency (filled circles; N = 15) and no-masker data (open circles; N = 19), with similar average time constants (~83 - 85-ms). Statistical outliers (S10 and S16) and data that could not be fitted (S6 and S12) were excluded. SEM for the data are indicated by error bars. (Color online).

the off-frequency conditions is because outliers were removed (S10 and S16) as previously mentioned, and data for two other subjects (S6 and S19) were not fit due to an overall flat response with precursor duration. Individual subject time constants can be found in Table 1 (both psychoacoustic and TEOAE data). Overall, the average time constants were approximately equivalent to one another.

TEOAE Data

The fitted average $\Delta TEOAE_m$ is shown in Fig. 2A, and $\Delta TEOAE_{m+p}$ is shown in Fig. 2B. The time constant (N = 15) for the averaged $\Delta TEOAE_m$ was 171.19 ± 45.11-ms ($R^2 = 0.91$). Data for four subjects (1, 3, 10, and 12) could not be fit due to horizontal (and sometimes oscillating) patterns, and thus their data not included in the averaged data. The averaged responses of $\Delta TEOAE_m$ were all within 1 dB for all elicitor durations. The time constant for the averaged $\Delta TEOAE_{m+p}$ data (N = 17, data for two subjects could not be fit) was 78.73 ± 20.75-ms ($R^2 = 0.91$). Note that the averaged $\Delta TEOAE_{m+p}$ responses are larger than $\Delta TEOAE_m$ data. The average $\Delta TEOAE_{m+p}$ time constant was approximately equal to the psychoacoustic experiments. However, the time constant for the average $\Delta TEOAE_m$ data was longer than time constants estimated for the psychoacoustic or $\Delta TEOAE_{m+p}$ data.

DISCUSSION

In this study, the effect of ipsilateral broadband elicitor duration was explored psychoacoustically and physiologically using a forward-masking paradigm. Psychoacoustic results were consistent with gain reduction and not excitatory masking, as has been shown in previous studies, because of the differential effect of the elicitor with on- vs off-frequency maskers. There was quite a bit of variability in the individual time constants (see Table I). For example, time constants in either of the psychoacoustic experiments ranged from ~27 – 152-ms, while the time constants from the $\Delta TEOAE_m$ conditions ranged ~34 – 661-ms, and the $\Delta TEOAE_{m+p}$ ranged from ~22 - 182-ms. This is consistent with results from other studies showing that MOCR effects are variable across subjects [e.g., 7].

Time constants for the psychoacoustic data ranged from 29 to 154 ms, with an average of ~83-85 ms, In Roverud and Strickland (2014) [3], which measured gain reduction by on- and off-frequency tonal precursors as elicitors, the maximal effect for the on-frequency tonal elicitor on signal threshold occurred for the 50-ms precursor, and then plateaued or oscillated with increased precursor duration. In contrast, their off-frequency precursor continued to increase in effectiveness with precursor duration up to 150 ms. In that paper, time constants were fit to the data as part of a model that also included a temporal integration window and a delay in onset of gain reduction. The time constants ranged from approximately 28 to 76 ms. The model was able to predict the oscillation in gain reduction for on-frequency maskers for durations longer than 50 ms.

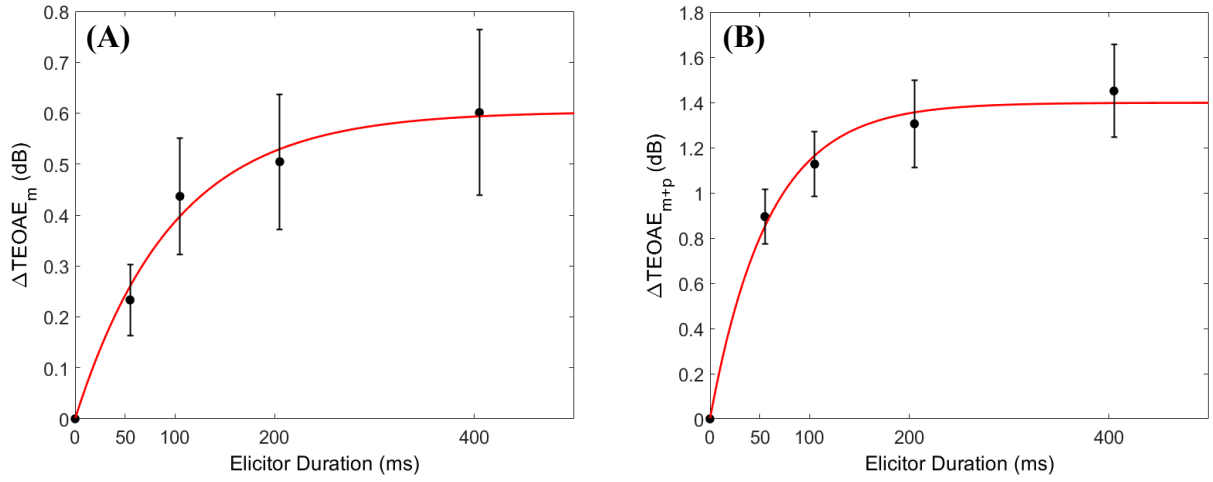


FIGURE 2. Fitted $\Delta TEOAE_m$ (A) and $\Delta TEOAE_{m+p}$ (B) data, as a function of elicitor duration. Note that the scales of the Y axis are different in A and B, with much larger shifts for $\Delta TEOAE_{m+p}$. Only individual data that could be fit with a time constant were included in the averages [(A) N = 15, (B) N = 17)]. SEM is indicated by the vertical error bars.

TABLE 1. Individual subject and averaged time constants with corresponding variance accounted for (R^2) for each condition. τ units are in milliseconds, and the ± error in the average is the SEM of the individual subject time constants. Asterisks indicate that the individual data was not included in the averaged time constant due to a poor R^2 (< 60%), and the '-' symbol indicates that a time constant could not be estimated from the data. 'N' is the subject total used in the average time constant.

Subject	Off-frequency (Avg: N = 15)		No-masker (Avg: N = 19)		$\Delta TEOAE_m$ (Avg: N = 15)		$\Delta TEOAE_{m+p}$ (Avg: N = 17)	
	τ	R^2	τ	R^2	τ	R^2	τ	R^2
1	*91.08*	*0.92*	*125.57*	*0.96*	*66.29**	*0.34**	*24.55*	*0.94*
2	*84.38*	*0.92*	*72.83*	*0.96*	*68.55*	*0.97*	*89.49*	*0.97*
3	*94.47*	*0.96*	*57.05*	*0.96*	*-*	*-*	*44.70*	*0.99*
4	*92.20*	*0.93*	*152.05*	*0.96*	*129.11*	*0.93*	*141.30*	*0.98*
5	*33.52*	*0.97*	*70.10*	*0.96*	*33.75*	*0.93*	*21.68*	*0.91*
6	*-*	*-*	*58.44*	*0.91*	*142.15*	*0.99*	*181.90*	*1.00*
7	*50.61*	*0.85*	*151.59*	*0.89*	*114.76*	*0.97*	*114.60*	*0.99*
8	*52.58*	*0.96*	*59.37*	*0.96*	*123.29*	*0.96*	*62.07*	*0.99*
9	*108.16*	*0.97*	*64.36*	*0.99*	*373.02*	*0.99*	*356.60*	*1.00*
10	*-*	*-*	*70.35*	*0.98*	*-*	*-*	*22.14*	*0.96*
11	*61.69*	*0.97*	*91.10*	*0.97*	*79.82*	*0.99*	*-*	*-*
12	*66.40*	*0.85*	*130.24*	*0.94*	*-*	*-*	*-*	*-*
13	*153.73*	*0.94*	*58.70*	*0.92*	*319.07*	*0.92*	*30.58*	*0.96*
14	*67.72*	*0.89*	*84.01*	*0.94*	*68.44*	*0.65*	*23.69*	*0.91*
15	*93.63*	*0.66*	*63.29*	*0.98*	*96.72*	*0.69*	*37.18*	*0.83*
16	*-*	*-*	*73.31*	*0.94*	*43.29*	*0.87*	*52.16*	*0.99*
17	*144.69*	*0.95*	*32.99*	*0.99*	*660.73*	*0.95*	*27.64*	*0.97*
18	*75.30*	*0.95*	*128.64*	*0.99*	*61.11*	*0.84*	*26.72*	*0.93*
19	*-*	*-*	*28.53*	*0.97*	*126.18*	*1.00*	*81.53*	*0.99*
Avg.	*84.68 ± 8.52*	*0.91*	*82.76 ± 8.52*	*0.96*	*171.19 ± 45.11*	*0.91*	*78.73 ± 20.75*	*0.96*

There are several factors that make it difficult to directly compare results with the current study to the Roverud and Strickland study. Precursor durations in [3] ranged from 10 to 150 ms, while in the present study precursor durations ranged from 50 to 800 ms. For example, the range of precursors used between studies were different and may be a factor in the time constant estimation, and off-frequency data in the Roverud and Strickland study did not always reach a plateau. An important distinction between the two studies is that the current study estimated time constants (63% of Y_{max}), while the other study [3] reported the maximal effect of the on-frequency tonal precursor (the 50-ms precursor). To find the maximal effect with broadband elicitors from our estimates, we can extrapolate (or visually inspect Fig. 1B) to find the duration needed to read Y_{max}, which is ~132-ms. While this varied across subjects, it appears that the duration needed to produce maximum gain reduction in the present study is between the time needed for on- and off-frequency precursors in the Roverud and Strickland study. This is consistent with a recent study [12], which found that broadband precursors had only slightly larger effects on signal threshold compared to on-frequency precursors as a function of precursor level when the precursor duration was fixed in duration (50-ms) and the broadband precursor level was calculated as the energy that should pass through a filter centered at the signal frequency. Since filter bandwidth and suppression may be changing with gain reduction, it might be expected that the effect of duration would be longer for a noise than for a tone elicitor. Overall, the data in the current study are consistent with gain reduction and show that the maximal effects of gain reduction occur within 200 ms of precursor onset for ipsilateral broadband stimulation. This is relatively short compared to the continuous or long duration elicitors that have typically been used to activate the MOCR in the psychoacoustic and physiological literature. Future studies using broadband elicitors may fully activate the MOCR with 150 - 200-ms durations, which could potentially save substantial data acquisition time.

The averaged psychoacoustic time constants were very similar to the averaged physiological time constants when phase was included in the response ($\Delta TEOAE_{m+p}$; ~79 ms), but not when phase was omitted ($\Delta TEOAE_m$; ~172-ms). As explained in the Results section, the $\Delta TEOAE_{m+p}$ produced substantially larger shifts compared to $\Delta TEOAE_m$. A roughly equivalent time constant between the two highly controlled experiments is consistent with a

common underlying mechanism in gain reduction via the MOCR. While the majority of studies have focused on measuring MOCR effects from OAEs by estimating the change in OAE magnitude only, the changes in phase have been less studied, and may provide additional details on MOCR effects in the cochlea [7]. For example, elicitor induced MOCR activation can produce a reduction in the OAE amplitude and/or a phase lead in the OAE, which may be consistent with broadened tuning of the auditory filter [13]. Changes in OAE phase from MOCR activation are still a topic of debate [7, 13], as are the relevance for auditory perception. Understanding how the auditory system adjusts to sound over time has also gained interest in hearing aid and cochlear implant research, as an MOCR inspired gain adjustment feature may improve speech intelligibility in noise, and the benefit seems to vary with the efferent time constant [14]. Further analysis on OAE phase changes with elicitor duration, and individual subject response variability are both future directions of study. Also, further analysis to determine whether physiological responses with elicitor duration can predict psychoacoustic responses, in terms of magnitude and individual time constants, are future directions. More advanced statistical procedures, such as linear mixed models, could be used in the future to account for the individual variation for when using inferential statistics.

ACKNOWLEDGMENTS

Supported by National Institutes of Health National Institute on Deafness and Other Communication Disorders (NIDCD)] Grant Nos. T32 DC016853 (WBS) and RO1 DC008327 (EAS) and R01 DC015989 (HB).

REFERENCES

1. James AL, Harrison RV, Pienkowski, M., Dajani HR, & Mount RJ (2005) Dynamics of real time DPOAE contralateral suppression in chinchillas and humans. International journal of audiology, 44, 118-129
2. Backus BC, & Guinan JJ (2006) Time-course of the human medial olivocochlear reflex. The Journal of the Acoustical Society of America, 119, 2889-2904
3. Roverud E, & Strickland, EA (2014) Accounting for nonmonotonic precursor duration effects with gain reduction in the temporal window model The Journal of the Acoustical Society of America, 135, 1321-1334
4. Salloom WB, & Strickland, EA (2021) The effect of broadband elicitor laterality on psychoacoustic gain reduction across signal frequency. *The Journal of the Acoustical Society of America*, *150*, 2817-2835
5. Allen JB (1986) Measurement of eardrum acoustic impedance. In Allen JB, Hall JL, Hubbard A, Neely ST, and Tubis A, editors, *Peripheral Auditory Mechanisms*, pages 44–51. Springer.
6. Groon KA, Rasetshwane, DM, Kopun JG, Gorga MP, & Neely, ST (2015) Air-leak effects on ear-canal acoustic absorbance. *Ear and hearing*, *36*, 155
7. Lilaonitkul W, Guinan Jr JJ (2012) Frequency tuning of medial-olivocochlear-efferent acoustic reflexes in humans as functions of probe frequency. *Journal of neurophysiology*, *107*, 1598-1611
8. Mertes IB, Goodman SS (2016) Within-and across-subject variability of repeated measurements of medial olivocochlear-induced changes in transient-evoked otoacoustic emissions. *Ear and Hearing*, *37*, e72-e84
9. Shera CA, Guinan JJ (1999) Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs. *The Journal of the Acoustical Society of America*, *105*, 782-798
10. Bharadwaj HM, Hustedt-Mai AR, Ginsberg HM, Dougherty KM, Muthaiah VPK, Hagedorn A, & Heinz, MG (2021) Cross-Species Experiments Reveal Widespread Cochlear Neural Damage in Normal Hearing. [In Press] *Communications Biology. Preprint available on bioRxiv:* https://doi.org/10.1101/2021.03.17.435900
11. Feeney MP, Keefe DH, Hunter LL, Fitzpatrick DF, Garinis AC, Putterman DB., & McMillan GP (2017) Normative wideband reflectance, equivalent admittance at the tympanic membrane, and acoustic stapedius reflex threshold in adults. *Ear and hearing*, *38*, e142
12. DeRoy Milvae, K, & Strickland, EA (2021) Behavioral Measures of Cochlear Gain Reduction Depend on Precursor Frequency, Bandwidth, and Level. *Frontiers in Neuroscience*, *15*.
13. Francis NA, & Guinan JJ (2010) Acoustic stimulation of human medial olivocochlear efferents reduces stimulus-frequency and click-evoked otoacoustic emission delays: Implications for cochlear filter bandwidths. *Hearing Research*, *267*, 36-45
14. Yasin I, Drga V, Liu F, Demosthenous A, & Meddis R. (2020) Optimizing speech recognition using a computational model of human hearing: effect of noise type and efferent time constants. *IEEE Access*, *8*, 56711-56719
15. Walsh KP, Pasanen EG, & McFadden D. (2010). Overshoot measured physiologically and psychophysically in the same human ears. *Hearing research*, *268*, 22-37

Lumped Element Models of Sound Conduction in the Human Ear: a Systematic Review

Irina Wils[1, a)] , Alexander Geerardyn[1, 2] , Tristan Putzeys[1,3], Kathleen Denis[4] and Nicolas Verhaert[1, 2]

[1] *KU Leuven, Department of Neurosciences, Research Group Experimental Oto-rhino-laryngology, Herestraat 49, Bus 721, B-3000, Leuven, Belgium*
[2] *University Hospitals Leuven, Department of Otorhinolaryngology, Head and Neck Surgery, Herestraat 49, B-3000 Leuven, Belgium*
[3] *KU Leuven, Department of Physics and Astronomy, Laboratory for Soft Matter and Biophysics, Celestijnenlaan 200D, Bus 2416, B-3000 Leuven, Belgium*
[4] KU Leuven, Department of Mechanical Engineering, Biomechanics Section, Andreas Vesaliusstraat 13, Bus 2600, B-3000 Leuven, Belgium
[a)] Corresponding author: irina.wils@kuleuven.be

Abstract. Lumped element models (LEMs) have been used to simulate and investigate the fundamental mechanisms of sound conduction in the human ear for decades. This systematic review aims to guide researchers to the most optimal LEM for the parameters under investigation and provide considerations for creating a new model. For this purpose, the literature was reviewed up to March 11, 2022, according to the PRISMA guidelines. We included 6 models via database searching and another 18 models via cross-referencing. The models differ based on six aspects. 1) The number of elements. 2) The type of elements: electrical or mechanical. 3) The parts of the ear: outer, middle or inner ear. 4) The type of conduction: air or bone conduction. 5) The method to determine the elements: reused from previous models, based on anatomical properties, or fit to experimental data. 6) Whether the model predictions are compared to empirical data. The quality of the LEMs is assessed by comparing the predicted middle ear transfer function, the tympanic membrane impedance, the energy reflectance, and the intracochlear pressures (scala vestibuli, scala tympani, and differential) with experimental data. Regarding air conduction, LEMs characterize the pathways from the outer to the inner ear and accurately predict all six aforementioned parameters. This contrasts with the few existing models for bone conduction, where only a part of the ear is modeled. In addition, bone conduction models often excel at predicting one observable parameter in particular, namely the intracochlear pressure. There is a need for complete bone conduction LEMs, incorporating accessible parameters, such as the ear canal pressure, for non-invasive pre-clinical testing and clinical fitting. Moreover, the existing models only simulate one bone conduction pathway, except Stenfelt's model, which simulates five pathways. Furthermore, a model that allows variation of external parameters, such as the implant's position and the stimulation amplitude, enables optimizing implant parameters and may improve the quality of life for people with hearing loss. To conclude, this review provides recommendations to determine the appropriate model for a specific application, including clinical applications.

INTRODUCTION

Sound conduction in the human ear occurs via multiple modes, including air conduction (AC), where air vibrations in the ear canal induce hearing sensation, and bone conduction (BC), where vibrations of the skull induce hearing sensation. AC vibrations are transferred to the basilar membrane via one main pathway. In contrast, BC vibrations are transferred to the basilar membrane via five main pathways: 1) compression of the ear canal, 2) relative motion of the middle ear ossicles due to inertia, 3) compression of the cochlear space, 4) inertia of the cochlear fluid and 5) secondary fluid paths between the inner ear and cranial cavity [1]. Hearing implants

that stimulate the inner ear via these BC pathways are frequently used in hearing rehabilitation. The main indication is conductive hearing loss as these devices bypass the dysfunctional AC pathway and send sound via the skull to the working inner ear [2]. Typically, these hearing implants provide vibrational energy to the skull with an actuator. Although, other hearing implants may also be coupled to specific parts of the ossicular chain or close to the cochlea [3].

Multiple objective measures can quantify hearing sensation and can be used to quantify the performance of implantable and conventional hearing aids. One of them is the middle ear transfer function (METF), the ratio of the stapes velocity to the ear canal pressure, which describes sound transmission through the middle ear. Another possible measure is the tympanic membrane impedance (Z_{TM}) or the pressure to volume velocity ratio at the tympanic membrane [4]. In addition, the energy reflectance (ER), which is the ratio of reflected to incident energy [5], can be measured. Lastly, intracochlear pressures (ICPs) – pressure in the scala vestibuli (P_{SV}), pressure in the scala tympani (P_{ST}), and the difference between them (P_{diff}) – can be used to quantify the performance of hearing aids as they correlate with the cochlear drive [6–10].

Lumped element models (LEMs) are a possible alternative for these measurements. A benefit of such a model is that it does not require cadaveric specimens and thus circumvents practical and ethical issues [11]. Furthermore, in contrast to experiments in cadaveric specimens, LEMs allow testing multiple implants in the same specimen or model; adapting parameters of the implant; estimating velocities and pressures at anatomical locations that are hardly accessible in a specimen; changing the design parameters of an implant continuously; studying the influence of one specific parameter, and investigating the ear without damaging it [12,13]. In addition, it illuminates fundamental mechanisms of sound transmission [14]. There already exist multiple LEMs that characterize sound conduction in the human ear. However, they differ in the type of LEM, parts of the ear that they model, type of conduction they simulate, and their choice of elements. Therefore, this systematic review aims to assess existing LEMs that simulate sound conduction in the human ear. It provides the model predictions, evaluates the quality of the models by comparing the predictions with experimental data, assesses the approaches to create a model, and provides a recommendation on how to create or choose the appropriate model for a specific application.

MATERIALS AND METHODS

On March 11, 2022, Scopus, PubMed and Embase were searched to identify studies that present a lumped element model (LEM) of sound conduction in the human ear. The used search strategy included terms related to LEMs, sound conduction, and the human ear. It resulted in a total of 101 records, and Appendix A provides the detailed search strategy for each database. Figure 1 shows the flow diagram of the selection of relevant records according to the PRISMA guidelines [15].

After removing duplicates, IW and AG independently screened the 52 unique records on title, abstract, and figures using Rayyan [16]. Records were excluded due to the following reasons: the record does not report a LEM (N=13); describes an animal ear (N=10); does not model sound but electrical conduction in the auditory nerves (N=3); has been updated (N=3); or does not model the ear, but the vocal cords, a microphone or an ear simulator (N=3). Disagreements were resolved by consensus-based discussion. Subsequently, full texts were sought for retrieval: three reports were not retrieved. The other 17 texts were assessed for eligibility by IW. Records were excluded if the reported LEM does not predict one of the studied parameters (n=4), is used to evaluate experimental data (n=3), is updated (n=3), or if the record does not report a LEM (n=1). Hence, 6 studies were included via the database search.

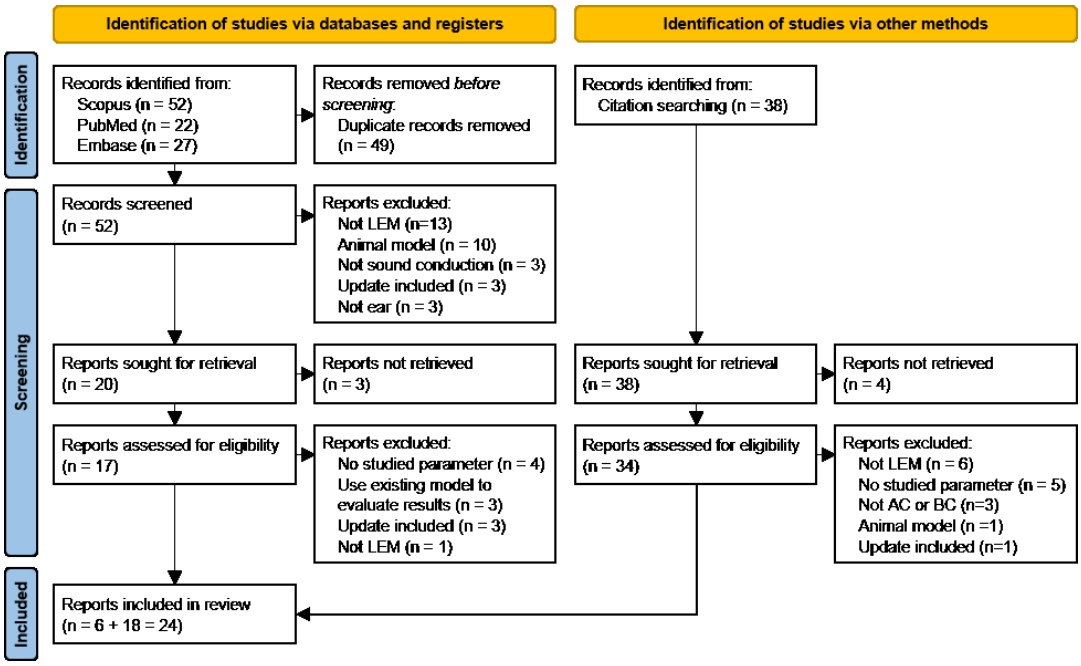


FIGURE 1. Overview of the selection of different articles using the PRISMA flow diagram [15]

Another 38 records were identified via citation searching, of which 34 were retrieved. A total of 16 records were excluded because the record does not report a LEM (n=6); the reported LEM does not predict one of the studied parameters (n=5); it does not model AC or BC (n=3); it models an animal ear (n=1); or it is updated (n=1). Thus, 18 records were included via citation searching, resulting in a total of 24 records included in the review.

To the best of the authors' knowledge, there is no validated tool to assess the risk of bias in LEMs. Therefore, we included all human ear models that predicted either METF, Z_{TM}, ER, or ICPs in this review. Ultimately, the quality of the LEMs is assessed by comparing the predicted METF to experimental data of Aibara et al. [17], Rosowski et al. [18], and Voss et al. [19] since this parameter is used in the ASTM standard to verify the quality of the AC pathway [20]. Furthermore, the Z_{TM} is compared to the ranges of Hudde [21], Rabinowitz [22], and Withnell et al. [23], while the ER is compared to the ranges of Hudde [23], Farmer-Fedor et al. [24], Feeney et al. [25], and Margolis [26]. In addition, the ICPs are compared to currently available experimental data. Regarding AC, these are data of Aibara et al. [17], Grossöhmichen et al. [27], Nakajima et al. [28] and Puria et al. [29]. For BC, these are the data of Fierens and Borgers et al. [30], Mattingly et al. [31], Putzeys et al. [32] and Stieger et al. [6].

For most records, the LEM is reproduced in Matlab (MathWorks, Massachusetts, USA) to obtain the predicted pressures and velocities. In contrast, we used reported pressures and velocities for the models of Keefe [33], O'Connor et al. [34], Stenfelt [35], and Xue et al. [4]. Regarding the model of Stenfelt [35], the model prediction for P_{SV} are provided at two positions, further referred to as A and B. The models of Hudde et al. [36], Peake et al. [37], Onchi [38], Voss et al. [39], and Withnell et al. [23] are not included in the comparison because they are not reproducible with the information presented in the paper, and they did not report the predicted METF, Z_{TM}, ER or ICPs. Neither is the model of Elliott et al. [40] since it does not model ear canal pressure. Regarding the

model of Rosowski et al. [41], only results for AC are included in the comparison with experimental data as the results of BC are not normalized to promontory velocity.

RESULTS

Figure 2 shows a schematic view of three LEMs, namely the LEM developed by Feng et al. [42] (further referred to as model 1), by Stenfelt [35] (further referred to as model 2), and by O'Connor et al. [34] (further referred to as model 3). A schematic view of all models is shown in Appendix B. Further, the following paragraphs and Table 1 compare the LEMs of the reviewed records, sorted by the number of elements used in the models.

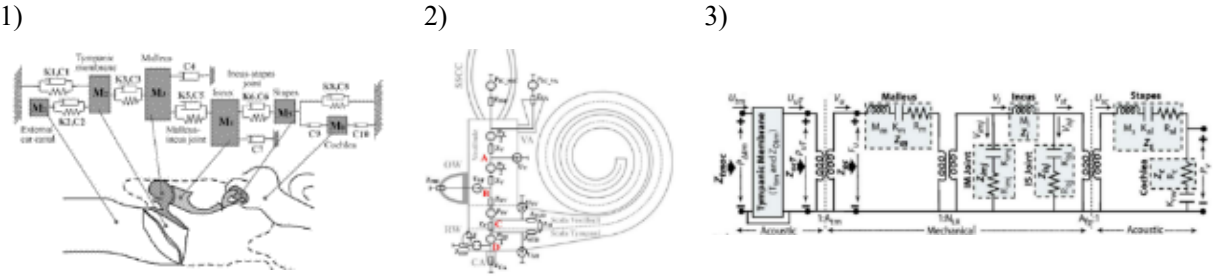


FIGURE 2. Schematic overview of the LEMs. 1) LEM with mechanical elements of the middle ear for AC [42]. 2) LEM with electrical elements of the inner ear for AC and BC [35]. 3) LEM with electrical elements of the whole ear for AC [34].

Regarding the types of LEM, some models use a mechanical representation, such as model 1 in Figure 2, while others use an electrical representation of elements, such as models 2 and 3. Yet, others use both electrical and mechanical elements. For example, the model of Keefe [33] uses electrical elements to model the outer and inner ear, while it uses mechanical elements for the middle ear. Also, Onchi [38] reports both a model with mechanical elements and a transformation to a model with electrical elements. As it is possible to transform mechanical elements into electrical and vice versa, neither is better. However, it is more intuitive to model pressures and volume velocities in the outer and inner ear with electrical elements, and forces and linear velocities in the middle ear with mechanical elements.

The LEMs also describe different parts of the ear. For example, model 1 describes the middle ear, model 2 describes the inner ear, and model 3 describes all three parts of the ear. Models that only describe a part of the ear in detail often use a Thévenin equivalent for the other parts, for example, the middle ear impedance (Z_{ME}) in model 2. This impedance enables normalizing the modeled ICPs to ear canal pressure. However, as this is only one impedance, it does not allow discrimination of middle ear ossicles. Depending on the purpose of the model, this can be sufficient, but if multiple parts of the ear fall within the research scope, a complete model of the whole ear is needed.

Every LEM simulates the AC pathway, but only some model BC and from those, only the one by Stenfelt [35] models all five main pathways. The models of Guan et al. [10], Hudde et al. [36], and Rosowski et al. [41] simulate, respectively, inertia of the inner ear fluid, inertia of the middle

[1)] Reprinted by permission from Springer Nature: Springer Nature Biomech. Model. Mechanobiol. Lumped parametric model of the human ear for sound transmission, Feng and Gan, © 2004
[2)] Reprinted from Stenfelt, Front. Neurol., 11, 966 (2020). © 2020 Stenfelt S.
[3)] Reprinted with permission from Xue, Liu, Wang, Yang, Zhao, and Huang, J. Acoust. Soc. Am., 147, 1478–1490 (2020). © 2020, Acoustic Society of America.

TABLE 1. Comparison of the models. Number of elements: total number of elements used in a model. Type LEM: whether the model consists of electrical or mechanical elements. Part ear: whether the LEM simulates the outer, middle or inner ear. Type conduction: whether the LEM simulates AC or BC. Choice elements: whether elements are reused from previous models, based on anatomical properties, or fit to experimental data. Compared with data: whether the model predictions are compared to experimentally obtained data. ✓: included in the model. Ø: not included in the model. ✓*: Modeled with Thévenin equivalent.

Author	Number of elements	Type LEM		Part Ear			Type conduction		Choice elements			Compared with experimental data
		Electrical	Mechanical	Outer	Middle	Inner	AC	BC	Previous model(s)	Anatomy based	Fit to data	
Liu et al. [43]	7	Ø	✓	✓	✓	✓*	✓	Ø	✓	✓	✓	Ø
Voss et al. [39]	7	✓	Ø	✓	✓	✓*	✓	Ø	✓	✓	✓	✓
Hudde et al. [36]	8	✓	Ø	✓	✓	✓*	✓	✓	✓	✓	✓	✓
Peake et al. [37]	9	✓	Ø	✓	✓	✓*	✓	Ø	✓	✓	✓	✓
Frear et al. [14]	10	✓	Ø	✓*	✓*	✓	✓	Ø	Ø	✓	✓	✓
Guan et al. [10]	10	✓	Ø	✓*	✓*	✓	✓	✓	✓	✓	✓	✓
Møller et al [44]	10	✓	Ø	✓*	✓*	✓*	✓	Ø	Ø	✓	✓	✓
Rosowski et al. [41]	11	✓	Ø	✓*	✓*	✓	✓	✓	✓	Ø	✓	Ø
Elliott et al. [40]	13	✓	Ø	Ø	✓*	✓	✓	Ø	✓	✓	Ø	✓
Marquardt et al. [45]	13	✓	Ø	✓*	✓*	✓	✓	Ø	✓	Ø	✓	✓
O'Connor et al. [34]	14	✓	Ø	✓	✓	✓	✓	Ø	✓	✓	✓	✓
Zwislocki [46]	18	✓	Ø	✓	✓	✓*	✓	Ø	Ø	Ø	✓	✓
Kringlebotn [47]	20	✓	Ø	✓	✓	✓*	✓	Ø	✓	✓	✓	✓
Lutman et al. [48]	20	✓	Ø	✓	✓	✓*	✓	Ø	✓	✓	✓	✓
Stenfelt [35]	20	✓	Ø	✓*	✓*	✓	✓	✓	✓	✓	Ø	✓
Onchi [38]	21	✓	✓	✓	✓	✓	✓	Ø	✓	Ø	✓	✓
Feng et al. [42]	22	Ø	✓	✓	✓	Ø	✓	Ø	✓	✓	✓	✓
Rosowski et al. [49]	22	✓	Ø	✓	✓	✓*	✓	Ø	✓	Ø	✓	✓
Xue et al. [4]	24	✓	Ø	✓	✓	✓	✓	Ø	✓	Ø	Ø	✓
Pascal et al. [13]	25	✓	Ø	✓	✓	✓*	✓	Ø	✓	✓	Ø	✓
Withnell et al. [23]	25	✓	Ø	✓	✓	✓*	✓	Ø	✓	✓	Ø	✓
Goode et al. [50]	27	✓	Ø	✓	✓	✓*	✓	Ø	✓	✓	✓	✓
Keefe [33]	32	✓	✓	✓	✓	✓	✓	Ø	✓	✓	✓	✓
Giguère et al. [51]	560	✓	Ø	✓	✓	✓	✓	Ø	✓	✓	✓	Ø

ear ossicles, and compression of the cochlear walls. To the best of the authors' knowledge, no LEM simulates BC from actuator to ICP and for all five main pathways.

The reports used multiple strategies to determine elements and their values. Except for Frear et al. [14] and Zwislocki [46], all authors reused some elements of one or more previous models. Another strategy to determine elements and their value is to base them on anatomical data - for example, some LEMs use area of the tympanic membrane or mass of the stapes. In addition, most researchers also optimized values of some elements such that the predicted pressures and velocities fit experimentally obtained data.

Almost all studies (21 out of 24) compared some of their final LEMs predicted impedances, pressures, or velocities with experimentally obtained data. Appendix C gives a complete overview

of the reported comparisons between predicted and experimental data. In brief, most reports compare with experimental data from the literature, but some reports use data obtained by their research group. In addition, some reports only compared one specific impedance, pressure, or velocity [23,36–40,44,46,48], while others compared up to seven different values [14]. Furthermore, the sample size of the experimental data varies between 3 [36,38] and 68 [47,49] specimens or subjects.

DISCUSSION

To assess the quality of the included LEMs, this section compares their predicted METF, Z_{TM}, ER, or to experimental data available in the literature. Furthermore, this section discusses clinical applications, considerations when creating or choosing a LEM , and limitations of this systematic review.

Quality Assessment

To assess the quality of the LEMs, the fit between their predictions and the experimental data from the literature is evaluated in Table 2. Figures D.1 and D.2 show the frequency ranges for which the predictions are within the experimental ranges, while Tables D.1 and D.2 provide the values. In addition, Figures E.1, E.2 and E.3 show the predicted parameters together with the experimental data.

TABLE 2. Evaluation of the fit between model predictions and experimental data, based on the percentage of the predictions that lie within the experimental data on a logarithmic scale. +++: >90%, ++: >75%, +: >50%, ---: <10%, --: <25%, -: <50%, Ø: the model does not predict the parameter. METF: middle ear transfer function, Z_{TM}: tympanic membrane impedance, ER: energy reflectance, AC: air conduction, BC: bone conduction, P_{SV}: scala vestibuli pressure, P_{ST}: scala tympani pressure, P_{diff}: differential pressure.

Author	$\lvert METF \rvert$	$\angle METF$	$Re(Z_{TM})$	$Im(Z_{TM})$	$\lvert ER \rvert$	$\angle ER$	$AC: \lvert P_{SV} \rvert$	$AC: \angle P_{SV}$	$AC: \lvert P_{ST} \rvert$	$AC: \angle P_{ST}$	$AC: \lvert P_{diff} \rvert$	$AC: \angle P_{diff}$	$BC: \lvert P_{SV} \rvert$	$BC: \angle P_{SV}$	$BC: \lvert P_{ST} \rvert$	$BC: \angle P_{ST}$	$BC: \lvert P_{diff} \rvert$	$BC: \angle P_{diff}$
Liu et al. [43]	+	-	-	--	+++	-	+	-	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
Frear et al. [14]	+++	+	Ø	Ø	Ø	Ø	++	-	+	--	+++	--	Ø	Ø	Ø	Ø	Ø	Ø
Guan et al. [10]	+++	+	Ø	Ø	Ø	Ø	-	-	+	--	---	--	++	+	++	+	++	+
Møller [44]	Ø	Ø	++	+	-	---	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
Rosowski et al. [41]	Ø	Ø	---	-	+++	-	-	+	+++	-	--	--	Ø	Ø	Ø	Ø	Ø	Ø
Marquardt et al. [45]	+++	+	Ø	Ø	Ø	Ø	+++	+	+++	---	+++	-	Ø	Ø	Ø	Ø	Ø	Ø
O'Connor et al. [34]	+++	++	++	+	-	--	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
Zwislocki [46]	+++	+	-	+	--	---	+	-	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
Kringlebotn [47]	+++	+++	++	++	---	--	+++	++	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
Lutman et al. [48]	+++	+	+++	+	---	-	+	-	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
Stenfelt A [35]	Ø	Ø	Ø	Ø	Ø	Ø	+++	Ø	+	Ø	Ø	Ø	++	Ø	++	Ø	++	Ø
Stenfelt B [35]	Ø	Ø	Ø	Ø	Ø	Ø	+++	Ø	+	Ø	Ø	Ø	++	Ø	++	Ø	+	Ø
Feng et al. [42]	++	+++	+++	-	+	--	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
Rosowski et al. [49]	+++	++	++	++	---	--	+++	+	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
Xue et al. [4]	+++	+++	++	+	-	-	+++	+++	Ø	Ø	+++	-	Ø	Ø	Ø	Ø	Ø	Ø
Pascal et al. [13]	+	+	+++	+	---	---	+	+	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
Goode et al. [50]	++	+++	-	-	+	-	+	+	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
Keefe [33]	+++	+++	++	+	--	---	+++	+++	+++	+	+++	+	Ø	Ø	Ø	Ø	Ø	Ø
Giguère et al. [51]	+	-	+	+	---	---	Ø	Ø	Ø	Ø	++	+	Ø	Ø	Ø	Ø	Ø	Ø

The METF, the ratio of stapes velocity to ear canal pressure, is a descriptor of sound transmission through the middle ear [52]. According to the ASTM standard, the AC pathway is of good quality if the METF is within this range Rosowski et al. [18] for frequencies between 0.25 and 4 kHz [20]. Most predicted METFs lie within the range of the experimentally obtained data. However, the METF of Liu et al. [43] lies below the range and lacks the typical configuration. This can be caused by the fact that the model was only optimized such that predicted P_{SV} fitted to experimental data. Similarly, the model of Pascal et al. [13] focused on Z_{TM} and P_{SV}, which explains why its METF lies above the experimental ranges at frequencies below 400 Hz. Also, the model of Giguère et al. [51], where the original paper did not compare its predictions with experimental data, lies above the experimental ranges for frequencies between 440-1010 and 3760-5280 Hz. Regarding the phase of the METF, most models follow the trend of the experimental data, except the one from Liu et al. [43], which also does not follow the trend for the amplitude. The best fit for the phase is found in the models of Kringlebotn [47], Feng et al. [42], Xue et al. [4], Goode et al. [50], and Keefe [33]. These models have in common that they all modeled the external and middle ear in detail and fitted their model based on experimental data.

All impedances from the tympanic membrane to the inner ear determine the Z_{TM} or the ratio of pressure to volume velocity at the membrane. Therefore, it is useful to assess the quality of the LEMs based on this property. Some model predictions are below the ranges (Keefe [33], and Giguère et al. [51]), while another lies above the ranges (Goode et al. [50]), and even others show predictions above and below the ranges depending on the frequency (Liu et al. [43], Rosowski et al. [41], and Feng et al. [42]). In addition, some models show a resonance peak, while this is not visible in the experimental data (Møller [44], O'Connor et al. [34], Zwislocki [46], and Xue et al. [4]). The best fit is found for the models of Kringlebotn [47], Lutman et al. [48], Rosowski et al. [49], Xue et al. [4] and Pascal et al. [13]. These models have in common that they have at least 20 elements and that they model all three parts of the ear.

The energy reflectance is the ratio of the reflected to incident energy and is calculated as follows:

$$ER = \left(\frac{Z_{TM} - Z_{char}}{Z_{TM} + Z_{char}}\right)^2 \; with \; Z_{char} = \frac{\rho \cdot c}{S_e}$$

With Z_{char} the characteristic impedance of the ear canal, ρ the equilibrium air density, c the velocity of sound, and S_e the cross-sectional area of the ear canal. Most models show a high reflectance at the low and high frequencies and a low reflectance at the mid frequencies, similar to the experimental data. Nevertheless, the predicted ERs are in general lower than the experimental data. Only the model of Liu et al. [43] shows a reflectance that lies within the ranges and follows the same trend as the ranges. Regarding the phase, experimental data are only available for frequencies above 1 kHz. The model of Liu et al. [43] also lies in this phase range.

ICPs are a promising measurement tool to evaluate cochlear drive and thus hearing sensation. Aibara et al. [17], Grossöhmichen et al. [27], Nakajima et al. [53], and Puria et al. [29] measured these pressures in human cadaveric specimens for AC and normalized them to ECP. Amplitudes of the pressures predicted by most models follow the trend of the experimental data. However, the phase differs for many models. This is because most models are optimized to fit amplitude but not phase. What is also noteworthy is that the amplitudes of the P_{SV} and P_{diff} predicted by the LEMs of Guan et al. [10] and Rosowski et al. [41] are above the experimental ranges. This is potentially caused by their focus on BC. Further, the LEM of Keefe [33] is the only one that models P_{SV}, P_{ST}, and P_{diff}, and for which all pressures are between the experimental ranges. This can be explained

by the fact that it models all three parts of the ear, has many elements (32), uses a combination of the three fitting strategies, and in contrast to other LEMs, it also includes an impedance for middle ear cavities.

Similar to AC, ICPs are also a promising tool to evaluate cochlear drive and hearing sensation for BC. Here, pressures are normalized to promontory velocity (V_{prom}). Again, the predicted data fit well within the experimental ranges. Regarding the model of Stenfelt [35], the fit to Pdiff improves when P_{SV} is predicted at position A rather than B. A limitation of both models is that they only simulate the inner ear and not the actuator to the inner ear. Furthermore, the model of Guan et al. [10] only simulates one of the five main pathways, but can predict the amplitudes and the phases of the three pressures. In comparison, the one from Stenfelt [35] simulates the five main pathways, but the combination of the five pathways is determined by the root of the summed squares, thus eliminating the phase information.

Clinical Applications of LEMs

LEMs can be used to investigate multiple clinical applications, such as changes in hearing sensitivity caused by clinical pathologies or hearing aids. Table 3 gives an overview of these applications and the LEMs that simulate them.

TABLE 3. Clinical applications and LEMs that simulate them.

Clinical application	LEMs
Tympanic membrane perforation	*Voss et al. [39], Goode et al.* [50]
Otosclerosis	*Goode et al.* [50]*, Zwislocki et al.* [46]
Round window reinforcement	*Elliott et al. [40]*
Superior canal dehiscence and third window effect	*Rosowski et al. [41], Stenfelt* [35]*, Xue et al. [4]*
Round window stimulation	*Frear et al. [14], Xue et al. [4]*

Considerations when Creating a New Model or Determining the Appropriate Existing Model for a Specific Application

Depending on the model's purpose, some characteristics can be beneficial or unfavorable. For example, including all ear parts makes a model more complete but increases computational complexity. Thus, when investigating only a part of the ear, excluding the other parts or replacing them with a Thévenin equivalent is useful. Similarly, a large number of elements increases both accuracy and complexity. Also, when only one type of stimulation is required, it is preferable to only simulate that type. Regarding bone conduction, if modeling one or a few pathways results in an accurate simulation, omitting pathways can be considered to reduce the complexity. Further, if elements exist in previous models, they can be reused. Otherwise, they can be determined by anatomical properties or optimized such that a fit is found between predicted and experimental measures. This measure should be chosen depending on the application.

To determine the appropriate existing model, the model should be able to predict the parameters under investigation. This can be verified in Figures 3 and 4 or Appendix D, which gives the exact values of the frequency ranges. If two models predict the parameters equally well, choosing the model with the least elements is recommended to reduce the computational cost. Hence, the models in the figures are sorted by the number of elements. Lastly, if the application includes a pathology, choosing a model that includes this pathology is recommended, as mentioned above.

Limitations

This review compared the METF, Z_{TM}, ER and ICPs - values often used to evaluate hearing sensation - of the LEMs with experimentally obtained data. However, the models can also predict other characteristics, which are not discussed in this review. Examples of these characteristics are cochlear impedance (ratio of pressure to volume velocity in the scala vestibuli) [33], ICPs relative to stapes velocity, and mechanical impedance of the mastoid (ratio of output force from a BC actuator to velocity at the mastoid) [54]. Therefore, when a LEM cannot accurately predict the METF, Z_{TM}, ER or ICPs, it possibly predicts these other characteristics accurately and vice versa.

Further, this review evaluated LEMs based on the results for non-pathologic ears, while some models also simulate pathologic ears [4,35,39–41,46,50]. In addition, this review only evaluates AC and BC, while models that simulate other stimulation types, such as round window stimulation [4,14] and mechanical stimulation of the middle ear [55], exist too. Lastly, this review did not include animal models, although they resemble human models [8,56–58].

CONCLUSION

This systematic review provides an overview of LEMs that simulate sound conduction in the normal human ear. Regarding AC, some LEMs simulate the ear from outer to inner ear and can accurately predict the METF, Z_{TM}, ER and ICPs. Regarding BC, however, LEMs only model a part of the ear, and most of them only model one of the five main pathways. Thus, a LEM that models BC from actuator to inner ear is still lacking and would increase insights into the human ear sound conduction pathways as well as the functioning of BC implants.

ACKNOWLEDGMENTS

The authors are grateful to Nathaniel Green and Christof Stieger for sharing their experimental data, Stefan Stenfelt for discussing and sharing data from his LEM, and John Rosowski for the critical review.

IW acknowledges funding by Flanders Innovation and Entrepreneurship (HBC.2020.2201), AG and NV by Research Foundation – Flanders (FWO 1SD3322N and FWO 1804816N).

REFERENCES

1. Stenfelt, *Hear. Res.*, **340**, 135–143 (2016) http://dx.doi.org/10.1016/j.heares.2015.10.014.
2. Reinfeldt, Taghavi, Eeg-Olofsson, and Hakansson, *Med. Devices Evid. Res.*, **8**, 79 (2015) http://www.dovepress.com/new-developments-in-bone-conduction-hearing-implants-a-review-peer-reviewed-article-MDER.
3. Verhaert, Desloovere, and Wouters, *Otol. Neurotol.*, **34**, 1201–1209 (2013) https://journals.lww.com/00129492-201309000-00006.
4. Xue, Liu, Wang, Yang, Zhao, and Huang, *J. Acoust. Soc. Am.*, **147**, 1478–1490 (2020) http://asa.scitation.org/doi/10.1121/10.0000846.
5. Voss, and Allen, *J. Acoust. Soc. Am.*, **95**, 372–384 (1994) http://asa.scitation.org/doi/10.1121/1.408329.
6. Stieger, Guan, Farahmand, Page, Merchant, Abur, et al., *J. Assoc. Res. Otolaryngol.*, **19**, 523–539 (2018) http://link.springer.com/10.1007/s10162-018-00684-1.
7. Niesten, Stieger, Lee, Merchant, Grolman, Rosowski, et al., *Audiol. Neurotol.*, **20**, 62–71 (2015) https://www.karger.com/Article/FullText/366512.
8. Dancer, and Franke, *Hear. Res.*, **2**, 191–205 (1980) https://linkinghub.elsevier.com/retrieve/pii/037859558090057X.
9. Borgers, van Wieringen, D'hondt, and Verhaert, *AIP Conf. Proc.*, **1965**, 050005 (2018) http://aip.scitation.org/doi/abs/10.1063/1.5038471.
10. Guan, Cheng, Galaiya, Rosowski, Lee, and Nakajima, *Sci. Rep.*, **10**, 16564 (2020) http://www.nature.com/articles/s41598-020-73565-4.
11. Chang, Kim, and Stenfelt, *J. Acoust. Soc. Am.*, **140**, 1635–1651 (2016) http://dx.doi.org/10.1121/1.4962443.
12. Chang, and Stenfelt, *Trends Hear.*, **23**, 233121651983605 (2019) http://journals.sagepub.com/doi/10.1177/2331216519836053.
13. Pascal, Bourgeade, Lagier, and Legros, *J. Acoust. Soc. Am.*, **104**, 1509–1516 (1998) http://asa.scitation.org/doi/10.1121/1.424363.
14. Frear, Guan, Stieger, Rosowski, and Nakajima, *Hear. Res.*, **367**, 17–31 (2018) https://doi.org/10.1016/j.heares.2018.06.019.
15. Page, McKenzie, Bossuyt, Boutron, Hoffmann, Mulrow, et al., *BMJ*, **372**, n71 (2021)

https://www.bmj.com/lookup/doi/10.1136/bmj.n71.
16. Ouzzani, Hammady, Fedorowicz, and Elmagarmid, *Syst. Rev.*, **5**, 210 (2016) http://dx.doi.org/10.1186/s13643-016-0384-4.
17. Aibara, Welsh, Puria, and Goode, *Hear. Res.*, **152**, 100–109 (2001) https://linkinghub.elsevier.com/retrieve/pii/S0378595500002409.
18. Rosowski, Chien, Ravicz, and Merchant, *Audiol. Neurotol.*, **12**, 265–276 (2007) https://www.karger.com/Article/FullText/101474.
19. Voss, Rosowski, Merchant, and Peake, *Hear. Res.*, **150**, 43–69 (2000) https://linkinghub.elsevier.com/retrieve/pii/S0378595500001775.
20. ASTM, (2014) www.astm.org.
21. Hudde, *J. Acoust. Soc. Am.*, **73**, 242–247 (1983).
22. Rabinowitz, *J. Acoust. Soc. Am.*, **70**, 1025–1035 (1981) http://asa.scitation.org/doi/10.1121/1.386953.
23. Withnell, and Gowdy, *J. Assoc. Res. Otolaryngol.*, **14**, 611–622 (2013) http://link.springer.com/10.1007/s10162-013-0407-y.
24. Farmer-Fedor, and Rabbitt, *J. Acoust. Soc. Am.*, **112**, 600–620 (2002) http://asa.scitation.org/doi/10.1121/1.1494445.
25. Feeney, Grant, and Mills, *Ear Hear.*, **30**, 391–400 (2009) https://journals.lww.com/00003446-200908000-00002.
26. Margolis, Saly, and Keefe, *J. Acoust. Soc. Am.*, **106**, 265–280 (1999) http://asa.scitation.org/doi/10.1121/1.427055.
27. Grossöhmichen, Salcher, Püschel, Lenarz, and Maier, *Biomed Res. Int.*, **2016**, 1–10 (2016) http://www.hindawi.com/journals/bmri/2016/6059479/.
28. Nakajima, Ravicz, Merchant, Peake, and J. Rosowski, *Hear. Res.*, **204**, 60–77 (2005) https://linkinghub.elsevier.com/retrieve/pii/S0378595505000195.
29. Puria, Peake, and Rosowski, *J. Acoust. Soc. Am.*, **101**, 2754–2770 (1997) http://asa.scitation.org/doi/10.1121/1.418563.
30. Fierens, Borgers, Putzeys, Walraevens, Van Wieringen, and Verhaert, *Biomed Res. Int.*, **2022**, 1–16 (2022) https://www.hindawi.com/journals/bmri/2022/9079903/.
31. Mattingly, Banakis Hartl, Jenkins, Tollin, Cass, and Greene, *Ear Hear.*, **41**, 312–322 (2020) https://journals.lww.com/00003446-202003000-00010.
32. Putzeys, Borgers, Fierens, Walraevens, Van Wieringen, and Verhaert, *Hear. Res.*, (2022).
33. Keefe, *J. Acoust. Soc. Am.*, **137**, 2698–2725 (2015) http://dx.doi.org/10.1121/1.4916592.
34. O'Connor, and Puria, *J. Acoust. Soc. Am.*, **123**, 197–211 (2008) http://asa.scitation.org/doi/10.1121/1.2817358.
35. Stenfelt, *Front. Neurol.*, **11**, 966 (2020) https://www.frontiersin.org/article/10.3389/fneur.2020.00966/full.
36. Hudde, and Weistenhöfer, *Funct. Mech. Norm. Dis. Reconstr. Middle Ears*, 39–58 (2000).
37. Peake, Rosowski, and Lynch, *Hear. Res.*, **57**, 245–268 (1992) https://linkinghub.elsevier.com/retrieve/pii/037859559290155G.
38. Onchi, *J. Acoust. Soc. Am.*, **33**, 794–805 (1961) http://asa.scitation.org/doi/10.1121/1.1908801.
39. Voss, Rosowski, Merchant, and Peake, *J. Acoust. Soc. Am.*, **110**, 1445–1452 (2001) http://asa.scitation.org/doi/10.1121/1.1394196.
40. Elliott, Ni, and Verschuur, *Hear. Res.*, **341**, 155–167 (2016) http://dx.doi.org/10.1016/j.heares.2016.08.006.
41. Rosowski, Songer, Nakajima, Brinsko, and Merchant, *Otol. Neurotol.*, **25**, 323–332 (2004) http://journals.lww.com/00129492-200405000-00021.
42. Feng, and Gan, *Biomech. Model. Mechanobiol.*, **3**, 33–47 (2004) http://journals.kluweronline.com/article.asp?PIPS=486601.
43. Liu, and Neely, *J. Acoust. Soc. Am.*, **127**, 2420–2432 (2010) http://asa.scitation.org/doi/10.1121/1.3337233.
44. Møller, *J. Acoust. Soc. Am.*, **33**, 168–176 (1961) http://asa.scitation.org/doi/10.1121/1.1908610.
45. Marquardt, and Hensel, *J. Acoust. Soc. Am.*, **134**, 3730–3738 (2013) http://asa.scitation.org/doi/10.1121/1.4824154.
46. Zwislocki, *J. Acoust. Soc. Am.*, **34**, 1514–1523 (1962) http://asa.scitation.org/doi/10.1121/1.1918382.
47. Kringlebotn, *Scand. Audiol.*, **17**, 75–85 (1988) http://www.tandfonline.com/doi/full/10.3109/01050398809070695.
48. Lutman, and Martin, *J. Sound Vib.*, **64**, 133–157 (1979) https://linkinghub.elsevier.com/retrieve/pii/0022460X79905789.
49. Rosowski, and Merchant, *Am. J. Otol.*, **16**, 486–497 (1995) http://www.ncbi.nlm.nih.gov/pubmed/8588650.
50. Goode, Killion, Nakamura, and Nishihara, *Am. J. Otol.*, **15**, 145–154 (1994) https://journals.lww.com/otology-neurotology/toc/1994/03000.
51. Giguère, and Woodland, *J. Acoust. Soc. Am.*, **95**, 331–342 (1994) http://asa.scitation.org/doi/10.1121/1.408366.
52. Ravicz, and Rosowski, *J. Acoust. Soc. Am.*, **134**, 2852–2865 (2013) http://asa.scitation.org/doi/10.1121/1.4818745.
53. Nakajima, Dong, Olson, Merchant, Ravicz, and Rosowski, *J. Assoc. Res. Otolaryngol.*, **10**, 23–36 (2009) http://link.springer.com/10.1007/s10162-008-0150-y.
54. Nie, Wang, Zheng, Xu, Li, Wang, et al., *J. Acoust. Soc. Am.*, **151**, 1434–1448 (2022) https://asa.scitation.org/doi/10.1121/10.0009618.
55. Rusinek, and Weremczuk, *Meccanica*, **55**, 2609–2621 (2020) http://link.springer.com/10.1007/s11012-020-01226-x.
56. Dallos, *J. Acoust. Soc. Am.*, **48**, 489–499 (1970) http://asa.scitation.org/doi/10.1121/1.1912163.
57. Lemons, and Meaud, *J. Acoust. Soc. Am.*, **140**, 2735–2753 (2016) http://dx.doi.org/10.1121/1.4964707.
58. Puria, and Allen, *J. Acoust. Soc. Am.*, **104**, 3463–3481 (1998) http://asa.scitation.org/doi/10.1121/1.423930.
59. Voss, Effects of tympanic membrane perforations on the middle ear sound transmission: measurements, mechanisms and models, Ph.D. thesis, Massachusetts Institute of Technology, 1998.
60. Guinan, and Peake, *J. Acoust. Soc. Am.*, **41**, 1237–1261 (1967) http://asa.scitation.org/doi/10.1121/1.1910465.
61. O'Connor, and Puria, *J. Acoust. Soc. Am.*, **120**, 1517–1528 (2006) http://asa.scitation.org/doi/10.1121/1.2221414.
62. Shaw, and Stinson, *JASA*, (1981).
63. Shaw, *J. Acoust. Soc. Am.*, **56**, 1848–1861 (1974) http://asa.scitation.org/doi/10.1121/1.1903522.
64. Kringlebotn, and Gundersen, *J. Acoust. Soc. Am.*, **77**, 159–164 (1985) http://asa.scitation.org/doi/10.1121/1.392280.
65. Gan, Wood, and Dormer, *Otol. Neurotol.*, **25**, 423–435 (2004) http://journals.lww.com/00129492-200407000-00005.
66. Voss, Rosowski, Shera, and Peake, *J. Acoust. Soc. Am.*, **107**, 1548–1565 (2000) http://asa.scitation.org/doi/10.1121/1.428440.
67. Puria, Rosowski, and Peake, *Symp. Biophys. Hair Cell Sens. Syst.*, 345–351 (1993).
68. Gyo, Aritomo, and Goode, *Acta Otolaryngol.*, **103**, 87–95 (1987) http://www.tandfonline.com/doi/full/10.3109/00016488709134702.
69. Puria, *J. Acoust. Soc. Am.*, **113**, 2773–2789 (2003) http://asa.scitation.org/doi/10.1121/1.1564018.

APPENDICES

Appendix A – Search Strategy

TABLE A.1. Pubmed search strategy

Number	Search term
#4	#1 AND #2 AND #3
#3	"hearing"[MeSH Terms] OR "hearing"[All Fields] OR "hearings"[All Fields] OR ("ear"[MeSH Terms] OR "ear"[All Fields]
#2	"sound conduct*"[All Fields] OR "sound transmi*"[All Fields] OR "sound pressure"[All Fields] OR "air conduct*"[All Fields] OR "bone conduct*"[All Fields] OR "ear acoustics"[All Fields]
#1	"lumped parameter*"[All Fields] OR "lumped element*"[All Fields] OR "electric model*"[All Fields] OR "electrical model*"[All Fields] OR "electric analog"[All Fields] OR "electrical analog"[All Fields] OR "impedance model*"[All Fields]

TABLE A.2. Embase search strategy

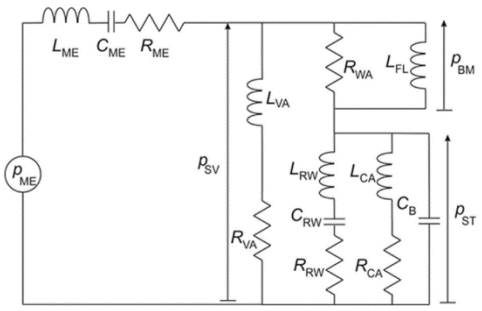
Number	Search term
#4	#1 AND #2 AND #3
#3	hearing OR ear
#2	“sound conduct*“ OR “sound transmi*” OR “sound pressure“ OR “air conduct*“ OR “bone conduct*“ OR “ear acoustics“
#1	“lumped parameter*” OR "lumped element*" OR "electric model*" OR "electrical model*" OR "electric analog" OR "electrical analog" OR “impedance model*”

TABLE A.3. Scopus search strategy

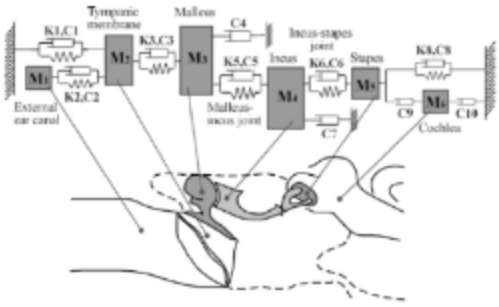
Number	Search term
#4	#1 AND #2 AND #3
#3	TITLE-ABS-KEY (hearing OR ear)
#2	TITLE-ABS-KEY ((sound PRE/0 conduct*) OR (sound W/0 transmi*) OR (sound W/0 pressure) OR (air PRE/0 conduct*) OR (bone PRE/0 conduct*) OR (ear PRE/0 acoustics))
#1	TITLE-ABS-KEY ((lumped PRE/1 parameter*) OR (lumped PRE/1 element*) OR (electric* PRE/0 model*) OR (electric* PRE/0 analog) OR (impedance PRE/0 model*))

Appendix B – Schematic View of LEMs

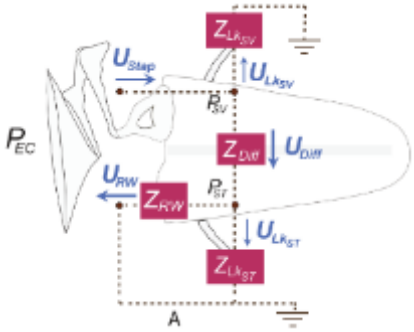
Elliott et al.[4] [40]



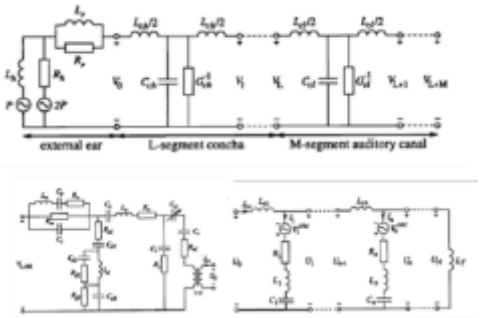
Feng et al.[5] [42]



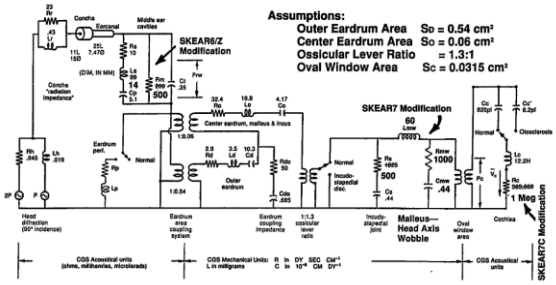
Frear et al.[6] [14]



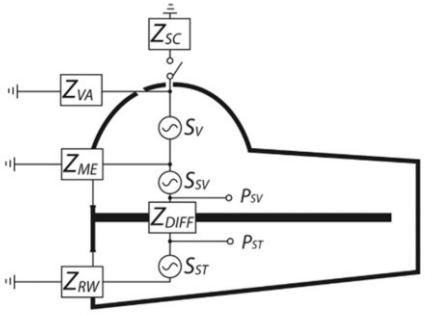
Giguère et al.[7] [51]



Goode et al.[8] [50]

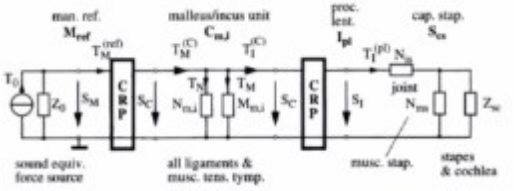


Guan et al.[9] [10]

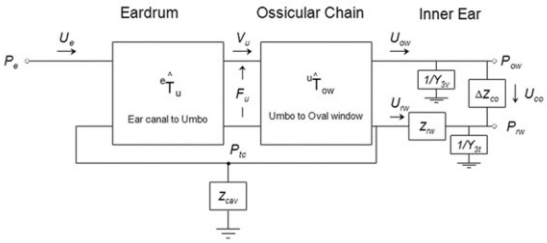


[4] Reprinted from Elliott, Ni, and Verschuur, Hear. Res., 341, 155–167 (2016). © 2016 Elliott, Ni, and Verschuur
[5] Reprinted by permission from Springer Nature: Springer Nature Biomech. Model. Mechanobiol. Lumped parametric model of the human ear for sound transmission, Feng and Gan, © 2004
[6] Reprinted from Hear. Res., 367, Frear, Guan, Stieger, Rosowski, and Nakajima, Impedances of the inner and middle ear estimated from intracochlear sound pressures in normal human temporal bones, 17-31, © (2018), with permission from Elsevier.
[7] Reprinted with permission from Goode, Killion, Nakamura, and Nishihara, New Knowledge about the function of the human middle ear: development of an improved analog model, Am. J. Otol., 15, 2, 145-154 © 1994, Wolters Kluwer Health, Inc.
[8] Reprinted with permission from Giguère, and Woodland, J. Acoust. Soc. Am., 95, 331–342 (1994). © 1994, Acoustic Society of America.
[9] Reprinted from Guan, Cheng, Galaiya, Rosowski, Lee, and Nakajima, Sci. Rep., 10, 16564 (2020). © 2020 Guan, Cheng, Galaiya, Rosowski, Lee, and Nakajima

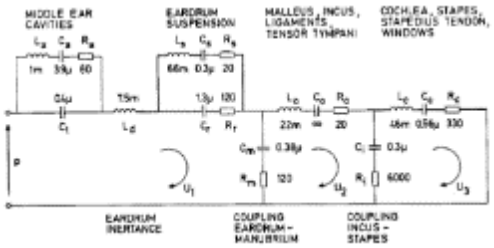
Hudde et al.[10] [36]



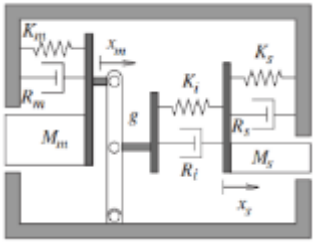
Keefe[11] [33]



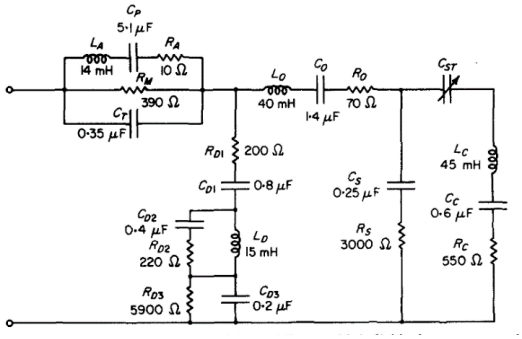
Kringlebotn[12] [47]



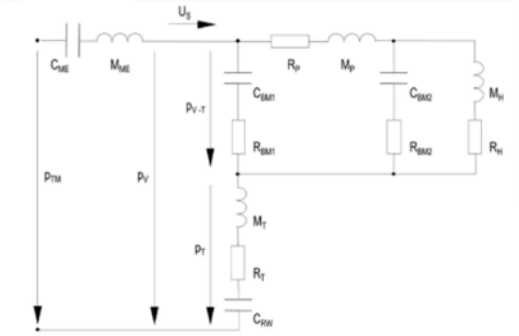
Liu et al.[13] [43]



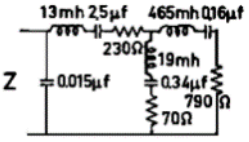
Lutman et al.[14] [48]



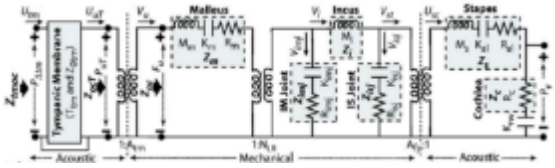
Marquardt et al.[15] [45]



Møller et al.[16] [44]

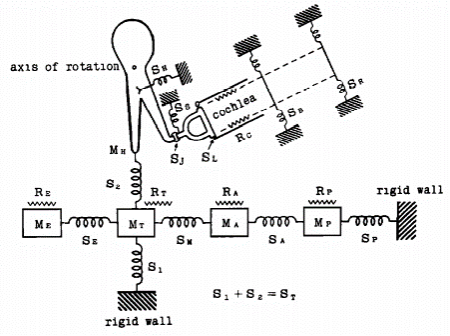


O'Connor et al.[17] [34]

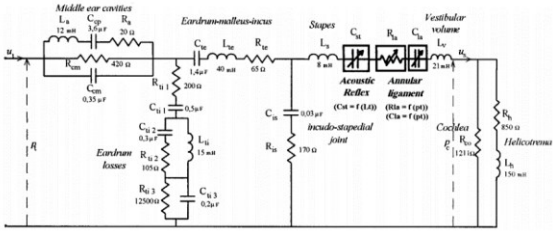


[10] Send request to reprint to publisher
[11] Reprinted with permission from Keefe, J. Acoust. Soc. Am., 137, 2698–2725 (2015). © 2015, Acoustic Society of America.
[12] Network Model for the Human Middle Ear, Kringlebotn, Scandinavian Audiology, 1988, reprinted by permission of the publisher (Taylor & Francis Ltd, http://www.tandfonline.com).
[13] Reprinted with permission from Liu, and Neely, J. Acoust. Soc. Am., 127, 2420–2432 (2010) © 2010, Acoustic Society of America.
[14] Reprinted from J. Sound Vib., 64, Lutman and Martin, Development of an electroacoustic analogue model of the middle ear and acoustic reflex, 133-157, © (1979), with permission from Elsevier.
[15] Reprinted with permission from Marquardt, and Hensel, J. Acoust. Soc. Am., 134, 3730–3738 (2013). © 2013, Acoustic Society of America.
[16] Reprinted with permission from Møller, J. Acoust. Soc. Am., 33, 168–176 (1961). © 1961, Acoustic Society of America.
[17] Reprinted with permission from O'Connor, and Puria, J. Acoust. Soc. Am., 123, 197–211 (2008). © 2008, Acoustic Society of America.

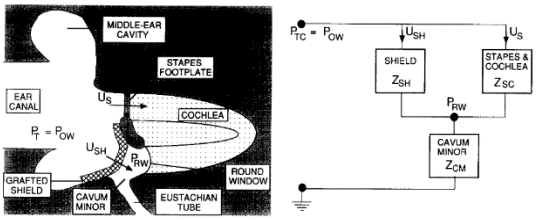
Onchi[18] [38]



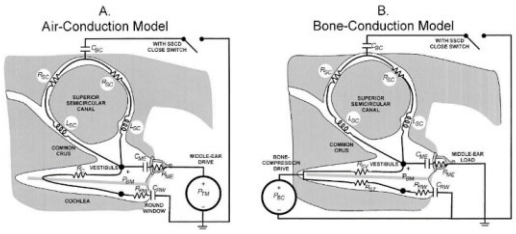
Pascal et al.[19] [13]



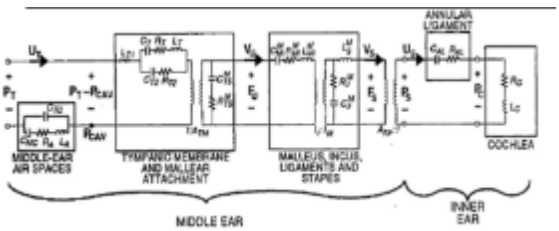
Peake et al.[20] [37]



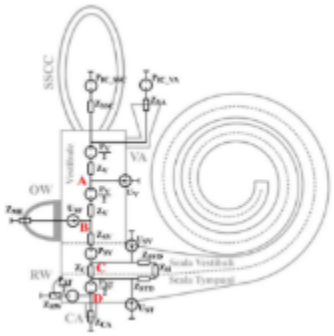
Rosowski et al.[21] [41]



Rosowski et al.[22] [49]

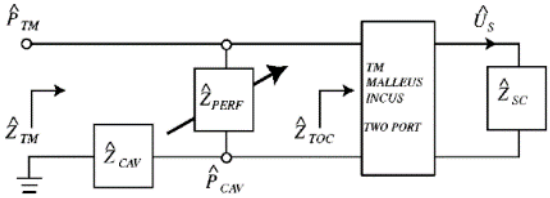


Stenfelt[23] [35]

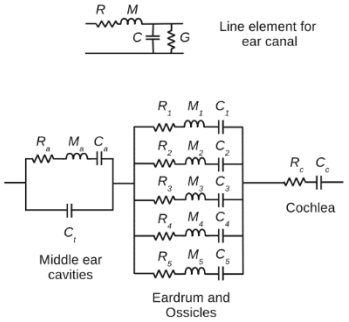


[18] Reprinted with permission from Onchi, J. Acoust. Soc. Am., 33, 794–805 (1961). © 1961, Acoustic Society of America.
[19] Reprinted with permission from Pascal, Bourgeade, Lagier, and Legros, J. Acoust. Soc. Am., 104, 1509–1516 (1998).. © 1998, Acoustic Society of America.
[20] Reprinted from Hear. Res., 57, Peake, Rosowski, and Lynch, Middle-ear transmission: Acoustic versus ossicular coupling in cat and human, 245-268, © (1992), with permission from Elsevier
[21] Reprinted with permission from Rosowski, Songer, Nakajima, Brinsko, and Merchant, Clinical, Experimental, and Theoretical Investigations of the Effect of Superior Semicircular Canal Dehiscence on Hearing Mechanisms, Otol. Neurotol., 25, 3, 323-332 © 2004, Wolters Kluwer Health, Inc.
[22] Reprinted with permission from Rosowski and Merchant, Mechanical and acoustic analysis of middle ear reconstruction., Am. J. Otol., 16, 4, 486-497 © 1995, Wolters Kluwer Health, Inc.
[23] Reprinted from Stenfelt, Front. Neurol., 11, 966 (2020). © 2020 Stenfelt S.

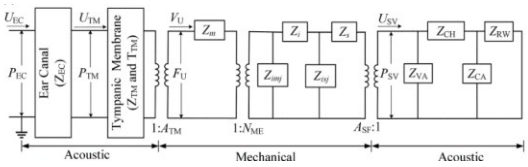
Voss et al.[24] [39]



Withnell et al.[25] [23]



Xue et al.[26] [4]



Zwislocki[27] [46]

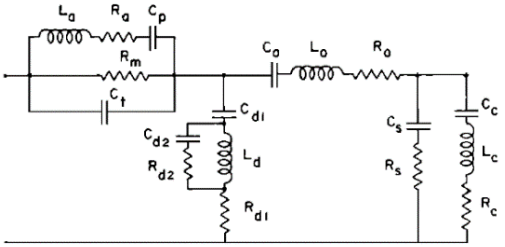


FIGURE B.1. Schematic overview of lumped element models

[24] Reprinted with permission from Voss, Rosowski, Merchant, and Peake, J. Acoust. Soc. Am., 110, 1445–1452 (2001). © 2001, Acoustic Society of America.
[25] Reprinted by permission from Springer Nature : Springer Nature, J. Assoc. Res. Otolaryngol., An Analysis of the Acoustic Input Impedance of the Ear, Withnell and Gowdy © 2013
[26] Reprinted with permission from Xue, Liu, Wang, Yang, Zhao, and Huang, J. Acoust. Soc. Am., 147, 1478–1490 (2020). © 2020, Acoustic Society of America.
[27] Reprinted with permission from Zwislocki, J. Acoust. Soc. Am., 34, 1514–1523 (1962). © 1962, Acoustic Society of America.

Appendix C – Summary of Reported - Comparison Between Predicted and Experimental Data

Table C.1 Summary of reported comparison between predicted and experimental data. With diff: differential, EC: ear canal, METF: middle ear transfer function, P: pressure, RW: round window, ST: scala tympani, SV: scala vestibuli, TM: tympanic membrane, U: volume velocity, V: linear velocity, x: linear displacement, Z: impedance.

Author (Year)	Compared data	Reference experimental data	Sample size	Type samples
Voss et al. [39]	$Z_{TM} = \frac{P_{TM}}{U_{TM}}$	*Voss* [59]	9	Human temporal bones
Hudde et al. [36]	V_{stapes}	*Hudde et al.* [36]	3	Human temporal bones
Peake et al. [37]	$\frac{U_{stapes}}{P_{EC}}$	*Guinan et al.* [60]	25	Anesthetized cats
Frear et al. [14]	$\frac{P_{diff}}{P_{EC}}$ $\frac{P_{SV}}{P_{EC}}$ $\frac{P_{ST}}{P_{EC}}$ $\frac{P_{SV}}{V_{stapes}}$ $\frac{P_{ST}}{V_{stapes}}$ $Z_{diff} = \frac{P_{SV} - P_{ST}}{U_{stapes}}$ $Z_{RW} = \frac{P_{ST}}{U_{stapes}}$	*Frear et al. [14]*	18	Human temporal bones
Guan et al. [10]	$Z_{diff} = \frac{P_{SV} - P_{ST}}{U_{stapes}}$ $Z_{cochlea} = \frac{P_{SV}}{U_{stapes}}$ $Z_{RW} = \frac{P_{ST}}{U_{stapes}}$	*Guan et al. [10]*	6	Human temporal bones
Møller et al. [44]	$Z_{TM} = \frac{P_{TM}}{U_{TM}}$	*Møller et al.* [44]	2	Normal hearing subjects
Elliott et al. [40]	$Z_{RW} = \frac{P_{ST}}{U_{stapes}}$	*Nakajima et al.*[53]	6	Human temporal bones

Marquardt et al. [45]	$\frac{P_{diff}}{P_{EC}}$ $\frac{P_{SV}}{P_{EC}}$ $\frac{P_{ST}}{P_{EC}}$ $Z_{diff} = \frac{P_{SV} - P_{ST}}{U_{stapes}}$ $Z_{RW} = \frac{P_{ST}}{U_{stapes}}$	*Nakajima et al.*[53]	6	Human temporal bones
O'Connor et al. [34]	$METF = \frac{V_{stapes}}{P_{EC}}$ $\frac{V_{stapes}}{V_{umbo}}$ $\frac{V_{incus}}{V_{umbo}}$ $\frac{V_{stapes}}{V_{incus}}$	*O'Connor et al.* [61] + *O'Connor et al. [34]*	4 + 12	Human temporal bones
Zwislocki [46]	$Z_{TM} = \frac{P_{TM}}{U_{TM}}$	*Zwislocki* [46]	8	Normal hearing and hearing impaired subjects
Kringlebotn [47]	$Z_{TM} = \frac{P_{TM}}{U_{TM}}$	*Shaw et al.* [62]	Record not obtained	
	$\frac{P_{EC}}{P_{free\ field}}$	*Shaw* [63]	19	Normal hearing subjects
	$\frac{x_{stapes}}{P_{EC}}$	*Kringlebotn et al.* [64]	68	Human temporal bones
Lutman et al. [48]	$Z_{TM} = \frac{P_{TM}}{U_{TM}}$	*Zwislocki* [46]	8	Normal hearing and hearing impaired subjects
Stenfelt [35]	$\frac{P_{SV}}{P_{EC}}$ $\frac{P_{ST}}{P_{EC}}$	*Nakajima et al.*[53] + *Niesten et al. [7]*	6 + 5	Human temporal bones
	$\frac{P_{SV}}{V_{promontory}}$ $\frac{P_{ST}}{V_{promontory}}$	*Stieger et al. [6]* + *Mattingly et al. [31]*	8 + 4	Human temporal bones + Human whole heads
Onchi [38]	$\frac{P_{SV}}{P_{EC}}$	*Onchi* [38]	3	Human temporal bones
Feng et al. [42]	$\frac{x_{TM}}{x_{stapes}}$	*Gan et al.* [65]	10	Human temporal bones

Rosowski et al. [49]	$Z_{TM} = \frac{P_{TM}}{U_{TM}}$	*Rabinowitz [22]* + *Hudde [21]*	4 + 6	Normal hearing subjects
	$\frac{U_{stapes}}{P_{EC}}$	*Kringlebotn et al.* [64]	68	Human temporal bones
Xue et al. [4]	$Z_{TM} = \frac{P_{TM}}{U_{TM}}$	*Voss et al.* [66] + *Hudde [21]* + *Farmer-Fedor et al.* [24]	13 + 6 + 6	Human temporal bones + Normal hearing subjects (2)
	$METF = \frac{V_{stap}}{P_{EC}}$ $Z_{cochlea} = \frac{P_{SV}}{U_{stapes}}$	*Aibara et al. [17]* + *O'Connor et al. [34]*	11 + 16	Human temporal bones
	$\frac{P_{SV}}{P_{EC}}$	*Aibara et al. [17]* + *Frear et al. [14]*	11 + 18	Human temporal bones
Pascal et al. [13]	$Z_{TM} = \frac{P_{TM}}{U_{TM}}$ $\frac{P_{SV}}{P_{EC}}$	*Puria et al.* [67]	Record not obtained	
Withnell et al. [23]	$Z_{TM} = \frac{P_{TM}}{U_{TM}}$	*Withnell et al. [23]*	6	Normal hearing subjects
Goode et al. [50]	$\frac{x_{stapes}}{x_{umbo}}$	*Gyo et al.* [68]	14	Human temporal bones
Keefe [33]	$Z_{TM} = \frac{P_{TM}}{U_{TM}}$	*Margolis et al.* [26]	20	Normal hearing subjects
	$Z_{cochlea} = \frac{P_{SV}}{U_{stapes}}$ $\frac{P_{SV}}{P_{EC}}$	*Nakajima et al.*[53] + *Puria* [69]	6 + 5	Human temporal bones
	$\frac{P_{ST}}{P_{EC}}$ $\frac{P_{diff}}{P_{EC}}$	*Nakajima et al.*[53]	6	Human temporal bones

Appendix D – Comparison of Model Predictions with Experimental data

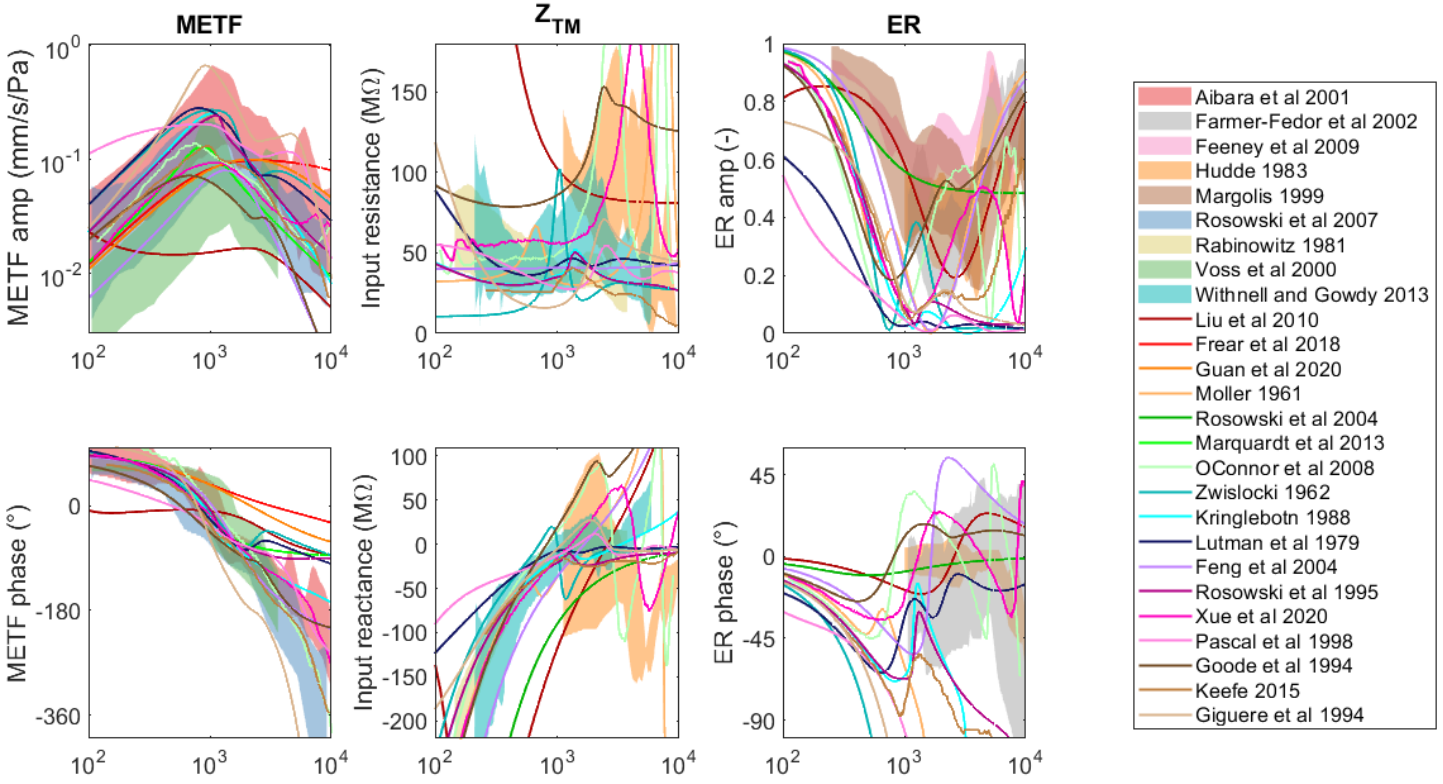


FIGURE D.1. Model predictions of 1) the METFs (stapes velocity normalized to ear canal pressure) compared with the experimental data of Aibara et al.[17], Rosowski et al. [18] and Voss et al.[19]. 2) the Z_{TM} (ratio of pressure to volume velocity at the tympanic membrane) compared with the experimental data of Hudde [21], Rabinowitz [22], and Withnell et al. [23]. 3) the ER (ratio of reflected to incident energy) compared with the experimental data of Hudde [21], Feeney et al. [25], Farmer-Fedor et al. [24], and Margolis et al. [26].

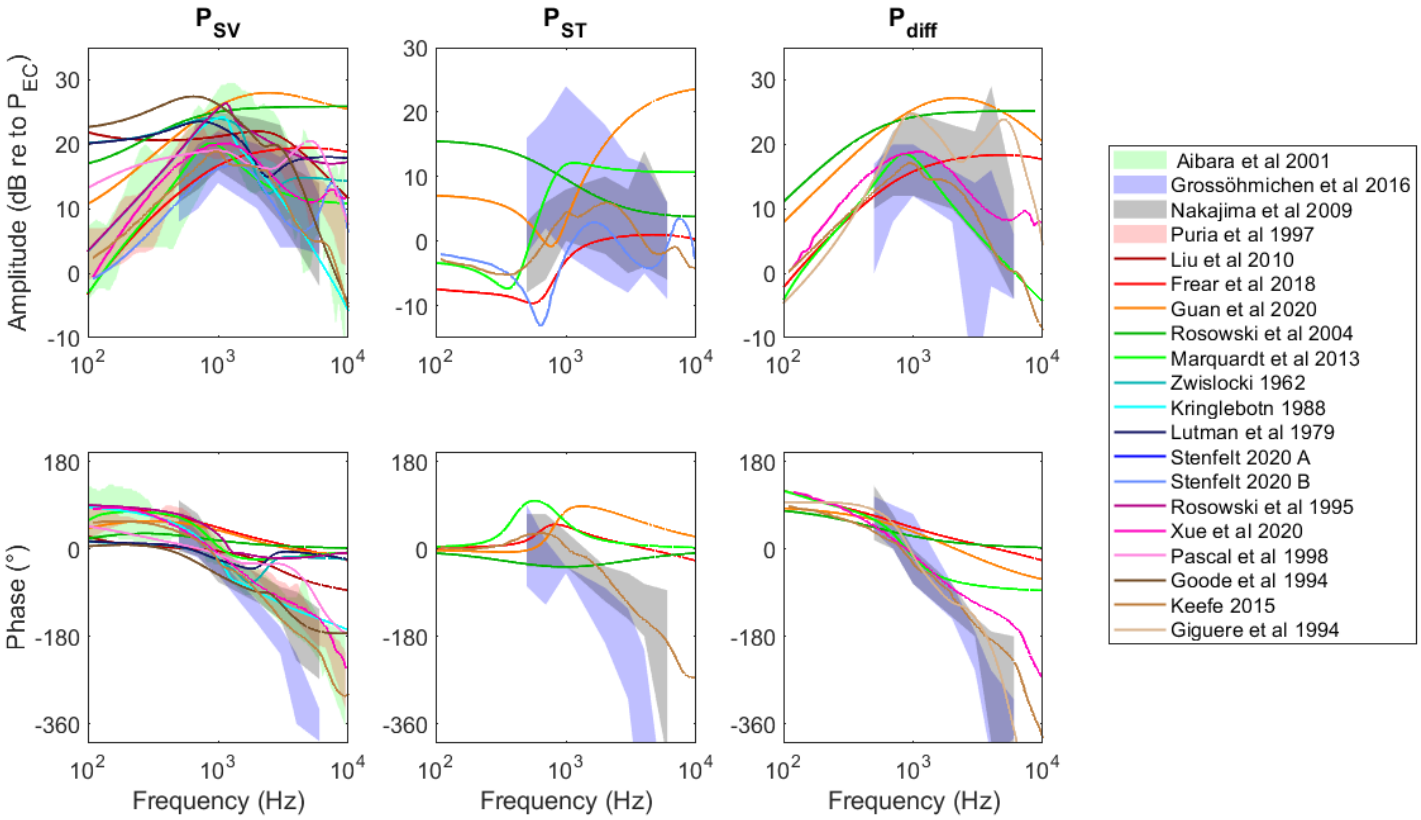


FIGURE D.2. Model predictions of ICPs (P_{SV}, P_{ST}, and P_{diff}) for AC normalized to ear canal pressure, compared with experimental data of Aibara et al. [17], Grossöhmichen et al. [27], Nakajima et al. [53], and Puria et al. [29].

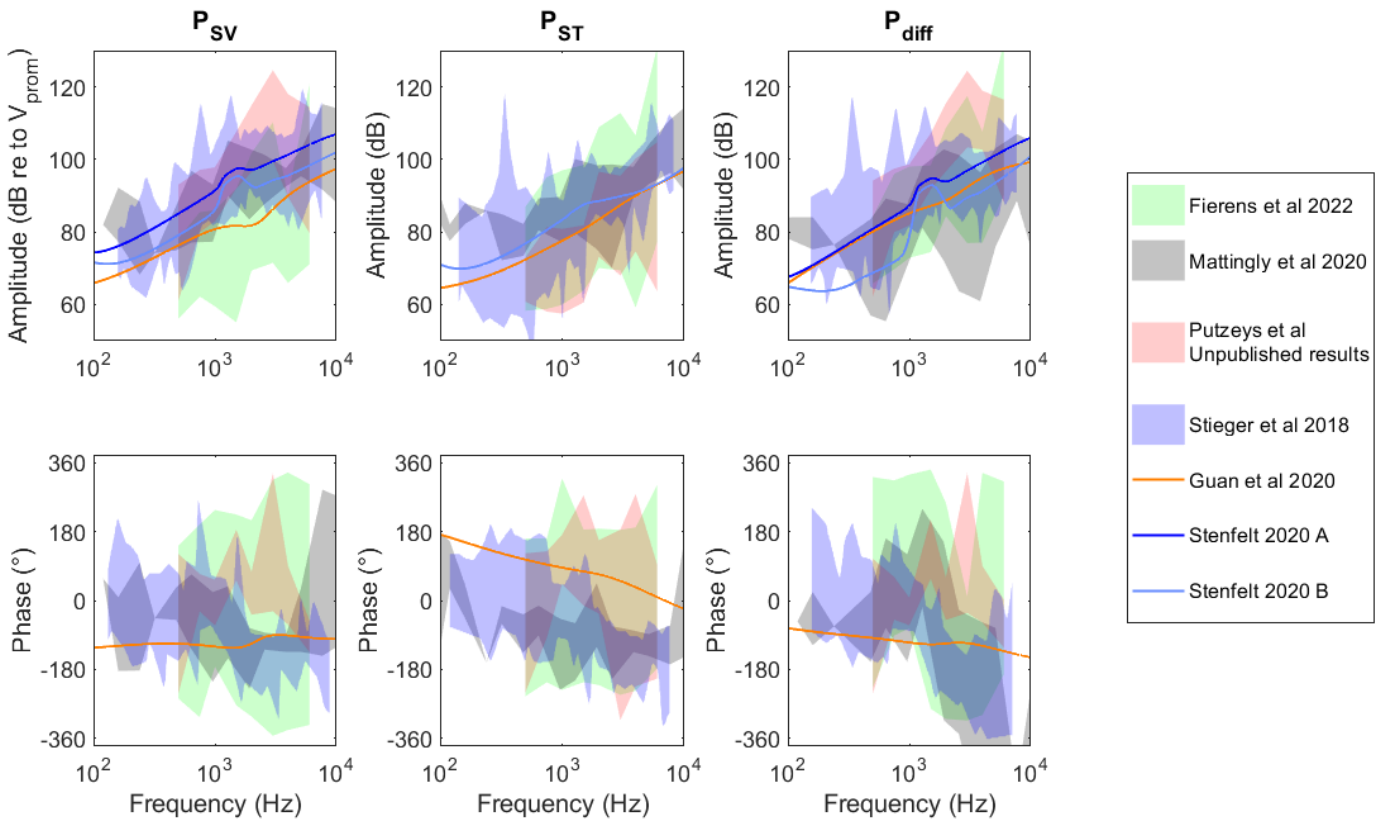


FIGURE 3. Model predictions of ICPs (P_{SV}, P_{ST}, and P_{diff}) for BC normalized to promontory velocity, compared with experimental data of Stieger et al. [6], Mattingly et al. [31], Fierens and Borgers et al. [30] end Putzeys et al. [32].

Appendix E – Frequency Ranges for which the Model Predictions are Within the Experimental Data Ranges

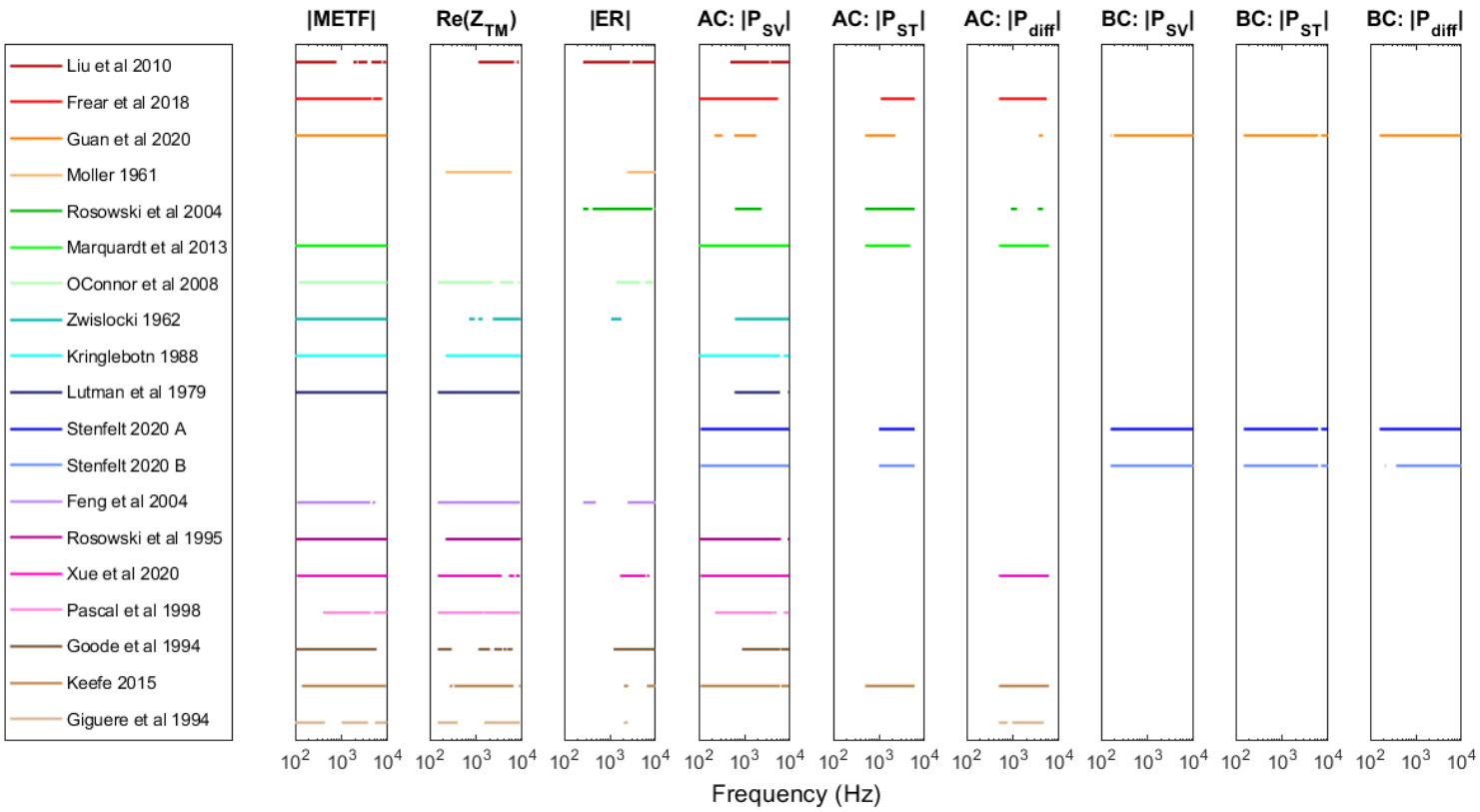


FIGURE E.1. Frequency ranges where the model predictions are within the experimental ranges: amplitudes, except for Z_{TM}, where resistance is shown. METF: middle ear transfer function, Z_{TM}: tympanic membrane impedance, ER: energy reflectance, AC: air conduction, BC: bone conduction, P_{SV}: scala vestibuli pressure, P_{ST}: scala tympani pressure, P_{diff}: differential pressure.

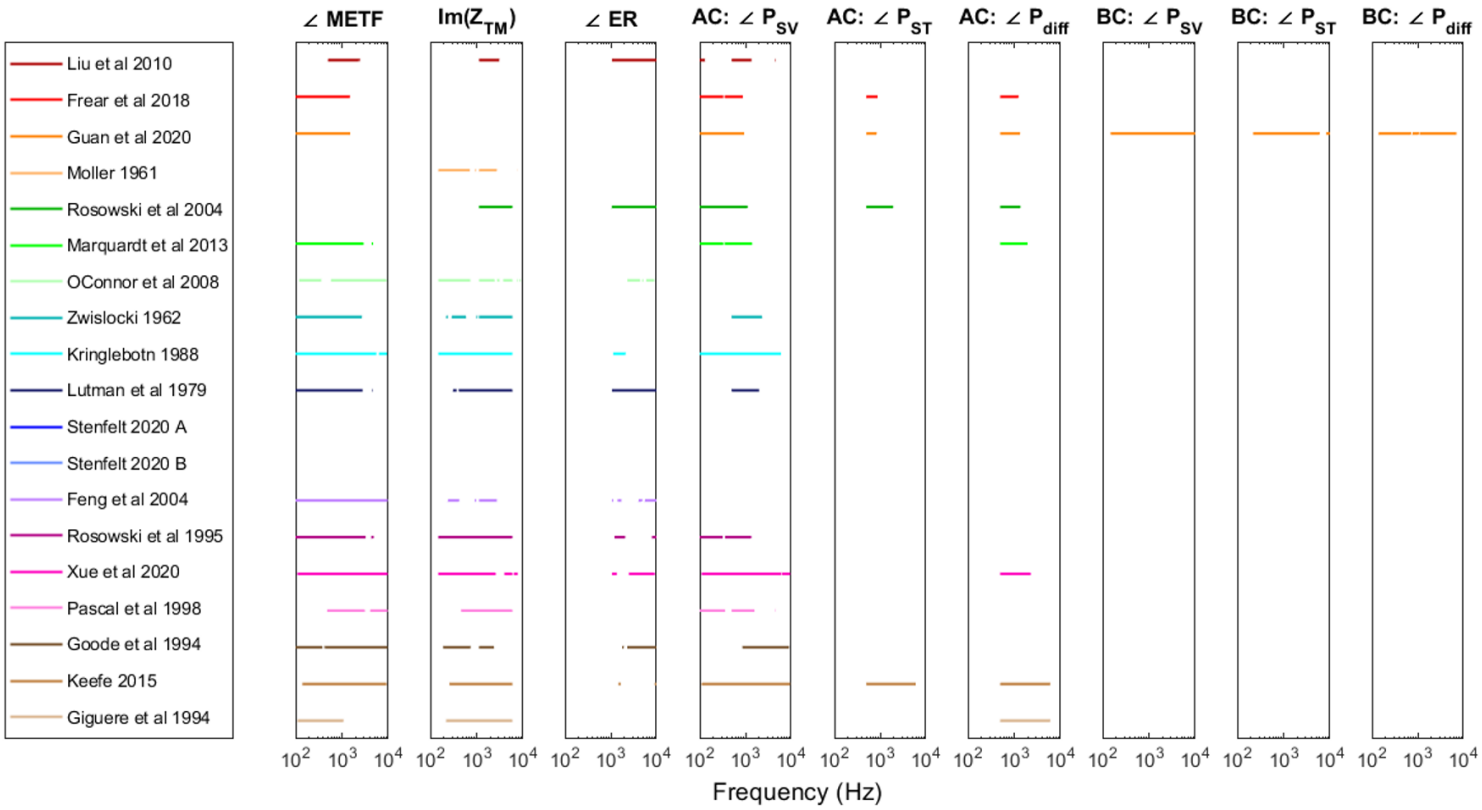


FIGURE E.2. Frequency ranges where the model predictions are within the experimental ranges: amplitudes, except for Z_{TM}, where reactance is shown. METF: middle ear transfer function, Z_{TM}: tympanic membrane impedance, ER: energy reflectance, AC: air conduction, BC: bone conduction, P_{SV}: scala vestibuli pressure, P_{ST}: scala tympani pressure, P_{diff}: differential pressure.

Table E.1 Frequency ranges for which the model predictions are within the experimental data ranges: amplitudes, except for Z_{TM}, where the resistance is shown. METF: middle ear transfer function, Z_{TM}: tympanic membrane impedance, ER: energy reflectance, AC: air conduction, BC: bone conduction, P_{SV}: scala vestibuli pressure, P_{ST}: scala tympani pressure, P_{diff}: differential pressure.

Author	|METF|	Re(Z_TM)	|ER|	AC: $\|P_{SV}\|$	AC: $\|P_{ST}\|$	AC: $\|P_{diff}\|$	BC: $\|P_{SV}\|$	BC: $\|P_{ST}\|$	BC: $\|P_{diff}\|$
Liu et al. [43]	[100 780] [1820 2170] [2280 3620] [4450 7390] [7900 9160]	[1130 6620] [7610 8350]	[260 2760] [3060 10000]	[490 3520] [3750 9460]					
Frear et al. [14]	[100 4510] [4760 7510]			[100 4340] [4360 5430]	[1090 6000]	[500 5330]			
Guan et al. [10]	[100 9720]			[220 330] [600 1830]	[500 2280]	[3620 4430]	[160 170] [190 10000]	[150 6000] [7020 10000]	[160 10000]
Møller [44]		[220 5890]	[2380 10000]						
Rosowski et al. [41]			[260 340] [420 8570] [9380 10000]	[620 2400]	[500 6000]	[910 1230] [3450 4480]			
Marquardt et al. [45]	[100 9720]			[100 9460]	[500 4810]	[500 6000]			
O'Connor et al. [34]	[120 9720]	[150 2280] [3290 6430] [8010 8980]	[1360 4590] [5910 8560]						
Zwislocki [46]	[100 9720]	[710 940] [1130 1380] [2290 9150]	[1070 1790]	[620 9460]					
Kringlebotn [47]	[100 9320]	[150 150] [220 9150]		[100 6000] [7140 9460]					
Lutman et al. [48]	[100 9720]	[150 8820]		[610 6010] [8990 9460]					
Stenfelt [35]				[110 9460]	[1000 6000]		[160 10000]	[150 6000] [7010 10000]	[160 9460]
Feng et al. [42]				[110 9460]	[1000 6000]		[160 10000]	[150 6000] [7010 10000]	[210 220] [370 10000]
Rosowski et al. [49]	[110 4170] [4590 4590] [4640 5390]	[150 8820]	[260 490] [2420 10000]						
Xue et al. [4]	[100 9720]	[150 150] [220 9150]		[100 6270] [8800 8860] [8890 9460]					
Pascal et al. [13]	[110 9720]	[150 3590] [5100 6680] [7270 8730]	[1640 1690] [1720 6110] [6460 6640] [6670 6700]	[110 9460]		[500 6000]			

			[6780 7320]						
Goode et al. [50]	[400 4290] [4900 9720]	[150 1460] [1510 8920]		[230 4250] [4370 5010] [7160 9460]					
Keefe [33]	[100 5750]	[150 300] [1130 2000] [2450 2450] [2480 3680] [3920 4530] [4770 6190]	[1200 10000]	[890 6140] [6410 9460]					
Giguère et al. [51]	[140 9160]	[270 310] [340 6500] [8500 9060] [9230 9500]	[2030 2500] [6450 10000]	[110 6000] [6410 9460]	[500 6000]	[500 6000]			

Table E.2 Frequency ranges for which the model predictions are within the experimental data ranges: phases, except for Z_{TM}, where the reactance is shown. METF: middle ear transfer function, Z_{TM}: tympanic membrane impedance, ER: energy reflectance, AC: air conduction, BC: bone conduction, P_{SV}: scala vestibuli pressure, P_{ST}: scala tympani pressure, P_{diff}: differential pressure.

Author	**∠METF**	**Im(Z_{TM})**	**∠ER**	**AC: ∠P_{SV}**	**AC: ∠P_{ST}**	**AC: ∠P_{diff}**	**BC: ∠P_{SV}**	**BC: ∠P_{ST}**	**BC: ∠P_{diff}**
Liu et al. [43]	[500 2320] [2350 2410] [2440 2460]	[1130 3060]	[1070 10000]	[100 130] [500 1350] [4420 4570]					
Frear et al. [14]	[100 1480]			[100 330] [350 880] [900 900]	[500 880]	[500 1240]			
Guan et al. [10]	[100 1510]			[100 940]	[500 840]	[500 1340]	[150 10000]	[100 100] [220 6000] [8260 10000]	[140 720] [770 1060] [1100 7040]
Møller [44]		[150 710] [930 980] [1130 2730] [7400 7460]							
Rosowski et al. [41]		[1130 5910]	[1060 10000]	[100 1120]	[500 1930]	[500 1360]			
Marquardt et al. [45]	[100 2920] [4400 4700]			[100 330] [350 1400]		[500 1940]			
O'Connor et al. [34]	[120 360] [580 9160]	[150 730] [1130 2440] [2760 3090] [3720 5890] [7300 7500]	[2310 4400] [4870 5150] [6010 8980]						

		[8430 8980]				
Zwislocki [46]		[220 240] [290 590] [990 1050]		[120 120]		
	[100 2710]	[1130 5910]		[500 2320]		
Kringlebotn [47]	[100 5530] [6300 9380]	[150 5910]	[1140 2120]	[100 6000]		
Lutman et al. [48]	[100 2790] [4490 4630]	[220 220] [310 370] [390 390] [410 5910]	[1070 10000]	[100 100] [120 120] [500 1990]		
Stenfelt [35]						
Feng et al. [42]	[100 9720]	[240 420] [910 980] [1130 2730]	[1070 1140] [1410 1710] [4060 4730] [4770 4900] [4980 5040] [5560 10000]			
Rosowski et al. [49]	[100 3230] [4290 4850]	[150 5910]	[1210 2070] [7980 10000]	[100 320] [360 1340]		
Xue et al. [4]	[110 9720]	[150 2580] [3990 5910] [6310 6310] [6350 7620]	[1070 1360] [2500 9150] [9200 9210] [9830 9860]	[110 6000] [6030 6030] [6260 9540]		[500 2270] [2290 2290]
Pascal et al. [13]	[480 3020] [3040 3150] [3210 3210] [4070 9720]	[460 5910]		[100 360] [500 1580] [4450 4570]		
Goode et al. [50]	[100 380] [400 410] [430 9720]	[190 740] [1130 2370]	[1780 1920] [2300 10000]	[860 9050]		
Keefe [33]	[140 9160]	[260 5910]	[1480 1660] [9150 10000]	[110 9540]	[500 6000]	[500 6000]
Giguère et al. [51]	[110 1080]	[220 5910]				[500 6000]

Can You Hear Me Now? Binaural Brainstem and Spatial Hearing Deficits in a Guinea Pig Model of Noise-Induced Cochlear Synaptopathy

Monica A. Benson[1,2,3], John Peacock[2], Nathaniel T. Greene[1,3], Daniel J. Tollin[1,2,3]

[1)]*Neuroscience Graduate Program, University of Colorado School of Medicine, Aurora, CO 80045, USA*
[2)]*Department of Physiology and Biophysics, University of Colorado School of Medicine, Aurora, CO 80045, USA*
[3)]*Department of Otolaryngology, University of Colorado School of Medicine, Aurora, CO 80045, USA*

Abstract. Noise-induced cochlear synaptopathy has been studied for over 20 years with no known diagnosis for this disorder within the human population. This type of "hidden hearing loss" induces a loss of synapses in the inner ear but no change in audiometric thresholds. Animal studies to this point have focused primarily on peripheral hearing measures to diagnose ribbon synapse loss. Here, we measure binaural hearing deficits in a model of noise-induced cochlear synaptopathy for the first time. Our results show that while common audiological assessments show no change in threshold post moderate noise exposure, measurable binaural hearing deficits persist long-term.

INTRODUCTION

Sensorineural hearing loss is typically diagnosed via permanently raised auditory thresholds due to cochlea dysfunction [19,26]. This can be easily assessed and diagnosed using traditional audiological tools such as subjective hearing tests (the audiogram) as well as objective tests of peripheral auditory function including otoacoustic emissions (OAEs) and auditory brainstem responses (ABRs) to assess hair cell and auditory nerve function, respectively. Perplexingly, up to 10% of people report difficulty hearing in noisy environments, such as classrooms and workplaces, despite normal audiometric thresholds [25,14,36,27]. The reasons for this are not yet known because the hearing difficulty is 'hidden' from routine hearing tests. Recent animal studies have shown that a moderate intensity noise exposure can cause temporary hearing threshold shifts, but permanent loss of synapses between inner hair cells (IHC) and the auditory nerve (ANF) in the cochlea [8,16]. This has been termed cochlear synaptopathy.

The cochlea contains IHCs that relay acoustical information about our environment to ANFs as well as outer hair cells (OHCs), which amplify the signal from IHCs to the cochlear nerve [15]. OHC function can be evaluated by measuring distortion product otoacoustic emissions (DPOAEs) [1]. DPOAE loss and its recovery post noise exposure has been shown in noise-induced synaptopathy studies [5,16] suggesting normal OHC functioning in synaptopathy. Ascending auditory pathway function can be assessed by ABRs. Each ABR peak roughly corresponds to synchronized neural activity at a specific point along the auditory pathway. The ABR waveforms thus provide insight into the functioning of the auditory brainstem; for example, wave I tells us about they synchronous firing of ANFs in the cochlea. While studies in mice have reported reliable and robust wave I amplitude depletion post noise exposure that persists despite recovery of audiometric and DPOAE thresholds, wave I has not proven to be a reliable biomarker for noise-induced cochlear synaptopathy in humans; wave I can be difficult to measure reliably even in humans that have perfectly normal hearing [13,18,11,15,3].

Individuals with histories of moderate noise exposure, but normal thresholds, can develop decreased ability to discriminate sounds in complex environments [30]. Numerous studies demonstrate the importance of binaural cues for localization and accurate perception of sounds like speech in the presence of noise [4]. Animal studies of synaptopathy to date have not measured binaural hearing abilities or binaural brainstem function**.** Brainstem binaural function can be assessed non-invasively via binaurally evoked ABRs by measuring the binaural interaction component (BIC) amplitude. The BIC is a residual waveform resulting from subtracting the sum of the monaurally evoked ABRs from the binaurally evoked ABRs. The BIC has been studied as a biomarker for binaural hearing abilities [2,6,7,17,33,35] and correlates with biaural behavioral abilities in normal and hearing-impaired subjects [17]. The BIC is known to vary with the two cues to horizontal sound localization: interaural time difference (ITD) and interaural level difference (ILD). These acoustical cues are generated by the filtering of sound by the head and body [9,10,23]. The BIC amplitude is maximum at 0 µs ITD, corresponding to a sound source directly in front of a subject, and systematically decreases in amplitude with increasing ITD and ILD [24,32]. Because the BIC and its source is the same in animals and humans [2,31], this physiological marker, while difficult to measure in humans, is a reliable

robust measurement in animals that we can track longitudinally and correlate with behavioral measures of binaural hearing to assess the effect of synaptopathy on binaural brainstem function for the first time.

Both the prevalence of cochlear synaptopathy in humans and its contributions to perceptual deficits remain unknown [25]. However, studies in post-mortem temporal bones have described IHC-ANF synapse loss with intact hair cells, suggesting that synaptopathy *can* occur in humans, despite its being unreliably detected using common clinical assays [34, 22]. Animal models of synaptopathy have been essential to understanding this pathology. However, all noise-induced synaptopathy studies in animals have focused on monaural hearing impairment despite implying binaural perceptual deficits [13, 15, 16,25]. We hypothesize that noise induced cochlear synaptopathy causes reduced peripheral temporal coding of sound leading to impaired binaural brainstem circuit function and spatial hearing abilities. The long-term sequelae of cochlear synaptopathy is a depletion of auditory nerve fibers; however, virtually nothing is currently known about how synaptopathy causes changes in the central auditory system and how these changes might cause auditory dysfunction. Here we test the hypothesis that noise-induced cochlear synaptopathy alters binaural brainstem circuit function and thus impacts binaural and spatial hearing capabilities. Our results establish a comprehensive understanding of noise-induced damage to the central auditory system despite normal audiograms.

METHODS

The noise exposure consists of a 4-8 kHz octave band noise at 106 dB SPL for 2 hours. 106 dB SPL is a commonly encountered level; it is approximately the noise level of a lawn mower. Based on prior studies in guinea pigs this protocol should induce cochlear synaptopathy [8,20,21,28,29]. We use 6-month-old adult pigmented male guinea pigs (*Cavia porcellus*). Animals are awake during the exposure. Experimental procedures complied with the guidelines of the University of Colorado Anschutz Medical Campus Animal Care and Use Committees and the National Institutes of Health.

ABR and DPOAE recordings were made prior to, 48 hours post, and 1, 2, 4, and 8 weeks following noise exposure. For ABR and DPOAE measurements, guinea pigs are anesthetized via an intramuscular injection of 70mg/kg ketamine and 7mg/kg xylazine. All experiments were performed in double-walled sound attenuating chambers (ETS Lindgren, Bronx, NY). ABR methods are described in Benichoux et al. (2018) and are only briefly described here. Stimuli were designed in custom built MATLAB scripts and generated evoked potentials recorded via an RME Fireface UCX soundcard (RME, Haimhausen, Germany). ABRs are collected via a TDT Rz4 4-channel amplifier (Tucker-Davis Technologies, FL, USA) using platinum subdermal needle electrodes (F-E2-12, Grass Technologies, West Warwick, RI). Here we need to obtain both reliable monaural wave 1 amplitude as well as BIC amplitude which require different electrode placements on the scalp. To do this, the TDT Rz4 simultaneously records monaural ABRs for each ear and binaural ABRs (for BIC) by having have one reference, one ground, and 3 recording electrodes, one on each mastoid and one on the nape. An apex electrode is placed between the ears and a ground electrode is placed on the hind leg. Recording monaural ABRs from the left and right ears simultaneously with binaural ABRs is novel in animal studies of synaptopathy. ABR thresholds are measured using tones (5 ms, 0.5 ms rise/fall), at a rate of ~30-40 Hz, at 1, 2, 4, 8, and 16 kHz, and at levels from 90 down to 20 dB SPL (lower if necessary) in 10 dB steps. Sound is generated with TDT MF1 speakers calibrated using Etymotic ER-7C probe mics (2,7). Both monaural and binaural (with 0 μs ITD) tone pips are used with 1000 repetitions for each condition to obtain a clean signal to noise ratio.

To non-invasively assess whether synaptopathy has altered binaural brainstem function, the BIC was measured at varying frequencies, intensities and at varying ITDs. Stimuli consisted of both tone-pips with varying intensities and frequencies and click stimuli with varying ITD. The clicks are presented at a rate of ~33 Hz. The ITD of the binaural signal was varied between ±1500 μs in steps of 250 μs. 1000 repetitions of the signal were presented for each stimulus condition.

DPOAEs were measured before and after noise exposure at the same time points as ABRs to establish any change and/or permanent loss of OHC function. DPOAE data was acquired with EMAV Otoacoustic Emssion Averager developed at Boys Town National Research Hospital [19]. Etymotic ER-10B+ microphone and Etymotic ER2 earphones are used to produce the stimulus tones and to measure the emissions. An RME Fireface UCX Sound card generates and records the signals. DPOAEs were recorded at six f2 frequencies (same as ABR): 1, 2, 4, 8, and 16 kHz; the ratio f2/f1 is set to 1.22. Each pair of stimulus frequencies was measured at 10 different sound pressure levels. The measurements begin by setting the level of f2 (L2) to 60 dB SPL and lowering it, in steps of 5 dB, to 15 dB. The level of f1 (L1) is determined by the formula L1 = 0.46*L2+41.

Our lab has demonstrated that the pre-pulse inhibition (PPI) of the acoustic startle response can reliably measure hearing abilities in guinea pigs [9]. Our lab has shown that guinea pigs, like humans, use both ITD and ILD to localize sound [9,10]. We can measure the ability of guinea pigs to discriminate sources of high-pass sounds, elucidating their use of ILDs, low-pass sounds, examining their use of ITDs, and the detection of sounds in the presence of background noise (spatial release from masking) by measuring prepulse inhibition of the acoustic startle reflex to study the effect of noise-induced synaptopathy on binaural and spatial hearing abilities. The method works as follows. Animals startle in response to an unanticipated brief (~5 ms) but loud (~110 dB) sound. The brevity of the sound ensures that it will not cause a hearing loss. The animal is placed forward-facing in a custom-built acoustically- transparent cage on top of a polycarbonate platform. Pre-pulse stimuli are played from one of the 25 loudspeakers (Morel MDT-20) spaced along a 1-m radius semicircular boom at 7.5° increments from -90° (right) to +90° (left) in front of the animal at 0° elevation. Animals are tested in the dark and monitored by IR camera to ensure they are oriented in the cage facing the center speaker (0° azimuth). Startle stimuli are played from a FaitalPRO HF102 compression driver positioned 25 cm above the animal. At the start of a test session, there is a ~2-minute acclimation phase for the animal to sit in on the platform in low-background noise. The magnitude of the startle is measured with an accelerometer placed on the bottom of the animal platform. If some detectable aspect of the sensory environment changes just a few 10s of ms prior to the delivery of the startle sound, such as a sound from a speaker in front of the animal, the startle response is inhibited and the reduction scales with the magnitude of the sensory cue change [9,37]. The PPI task can be used to test spatial hearing abilities before and after noise exposure.

We used PPI to measure hearing in noise ability, i.e., spatial release from masking, by having the animal detect a target sound (broadband chirps, see [9]) at varying levels from one spatial location in the presence of background noise presented from other spatial locations at a 70 dB SPL (Fig 5). The chirp is always played from the center speaker, towards which the animal faces. This test relates theoretically to our daily hearing requirements, as we focus on a talker while trying to ignore background noise. The angles of the masker (background noise) alternated from -22.5° (right) to 22.5° (left) in 7.5° increments because normal hearing guinea pigs can distinguish a change in sound location of broadband noise of 7.5° or less [9,10].

Lastly, we must confirm cochlear synaptopathy by visualizing a loss of ribbon synapses between IHCs and ANFs via immunohistochemistry. Cochleae are removed and bathed in 4% paraformaldehyde solution (PFA) for 2 hours and then decalcified in ethylenediaminetetraacetic acid (EDTA) for 4 weeks (EDTA changed daily). Once decalcified, cochleae are cut into basal, middle, and apex portions, and placed in 500 µl phosphate buffered saline (PBS) for antibody protocol. PBS is replaced with ~200 µl of Blocking solution (1% Triton X-100, 5% normal goat serum (NGS)) and placed on a rotator for 1 hr. Blocking solution is replaced with 100 µl of primary antibody solution (1% Triton X-100, 1% NGS), including Myosin VIIa (Proteus Biosciences #25-6790, 1:100 dilution) to label hair cells, mouse anti-Ctbp2 (C terminal binding protein 2) IgG1 (BD Biosciences #612044, 1:50 dilution) to label pre-synaptic ribbon synapses, and GluA2 (Millipore #MAB397, 1:1000 dilution) to label post-synaptic ribbon synapses. Samples are incubated overnight at 37°C. Antibody solution is removed, and samples undergo three 5-minute washes with PBS. PBS is replaced with ~100 µl of secondary antibody solution, Alexa Fluor 647 (Thermo Fisher #A21245, 1:200 dilution), Alexa Fluor 568 (Thermo Fisher #A21124, 1:1000 dilution), and Alexa Fluor 488 (Thermo Fisher #A21131, 1:500 dilution), protected from light, and incubated 1 hr at 37°C. The PBS washes and 1hr secondary incubation was repeated for amplification. Secondary antibody was followed by three 5-minute PBS washes in dark. Samples are mounted onto slides with ProLong Glass Antifade Mountant, ribbon synapses viewed on Zeiss LSM780 confocal with a 63x oil immersion lens, 1.40 NA.

RESULTS

The 2-hour noise exposure caused a temporary threshold shift as seen in the ABR threshold data (Fig. 1). DPOAE data at 8 kHz for the right ear of one animal is also shown in in Figure 1, only 4 timepoints are shown to simplify the data. DPOAEs are never lost completely (signal stays above the noise floor at al timepoints) and recover by 2 weeks post noise. Studies have reported that in noise-induced synaptopathy ABR-based thresholds and DPOAE measurements using similar methodology typically recover within 2 weeks post exposure [8,20], preliminary data (Fig. 1) confirms ABR recovery within 2 weeks following noise exposure and little change in DPOAEs. This also shows we induced no permanent hearing loss.

ABR wave I and BIC amplitudes were be examined to determine any monaural contributions to binaural processing; e.g., reduced amplitudes of these waves indicates reduced synchrony and/or numbers of neurons contributing to responses. Wave I measured in 2 guinea pigs at 90dB at 16kHz across timepoints is shown in Fig. 2. Previous studies have shown that Wave I is a reliable measure of cochlear synaptopathy in small rodents, but has not proven useful in humans. In our guinea pigs, Wave 1 showed a temporary loss that recovered by 8-weeks post noise exposure,

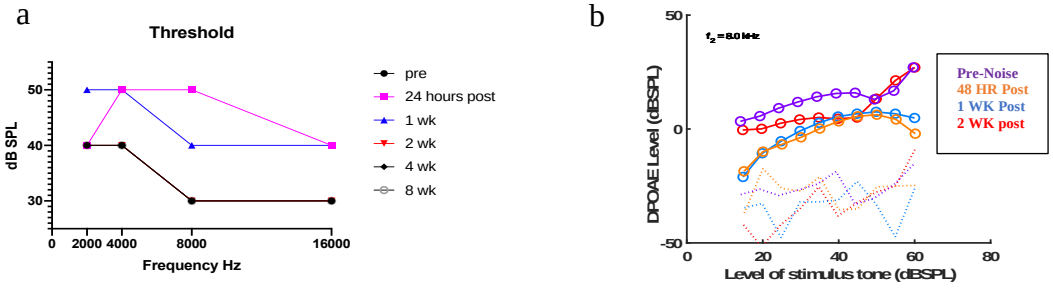


FIGURE 1. ABR and DPOAE measurements pre and post exposure in one animal. ABR measured auditory thresholds across frequencies showing recovery after a temporary threshold shift (a). DPOAE measurements relative to the noise floor across 4 timepoints are shown for f2 frequency of 8 kHz (b).

demonstrating that while Wave 1 may change immediately following noise, longitudinally it is a not a robust measure of cochlear synaptopathy. Binaural ABR data was also measured and BIC amplitude at 90 dB across the hearing frequency range is shown. Figure 3 shows the BIC measured in 2 animals at 16kHz across all timepoints. The BIC is depleted and never fully recovers through the 8-week timepoint, suggesting persisting binaural processing deficits at the level of the brainstem. Figure 4 shows BIC measured at varying ITDs at all timepoints. ITD is one cue to sound

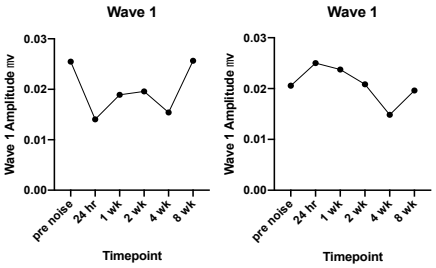


FIGURE 2. ABR wave 1 amplitude measured at 90dB in 2 animals at 16kHz. Amplitudes shown across timepoints and show recovery.

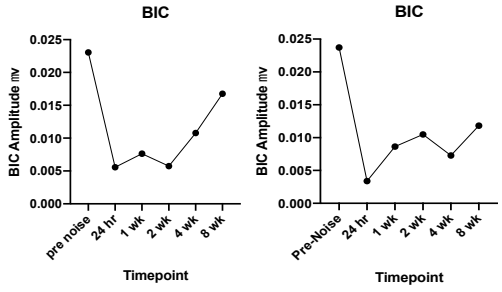


FIGURE 3. ABR BIC amplitude measured at 90dB in 2 animals at 16kHz. Amplitudes shown across timepoints and show depletion of the BIC.

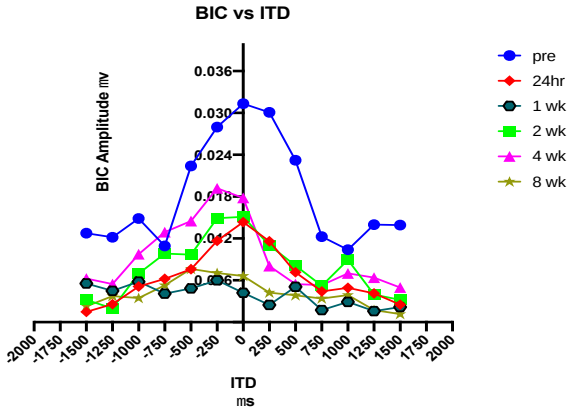


FIGURE 4. ABR BIC vs ITD at 90dB measured with click stimuli. Absolute value of BIC amplitude is plotted. BIC amplitude decreases with increasing ITD and diminishes over time post-noise exposure.

localization and the BIC is known to decrease with increasing ITDs. This data shows that BIC amplitude does decrease as ITDs increase and is depleted over time post noise-exposure exposing lasting binaural hearing deficits post moderate noise.

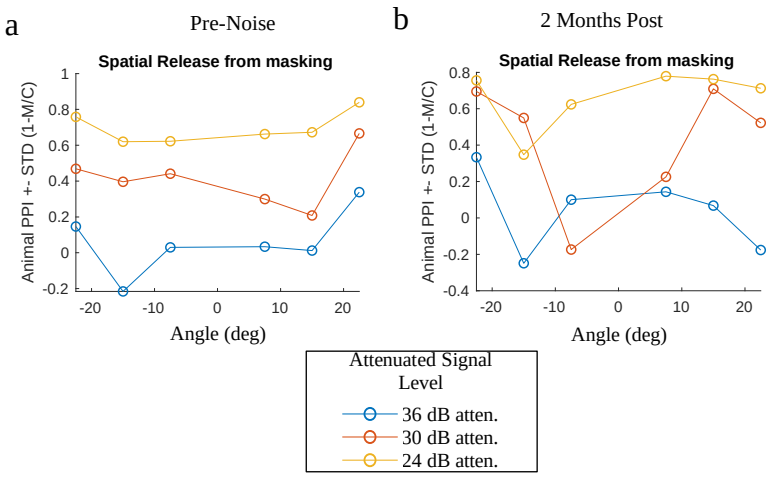


FIGURE 5. Spatial release from masking PPI task pre (a) and post (b) noise exposure.

Figure 5 shows an animal's average detection of a chirp signal in the presence of 70dB background noise presented at different locations around the animal. The target signal played from the center speaker had 3 different attenuation levels and was most easily detected (higher PPI values nearing 1.0) at 24 dB SPL attenuation (the loudest signal, yellow) and for the largest separations between noise and target (+/- 22.5°) as expected. As the background noise swaps between speakers at smaller angles along the horizontal plane, along with a softer chirp signal (higher attenuation), it should be more difficult for the animal to perceive and should result in a larger startle. Figure 5a shows pre-noise exposure hearing in noise ability for one animal. The animal's minimal audible angle is 7.5 degrees for the two louder signals (yellow and red datapoints). Figure 5b shows post-noise exposure hearing in noise ability for the same animal. 2 months post-noise the same animals can no longer detect the 30db attenuated signal (red) until the background noise is 15 degrees from the signal. The preliminary data in Fig. 5 suggests that spatial hearing worsens. These experiments assess how cochlear synaptopathy affects spatial hearing in noise behaviorally and indicate behavioral spatial hearing deficits post noise exposure.

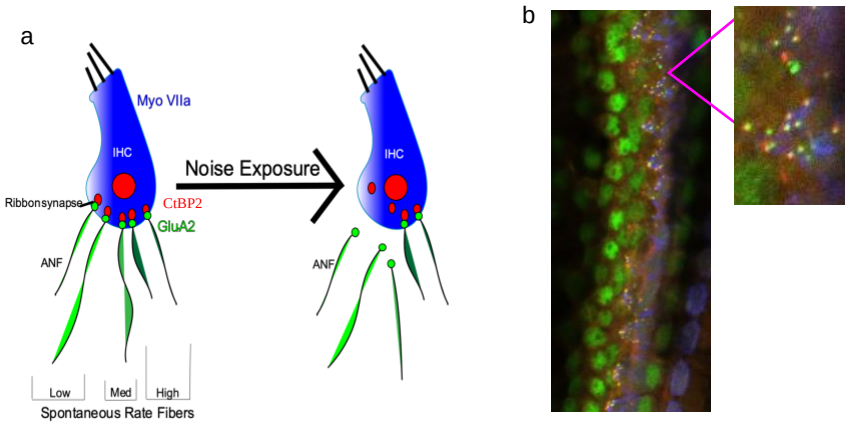


FIGURE 6. A schematic predicted cochlear histology post noise exposure (a) and example cochlear histology using pre- and post-synaptic markers in a noise-exposed guinea pig cochlea (b).

Figure 6a shows a schematic of our expectation regarding synaptopathy, that the pre (CtBP2) and post (GluA2) synaptic markers will be separated, and post synaptic marker will be lost post noise exposure. Figure 6b shows

preliminary cochlear histology. We can visualize the pre and post synaptic markers easily, next we will quantify the colocalization of these to determine synapse loss in our model of noise-induced cochlear synaptopathy.

CONCLUSIONS

No cochlear synaptopathy study to date has examined binaural hearing deficits. Here, we study a physiological binaural hearing marker, the BIC, and measured spatial hearing behavioral tasks to understand the central auditory deficits in a guinea pig model of synaptopathy. Monaural hearing showed no lasting deficits 8-weeks post noise, but the BIC amplitude was depleted and spatial hearing deficits were seen in these noise-exposed animals indicating central auditory deficits caused by cochlear synaptopathy. Recent studies [12,28] have shown synapse recovery post moderate noise exposure; this recovery happens mostly between 4- and 8-weeks post noise. Our physiological measurements were taken to 8 weeks post noise and our behavior was done following that timepoint to account for this possibility of synapse recovery. Histology and subsequent synapse counts are completed at least 20 weeks post noise exposure. Synapse counts are ongoing.

ACKNOWLEDGMENTS

This work is supported by NIH grants T32 DC012280 and R01 DC011555.

REFERENCES

[1] Abdala and Visser- Dumont (2001). Distortion Product Otoacoustic Emissions: A Tool for Hearing Assessment and Scientific Study. Volta Rev; 103(4): 281–302.

[2] Benichoux V, Ferber A, Hunt S, Hughes E, Tollin D (2018). Across Species "Natural Ablation" Reveals Brainstem Sources of a Noninvasive Biomarker of Binaural Hearing. Journal of Neuroscience, 38(40):8563-8573.

[3] Bharadwaj HM, Mai AR, Simpson JM et al. (2019). Non-invasive assays of cochlear synaptopathy-candidates and considerations. Neuroscience doi: 10.1016/j.neuroscience.2019.02.031

[4] Bronkhorst, A. W. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions, Acustica 86, 117–128.

[5] Coad, G., Long, G. R., Welch, D., and Thorne, P. R. (2013). Relationship of distortion product otoacoustic emission components to psychoacoustic measures of noise induced hearing loss. Journal of the Acoustical Society of America, 19, 1-9.

[6] Dobie, R.A., Berlin, C.I., (1979). Binaural interaction in brainstem-evoked responses. Archives of Otolaryngology, 105:391-398.

[7] Ferber AT, Benichoux V, Tollin DJ (2016). Test-Retest Reliability of the Binaural Interaction Component of the Auditory Brainstem Response. Ear Hear, 37(5):e291-e301

[8] Furman, A. C., Kujawa, S. G., and Liberman, M. C. (2013). Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates. Journal of Neurophysiology, 110(3), 577–86.

[9] Greene NT, Anbuhl KL, Ferber AT, DeGuzman M, Allen PD, Tollin DJ (2018). Spatial hearing ability of the pigmented Guinea pig (Cavia porcellus): Minimum audible angle and spatial release from masking in azimuth. Hear Res 365:62 76.

[10] Greene NT, Anbuhl KL, Williams W, Tollin DJ (2014). The acoustical cues to sound location in the guinea pig (Cavia porcellus). Hear. Res;316:1–15

[11] Guest, H, Munro, K, Prendergast, G, Plack, CJ (2019). Reliability and interrelations of seven proxy measures of cochlear synaptopathy. Hear Res 375:34–43

[12] Hickman, T. T., Hashimoto, K., Liberman, L. D., & Liberman, M. C. (2021). Cochlear Synaptic Degeneration and Regeneration After Noise: Effects of Age and Neuronal Subgroup. *Frontiers in cellular neuroscience*, *15*, 684706. https://doi.org/10.3389/fncel.2021.684706

[13] Hickox AE, Larsen E, Heinz MG, (2017). Translational issues in cochlear synaptopathy. *Hear Res.* 349:164–171.

[14] Hind SE, Haines-Bazrafshan R, Benton CL, Brassington W, Towle B, and Moore DR (2011). Prevalence of clinical referrals having hearing thresholds within normal limits. International Journal of Audiology 50(10):708–716.
[15] Kohrman DC, Wan G, Cassinotti L, Corfas G. (2019). Hidden Hearing Loss: A Disorder with Multiple Etiologies and Mechanisms. Cold Spring Harb Perspect Med. pii: a035493.
[16] Kujawa SG, and Liberman MC (2009). Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss. J Neurosci 29: 14077– 14085.
[17] Laumen, G., Ferber, A.T., Klump, G.M. & Tollin, D.J. (2016). The physiological basis and clinical use of the binaural interaction component of the auditory brainstem response. *Ear Hearing*, 37, e276– e290.
[18] Le Prell, CG (2018). Effects of noise exposure on auditory brainstem response and speech in noise tasks: a review of the literature. International Journal of Audiology. DOI: 10.1080/14992027.2018.1534010
[19] Liberman MC, Dodds LW (1984). Single-neuron labeling and chronic cochlear pathology. II. Stereocilia damage and alterations of spontaneous discharge rates. Hear Res 16: 43–53.
[20] Lin, H. W., Furman, A. C., Kujawa, S. G., and Liberman, M. C. (2011). Primary neural degeneration in the guinea pig cochlea after reversible noise-induced threshold shift. JARO - Journal of the Association for Research in Otolaryngology, 12, 605–616.
[21] Liu L, Wang H, Shi L, Almuklass A, He T, Aiken S, Bance M, Yin S, Wang J. (2012). Silent damage of noise on cochlear afferent innervation in guinea pigs and the impact on temporal processing. PLoS ONE 7: e49550. doi:10.1371/ journal.pone.0049550
[22] Makary, C.A., Shin, J., Kujawa, S.G., Liberman M.C., Merchant, S.N. (2011). Age-related primary cochlear neuronal degeneration in human temporal bones. J Assoc Res Otolaryngol 12, 711e717.
[23] Maier JK, Klump GM (2006). Resolution in azimuth sound localization in the Mongolian gerbil (Meriones unguiculatus). J Acoust Soc Am.119(2):1029-36.
[24] McPherson DL, Starr A (1995) Auditory time-intensity cues in the binaural interaction component of the auditory evoked potentials. Hear Res 89:162–171.
[25] Plack, C. J., Barker, D., & Prendergast, G. (2014). Perceptual consequences of "hidden" hearing loss. *Trends in hearing*, *18*, 2331216514550621.
[26] Saunders, J. C. (1985). The anatomical consequences of acoustic injury: A review and tutorial. The Journal of the Acoustical Society of America, 78(3), 833.
[27] Saunders, G. H., and Haggard, M. P. (1989). The clinical assessment of obscure auditory dysfunction--1. Auditory and psychological factors. Ear and Hearing, 10, 200–208.
[28] Shi L, Liu L, He T, Guo X, Yu Z, Yin S, Wang J (2013). Ribbon synapse plasticity in the cochleae of guinea pigs after noise-induced silent damage. PLoS One. 8(12); e81566
[29] Song Q, Shen P, Li X, Shi L, Liu L, Wang J, Yu Z, Stephen K, Aiken S, Yin S, Wang J (2016). Coding deficits in hidden hearing loss induced by noise: the nature and impacts. Sci Rep; 6:25200.
[30] Stone, M. A., Moore, B. C. J., and Greenish, H. (2008). Discrimination of envelope statistics reveals evidence of sub-clinical hearing damage in a noise exposed population with "normal" hearing thresholds. International Journal of Audiology, 47, 737–750.
[31] Tolnai S, Klump GM. Evidence for the origin of the binaural interaction component of the auditory brainstem response. *Eur J Neurosci*. 2020;51(2):598–610.
[32] Ungan P, Yağcioğlu S, Ozmen B (1997) Interaural delay-dependent changes in the binaural difference potential in cat auditory brainstem response: implications about the origin of the binaural interaction component. Hear Res 106:66–82.
[33] Van Yper L.N., Vermeire K., De Vel E.F., Beynon A.J., Dhooge I.J. (2016) Age-Related Changes in Binaural Interaction at Brainstem Level. Ear & Hearing, Jul-Aug;37(4):434-42.
[34] Viana, L.M., O'Malley, J.T., Burgess, B.J, Jones DD, Oliveira CA, Santos F, Merchant SN, Liberman LD, Liberman MC (2015). Cochlear neuropathy in human presbycusis: confocal analysis of hidden hearing loss in post-mortem tissue. Hear. Res. 327, 78e88.
[35] Wrege, K.S., Starr, A. (1981), Binaural interaction in human auditory brainstem evoked potentials, Arch. Neurol. https://doi.org/10.1001/archneur.1981. 00510090066008.
[36] Zhao, F., and Stephens, D. (2007). A critical review of King-Kopetzky syndrome: Hearing difficulties, but normal hearing? Audiology Medicine, 5(2), 119-124.
[37] Longenecker, F., Alghamdi, F., Rosen M.J., Galazyuk, A.V. (2016). Prepulse inhibition of the acoustic startle reflex vs auditory brainstem response for hearing assessment. Hearing Research, 339, 80-93.

Whole Stimulus DPOAE Analysis

Paul D. Teal,[1, a)] Christopher A. Shera,[2, b)] and Carolina Abdala[2, c)]

[1)] *Victoria University of Wellington, New Zealand*

[2)] *University of Southern California*

[a)] *Corresponding author: paul.teal@vuw.ac.nz*

[b)] *Electronic mail: christopher.shera@usc.edu*

[c)] *Electronic mail: carolina.abdala@med.usc.edu*

Abstract. Distortion Product Otoacoustic Emissions (DPOAE) offer great potential for hearing diagnosis, but are complicated by the interaction of components generated by different mechanisms. Separation of these components from DPOAE measurements may allow conclusions to be drawn about the functionality of these separate mechanisms of a cochlear. However, the signal processing methods for performing this separation are imperfect. Existing methods are based on time windowing of DPOAE generated from frequency sweep stimuli.

This paper presents a method in which the entire spectra of both distortion (D) and reflection (R) components are simultaneously estimated. This approach has several advantages. Firstly, the method removes the need for a compromise between frequency precision and signal to noise ratio. Secondly, the method can be made to include models of the stimulus signals, so that the stimulus does not strongly interfere with the estimation process. Thirdly, the method can be arranged to make efficient use of data that has been corrupted by measurement artefacts. Fourthly, the method can be easily adapted to track DPOAEs that are changing in response to chemical or acoustic treatments.

The basic modelling assumptions made are that the sum of R and D measurements can be represented as the sum of convolutions with the stimulus signal, the frequency representation of the D component is more smooth than the R component, and that a reasonable estimate of the noise level in the signal is available. These assumptions are combined into a linear convex problem.

In this paper we compare the proposed approach with three other methods. While it is not superior to the earlier methods at every frequency, it does offer some improvement, particularly with regards reducing the contamination of D by R.

INTRODUCTION

Distortion product otoacoustic emissions (DPOAE) are widely used for research and clinical practice, the latter particularly in infant hearing screening programs. DPOAE typically exhibits rapid variation with frequency, which has been termed "fine structure" and is thought to arise from the interaction of two propagating waves [1, 2, 3, 4]. The first of these is the distortion (D) component, which is generated near the characteristic place of the higher of the two stimulus frequencies, and which propagates both backwards towards the stapes, and forwards towards the apex. The second, reflection (R), component arises from the reflection of the first component near the characteristic place of the distortion product frequency. Separation of these two components considerably enhances the information contained in the DPOAE for both research and diagnostic purposes [5]. Typically this separation has been performed by time windowing of DPOAE generated from frequency sweep stimuli [6, 7], usually accompanied by a least-square fit [8]. In effect each window generates an estimate of each of the D and R components in a narrow frequency band.

This paper presents a method in which the entire spectra of both D and R components are simultaneously estimated. This approach has several advantages. Firstly, the method removes the need for a compromise between high frequency precision (obtained by using a narrow window) and high signal to noise ratio (obtained by using a wide window). Secondly, the method can be made to include models of the stimulus signals, so that these do not interfere with the estimation process (although in the method presented here, these models are not included). Thirdly, the method can be arranged to make efficient use of data that has been corrupted by measurement artefacts. Fourthly, the method can be easily adapted to track DPOAEs that are changing in response to chemical or acoustic treatments [9, 10].

In this paper we compare the proposed approach with three other methods: least squares fitting (LSF) [8], inverse Fourier Transform (IFFT) [11, 12, 13], and wavelet transform (CWT) [14, 15] on simulated data using the approach of [7]. The proposed approach is not superior to the earlier methods at every frequency, but the whole-stimulus approach offers improvement in many parts of the extracted spectra, particularly with regards reducing the contamination of D by R.

METHOD

We make the following assumptions about the DPOAE measurement data and the process by which it is generated:

We first assume that the received DPOAE consists of just three components: a distortion component, a reflection component and normal white noise. We also assume that both of the first two components can be modelled as a linear time-invariant system: i.e., resulting from a simple convolution with a stimulus signal. This stimulus is a "virtual" stimulus signal, since the distortion product itself is not used as a stimulus, but it is assumed that since the components are known, this distortion product virtual stimulus is known.

The distortion component is assumed to be a relatively smooth function of frequency, whereas the reflection component may contain considerably greater variation with frequency.

We further assume that the DPOAE is sampled at a sufficiently high rate that all phenomena of interest can be represented by discrete time operations (i.e., temporal aliasing of any distortion products of interest does not occur).

In the following derivation, we further assume that the starting phases of the primary tones are rotated systematically between sweeps so that the stimuli are cancelled in the average while leaving the distortion product response unchanged (phase-rotation averaging [16]). While this is not a necessary assumption for the approach presented here to be effective, it simplifies the modelling process and presentation.

The observed signal s is given by the sum of measurement noise η and the convolutions of the virtual stimulus v with the time domain impulse responses of the distortion and reflection components d and r:

$$s = v * (d + r) + \eta = \left[\begin{array}{cc} V_\mathrm{d} & V_\mathrm{r} \end{array}\right] x + \eta \tag{1}$$

where $x = [\, d^* \;\; r^* \,]^*$ is a vector of the separated responses, where $()^*$ represents the adjoint (transpose) operator, and V_d and V_r are appropriately size convolution matrices generated from the stimulus v. It is typically advantageous to make d shorter than r because doing so automatically imposes a degree of smoothness on the frequency domain representation D of the distortion component (for conversion to the frequency domain, the vector representing d is zero-padded). Using the assumption that the noise η is normal and white, the maximum likelihood (ML) estimate [17] of x is given by the minimiser x of

$$\left\| \left[\begin{array}{cc} V_\mathrm{d} & V_\mathrm{r} \end{array}\right] x - s \right\|_2^2 \tag{2}$$

where $\| \cdot \|_2^2$ represents the square of the ℓ-2 norm, i.e., the sum of the squares of the elements of the given vector.

The observation s is usually the average of the data collected from many frequency sweeps. It is frequently the case that impulsive noise sources (such as movement of the experimental subject) can corrupt the collection of data. Traditionally the entirety of a corrupted sweep has been discarded from the averaging process. However, if we remove only the corrupted data points (although in sets of three sweeps if we rely on phase rotation averaging, as mentioned above), we can obtain an estimate of s in which different components have different noise variance. Deletion of a section of the response in one of the measurements means that the variance of the average of that section is larger. In this case, the ML estimate of x is obtained using diagonal weighting matrix W_s:

$$\min_x \left\| \left(\begin{array}{c} W_\mathrm{s} \end{array} \right) \left(s - \left[\begin{array}{cc} V_\mathrm{d} & V_\mathrm{r} \end{array}\right] x \right) \right\|_2^2 = \min_x \| s' - Y' x \|_2^2 \tag{3}$$

where $s' = W_\mathrm{s} s$ and $Y' = W_\mathrm{x}[V_\mathrm{d}\ V_\mathrm{r}]$.

The matrix Y' in (3) is at most only half-rank, so more information is required for x to be defined. The following paragraphs explain how the assumptions introduced at the start of this section can be expressed as numerical constraints that enable this ambiguity to be resolved.

The conventional method for improving the stability and resolving ambiguity of the solution to an inverse problem such as (3) is to use regularisation, which can take various forms. Commonly used regularisation terms include constraints on the signal energy, on the first or second derivatives, or on the signal entropy.

The energy regularised problem will typically be the minimiser x of a problem of the form:

$$\| s' - Y' x \|_2^2 + \lambda_0 \| x \|_2^2 \tag{4}$$

Frequency energy

We may have cause to believe that certain frequencies of the D and R DPOAE components contain more energy than others. In this case, a frequency weighted regularisation may be appropriate. This will take the form:

$$\|W_{\mathrm{d}}Fd\|_2^2+\|W_{\mathrm{r}}Fr\|_2^2=\left\|\begin{pmatrix} W_{\mathrm{d}}F_{\mathrm{r}} & 0 \\ W_{\mathrm{d}}F_{\mathrm{i}} & 0 \\ 0 & W_{\mathrm{r}}F_{\mathrm{r}} \\ 0 & W_{\mathrm{r}}F_{\mathrm{i}} \end{pmatrix} x\right\|_2^2=\|D_0'x\|^2 \tag{5}$$

where W_{d} and W_{r} are diagonal matrices containing entries that are the inverse of the expected energy levels in the D and R components respectively, and F_{r} and F_{i} are the real and imaginary parts respectively of an appropriately sized DFT matrix. This term is then multiplied by an appropriate regularisation scaling factor and added to the expression to be minimised.

Derivative regularisation: Regularisation using the first or second derivative is appropriate when it is known that the signals being estimated are smooth. One of our assumptions is that the D component is smooth in the frequency domain. If L_s is the length of vector s, then the first derivative can be approximated using a $(L_s-1)\times L_s$ first difference matrix, and the second derivative using a $(L_s-2)\times L_s$ second difference matrix:

$$D_1=\begin{bmatrix} -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & & 0 \\ \vdots & & \ddots & \ddots & 0 \\ 0 & 0 & \dots & -1 & 1 \end{bmatrix} \qquad D_2=\begin{bmatrix} 1 & -2 & 1 & 0 & \dots & 0 \\ 0 & 1 & -2 & 1 & & 0 \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 1 & -2 & 1 \end{bmatrix} \tag{6}$$

Minimisation of the first derivative of the *frequency* domain distortion component uses regularisation of the form:

$$\left\|\begin{pmatrix} D_1F_{\mathrm{r}} & 0 \\ D_1F_{\mathrm{i}} & 0 \end{pmatrix}\begin{pmatrix} d \\ r \end{pmatrix}\right\|_2^2=\left\|D_1'x\right\|_2^2 \tag{7}$$

and the second derivative has an analogous form.

The combined minimisation problem

Combining the minimisation of the squared error and several regularisation terms leads to a minimisation of an expression of the form:

$$\|Y'x-s'\|_2^2+\lambda_0\|D_0'x\|_2^2+\lambda_1\|D_1'x\|_2^2+\lambda_2\|D_2x\|_2^2 \tag{8}$$

which can alternatively be expressed as:

$$\left\|\begin{pmatrix} Y' \\ \sqrt{\lambda_0}D_0' \\ \sqrt{\lambda_1}D_1' \\ \sqrt{\lambda_2}D_2' \end{pmatrix}x-\begin{pmatrix} s' \\ 0 \\ 0 \\ 0 \end{pmatrix}\right\|_2^2=\|Ax-b\|_2^2 \tag{9}$$

The minimiser of (9) is the well known solution $x=(A^*A)^{-1}A^*b$. There are several issues with such a solution. The first is that the matrix A is likely to be far too large to represent in the memory of a standard computer. Even the smaller matrix A^*A may be too large to create. This problem could be overcome by the use of iterative matrix methods such as the conjugate gradient (or preconditioned conjugate gradient) method.

Another significant problem is the choice of balance of the regularisation terms. It is difficult to know in advance what choice of regularisation coefficients λ_i will lead to a good solution. Furthermore, for some choices of λ_i, (for instance, to ensure the smoothness with frequency of the D component) the problem can become poorly conditioned.

However, it is much more simple to find values that each of the constraints are likely to take. So rather than solving

$$\min_x\left(\|A_0x-b\|^2+\sum_{i=1}^{N}\lambda_i\|A_ix\|^2\right) \tag{10}$$

we would actually find it easier to solve a problem like:

$$\min_x\|A_0x-b\|^2 \qquad \text{s.t. } \|A_ix\|<\gamma_i \qquad i\in\{1,\dots,N\} \tag{11}$$

in which the signals are constrained to comply with certain bounds on their energies or roughness. It is comparatively easy from samples of data to find suitable values for γ_i in (11). However, solving (11) is typically performed using the method of Lagrange multipliers, which leads to an expression somewhat like (10), along with a multidimensional optimisation over the various coefficients λ_i, which can be computationally very difficult to solve.

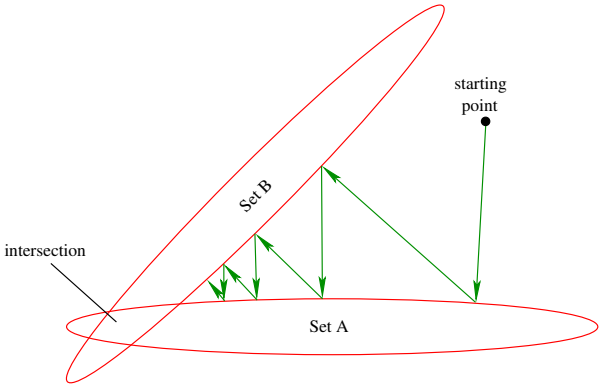


FIGURE 1. The *alternating projection* or *projection onto convex sets* method.

Projection onto convex sets (POCS)

Rather than minimising the square error, it actually sufficient to simply ensure that the squared error is no larger than is consistent with what is known about the noise level in the observations, since we have additional criteria with which we know that the signal must comply. We can thus express the required estimate as the solution to the problem

$$\text{find}_x \text{ s.t. } \|A_i x - b_i\|^2 \leq \gamma_i \qquad i \in \{1,\ldots,N\} \tag{12}$$

where most of the values of b_i are zero. The set of solutions for each of these constraints is a convex set. Such a problem can be solved using the method of *alternating projection* or *projection onto convex sets* [18] (POCS), the idea of which is illustrated in Fig. 1 for just two sets: the current solution is projected onto each of the constraining sets sequentially. Provided that the sets are convex, and an intersection of all the sets does indeed exist, the procedure is known to converge. It is not guaranteed to converge on the point of the intersection that is closest to the initialisation. Dykstra's algorithm [19] does provide this guarantee, and could have an advantage over POCS if we have an initial estimate x_0 of the solution that has some intrinsic merit, such as a solution obtained using a different method. This possibility is not investigated further in this paper.

The projection at stage i of the POCS algorithm (for each of the conditions of (12)), given the current solution x_0 from the previous stage, can be expressed as

$$\min_x \|x - x_0\|_2^2 \qquad \text{s.t. } \|A_i x - b_i\|_2 < \gamma_i \tag{13}$$

Attempts to solve this problem using Lagrange multipliers do not lead to useful results. However, (13) can equivalently be expressed as the proximity operator

$$\text{prox}_i(x) = \underset{x}{\text{argmin}} \left(I\left(\|Ax - b\| < \gamma\right) + \frac{1}{2}\|x - x_0\|_2^2 \right) \tag{14}$$

where $I()$ is the indicator function, taking the value 0 if its argument is true, and $+\infty$ otherwise. This is the form of problem which can be solved using the fast iterative shrinkage-thresholding algorithm (FISTA) [20], for which the objective function is $f(x) + g(x)$ where f is smooth and convex (in this case, the square of the distance $\|x - x_0\|$), and g is convex, but not necessarily smooth. In this case, the indicator function is not smooth.

The FISTA algorithm can be expressed as repetition (over the index k) of the steps:

$$x_k = \underset{x}{\text{argmin}} \left\{ g(x) + \frac{L}{2} \left\| y_k - \frac{1}{L}\nabla f(y) \right\|_2 \right\} \tag{15}$$

$$t_{k+1} = \frac{1 + \sqrt{1 + 4t_k^2}}{2} \tag{16}$$

$$y_{k+1} = x_k + \frac{t_k - 1}{t_{k+1}}(x_k - x_{k-1}) \tag{17}$$

Constraint	γ	L
Data consistency (20 dB SNR)	3 000	6×10^{10}
Data consistency (10 dB SNR)	9 000	6×10^{10}
Data consistency (6 dB SNR)	14 000	6×10^{10}
Energy of D	100	60 000
Energy of R	20	60 000
First difference of D	0.4	3
First difference of R	0.5	300
Second difference of D	0.001	0.0002
Second difference of R	0.02	1

TABLE I. Data constraints and Lipschitz constants.

where L is a Lipschitz constant i.e., a positive real number for which $\|Ax\|^2 \leq L\|x\|^2$ for all x.

To be able to solve the first of these stages, we need to be able to perform multiplication by both A and its adjoint: i.e., to be able to calculate Ax and A^*b for any x and b. For the squared error and regularisation terms, it is easy to perform multiplication by A using fast Fourier transforms and difference operations, without creating the very large matrix A. Multiplication by the adjoint A^* can also be accomplished without creating the matrix using the fact that the DFT matrix is (scaled) unitary i.e., $F^{-1} = \frac{1}{N}F^{*T}$. Multiplication by the transpose of the first difference matrix $D_{1,N}$ and by the transpose of the second difference matrix $D_{2,N}$ for length N sequences is performed by the operations:

$$D^*_{1,N} = \left[\begin{array}{c} -1 \quad 0 \ \ldots \ 0 \\ \hline -D_{1,(N-1)} \\ \hline 0 \quad \ldots \quad 0 \ 1 \end{array}\right] \qquad D^*_{2,N} = \left[\begin{array}{c} -1 \ \ 0 \quad \ldots \ 0 \\ -2 \ \ 1 \quad 0 \ \ldots \ 0 \\ \hline D_{2,(N-2)} \\ \hline 0 \ \ldots \quad 0 \ \ 1 \ -2 \\ 0 \ \ldots \quad \ 0 \ \ 1 \end{array}\right] \tag{18}$$

both of which can also be readily performed without creating the matrix.

The FISTA algorithm requires knowledge of the Lipschitz constants L for each of the sets. Iteration of the process $x_n = A^*Ax_n/\|A^*Ax_n\|$ rapidly converges on the eigenvector corresponding to the largest eigenvalue of A^*A, and $L_n = \|x_n\|/\|x_{n-1}\|$ to the largest eigenvalue, which is the required Lipschitz constant. This process is straightforward since we have already seen that multiplication by both A and A^* can be readily performed.

Constraint values: Sets of DPOAE sweep waveforms spanning 0.5–8 kHz were simulated using the method of [21]. The model parameters were varied randomly (variance 25%) from means chosen to match normative human data [22, 23]. Analysis of the properties of the D and R components suggests the constraints presented in Table I. It is worth noting that the smoothness of the D component appears to be more different from the R component when compared using the second derivatives than when comparing the first derivatives. In real DPOAE measurements the noise floor can be estimated using the difference between adjacent sweep pairs.

RESULTS

The algorithm was evaluated for many different simulated signals, and multiple signal-to-noise ratios (SNR) between 0 and 20 dB. In most cases, compliance with most of the the specified constraints was quickly obtained. The exceptions were the data consistency constraint, and the energy of R constraint. These constraints appeared to be less compatible, and hence most of the the computation time until convergence was spent alternating between just these two constraints.

One typical result for SNR of 6 dB is shown in Figure 2. The estimation of D appears to be quite robust, even for much lower SNR than 6 dB. This is also true of the *shape* of the R component, but the amplitude of the R component tends to be under-estimated at low SNR.

The unmixing errors were expressed as frequency dependent coefficients δ and ρ using the method of [7]. $|\delta|$ can be described as indicating the level of contamination of R by D, and $|\rho|$ as indicating the level of contamination of D by R. The logarithm of these values, averaged over 15 sets of synthetic DPOAEs are shown in Figure 3, along with the same calculation for the least square fit (LSF) [8], inverse Fourier transform (IFFT) [11] and continuous wavelet transform (CWT) [14, 15] methods. It is clear that for a large part of the frequency range, the value of $|\rho|$ for the proposed (POCS) method is considerably lower than the other methods, suggesting that the D component is being

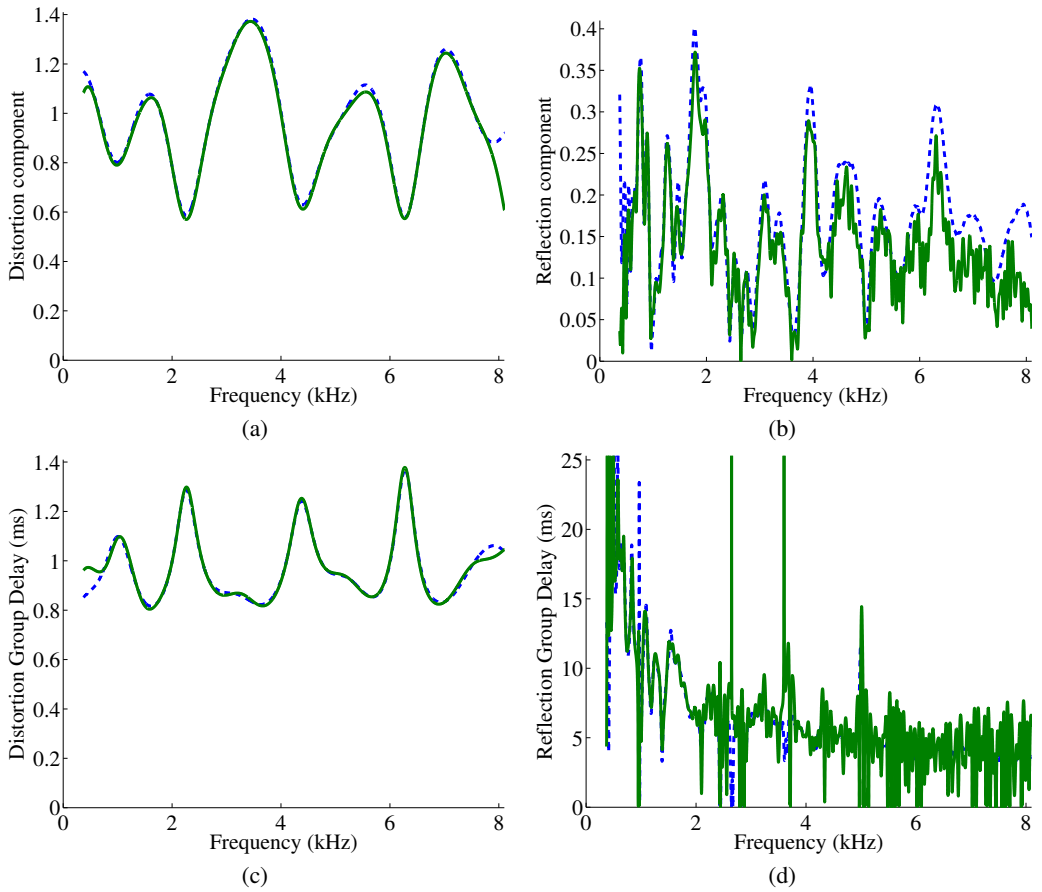


FIGURE 2. Results for observation SNR of 6 dB. The true (simulated) values are dashed, while the estimated values are solid curves. (a) and (b) show the actual and estimated distortion components, while (c) and (d) show the corresponding group delays.

more reliably estimated. There does not appear to be the same obvious advantage for $|\delta|$, but it is at least comparable to the other methods, except for frequencies under about 1.2 kHz.

CONCLUSIONS

It has been shown that the distortion and reflection components of DPOAEs can be separated on the basis of their relative energies and smoothness. No windowing of the data is required.

On the basis of a comparison with other methods based on synthetic data, the method appears to have an accuracy advantage in the estimation of the distortion (D) component.

An advantage of the proposed method is that appropriate weighting of the data fidelity criterion means that only the data points damaged by artefacts need to be discarded, rather than entire sweeps.

An important observation is that minimising the squared error of the reconstructed signal is not necessarily the best estimation strategy. Reducing the square error below the variance level of the noise is counter productive if there are other properties of the signal that can guide the separation process.

REFERENCES

1. M. Whitehead, B. Lonsbury-Martin, and G. Martin, "Evidence for two discrete sources of $2f_1 - f_2$ distortion-product otoacoustic emission in rabbit: I. differential dependence on stimulus parameters," J. Acoust. Soc. Am. **91**, 1587–1607 (1992), `https://doi.org/10.1121/1.402440`.

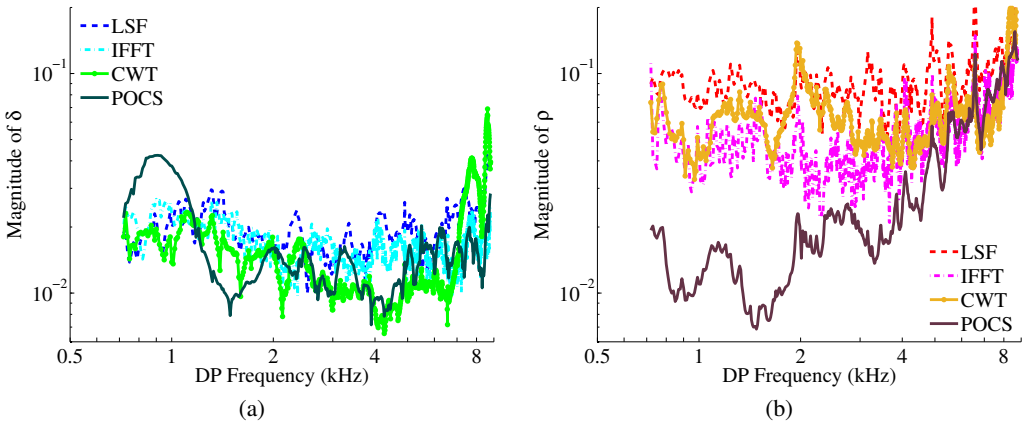


FIGURE 3. Demixing coefficients δ and ρ for the least squares fit (LSF), inverse Fourier transform (IFFT), continuous wavelet transform (CWT) and proposed projection onto convex sets (POCS) methods.

2. M. Whitehead, B. Lonsbury-Martin, and G. Martin, "Evidence for two discrete sources of $2f_1 - f_2$ distortion-product otoacoustic emission in rabbit: II. differential physiological vulnerability," J. Acoust. Soc. Am. **92**, 2662–2682 (1992), https://doi.org/10.1121/1.404382.
3. C. Talmadge, G. Long, A. Tubis, and S. Dhar, "Experimental confirmation of the two-source interference model for the fine structure of distortion product otoacoustic emissions," J. Acoust. Soc. Am. **105**, 275–292 (1999), https://doi.org/10.1121/1.424584.
4. A. Rao and G. Long, "Effects of aspirin on distortion product fine structure: Interpreted by the two-source model for distortion product otoacoustic emissions generation," J. Acoust. Soc. Am. **129**, 792–800 (2011), https://doi.org/10.1121/1.3523308.
5. C. Abdala, A. Ortmann, and C. Shera, "Reflection- and distortion-source otoacoustic emissions: Evidence for increased irregularity in the human cochlea during aging," J. Assoc. Research in Otolaryngology **19**, 493–510 (2018), https://doi.org/10.1007/s10162-018-0680-x.
6. R. Knight and D. Kemp, "Wave and place fixed DPOAE maps of the human ear," J. Acoust. Soc. Am. **109**, 1513–1525 (2001), https://doi.org/10.1121/1.1354197.
7. R. Kalluri and C. Shera, "Distortion-product source unmixing: a test of the two-mechanism model for DPOAE generation," J. Acoust. Soc. Am. **109**, 622–637 (2001), https://doi.org/10.1121/1.1334597.
8. G. Long, C. Talmadge, and J.Lee, "Measuring distortion product otoacoustic emissions using continuously sweeping primaries," J. Acoust. Soc. Am. **124**, 1613–1626 (2008), https://doi.org/10.1121/1.2949505.
9. G. Housley, R. Morton-Jones, S. Vlajkovic, R. Telang, V. Paramananthasivam, S. Tadros, A. Wong, K. Froud, J. Cederholm, Y. Sivakumaran, P. Snguanwongchai, B. Khakh, D. Cockayne, P. Thorne, and A. Ryan, "ATP-gated ion channels mediate adaptation to elevated sound levels," Proc. Nat. Acad. Sciences **110**, 7494–7499 (2012), https://doi.org/10.1073/pnas.1222295110.
10. J. Cederholm, A. Ryan, and G. Housley, "Onset kinetics of noise-induced purinergic adaptation of the 'cochlear amplifier'," Purinergic Signalling **15**, 343–355 (2019), https://doi.org/10.1007/s11302-019-09648-3.
11. C. Abdala, F. Guérit, P. Luo, and C. Shera, "Distortion-product otoacoustic emission reflection-component delays and cochlear tuning: estimates from across the human lifespan," J. Acoust. Soc. Am. **135**, 1950–1958 (2014), https://doi.org/10.1121/1.4868357.
12. C. Abdala, P. Luo, and C. Shera, "Optimizing swept-tones to measure DPOAEs in adult and newborn ears," J. Acoust. Soc. Am. **138**, 3785–3799 (2015), https://asa.scitation.org/doi/10.1121/1.4937611.
13. C. Abdala, P. Luo, and C. Shera, "Comparison of methods for DPOAE component separation," in *39th midwinter meeting of the Association for Research in Otolaryngology* (San Diego, 2016).
14. A. Moleti, F. Longo, and R. Sisto, "Time-frequency domain filtering of evoked otoacoustic emissions," J. Acoust. Soc. Am. **132**, 2455–2467 (2012), https://doi.org/10.1121/1.4751537.
15. G. Tognola, F. Grandori, and P. Ravazzani, "Time-frequency distributions of click-evoked otoacoustic emissions," Hear Res. **106**, 112–122 (1997), https://doi.org/10.1016/S0378-5955(97)00007-5.
16. M. Whitehead, B. Stagner, G. Martin, and B. Lonsbury-Martin, "Visualization of the onset of distortion-product otoacoustic emissions, and measurement of their latency," J. Acoust. Soc. Am. **100**, 1663–1679 (1996), https://doi.org/10.1121/1.416065.
17. S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice Hall, Englewood Cliffs, NJ, 1993).
18. H. Bauschke and J. Borwein, "On projection algorithms for solving convex feasibility problems," SIAM Rev. **38**, 367–426 (1996), https://dl.acm.org/doi/10.1137/S0036144593251710.
19. J. Boyle and R. Dykstra, "A method for finding projections onto the intersection of convex sets in Hilbert spaces," Lecture Notes in Statistics **37**, 28–47 (1986), https://doi.org/10.1007/978-1-4613-9940-7_3.
20. A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," SIAM J. Imaging Sciences **2**, 183–202 (2009), https://dl.acm.org/doi/10.1137/080716542.
21. C. Shera and C. Bergevin, "Obtaining reliable phase-gradient delays from otoacoustic emission data," J. Acoust. Soc. Am. **132**, 927–943 (2012), https://doi.org/10.1121/1.4730916.
22. C. Shera, J. Guinan, Jr., and A. Oxenham, "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements," Proc. Nat. Academy Sci. (PNAS) **99**, 3318–3323 (2002), https://doi.org/10.1073/pnas.032675099.
23. C. Abdala, S. Dhar, and R. Kalluri, "Level dependence of distortion product otoacoustic emission phase is attributed to component mixing," J. Acoust. Soc. Am. **129**, 3123–3133 (2011), https://doi.org/10.1121/1.3573992.

Temporal adjustment of interaural stimulation timing leads to improved sound localization but not to improved spatial release from masking in bimodal listeners

Stefan Zirn[1], Julian Angermeier[2], Werner Hemmert[3]

Author Affiliations

1University of Applied Sciences Offenburg, 2Faculty of Electrical Engineering, Medical Engineering and Computer Sciences, University of Applied Sciences Offenburg, Badstr. 24, 777652 Offenburg, Germany

2 Technical University of Munich

Author Email

Stefan.zirn@hs-offenburg.de

Abstract. In bimodal cochlear implant (CI) / hearing aid (HA) users a constant interaural time delay in the order of several milliseconds occurs due to differences in signal processing of the devices. For MED-EL CI systems in combination with different HA types, we have quantified the respective device delay mismatch.

In the current study, we investigate the effect of the device delay mismatch in actual bimodal listeners on sound localization accuracy. To deal with the device delay mismatch we delayed the CI stimulation according to the measured HA processing delay and two other values. To determine potential effects of procedural learning, we applied an A-B-B-A testing paradigm.

With all delay values highly significant improvements of the RMS error were observed compared to the test without the delay (14.7° improvement in average) in 9 bimodal listeners. Also the Signed Bias of sound localization improved significantly from 25.2° to 10.5° averaged across listeners.

At all times we also tested spatial release of masking (SRM) of speech (S) in noise (N). S and M were either collocated at 0° or S was situated at 0° and N at 90° at the HA side. In none of the bimodal listeners an improvement of SRM by the temporal adjustment of modalities was observed.

The results reveal that sound localization in bimodal listeners improves from a reduction of the device delay mismatch between CI and HA. With this form of temporal adjustment of modalities, spatial re-centering seems possible. However, transmission of temporal information via the CI and thus binaural processing seems to be not sufficient for improved SRM.

INTRODUCTION

Bimodal provision is very common in CI centers nowadays. In such cases one ear is provided with a cochlear implant (CI) and the contralateral ear with a conventional digital hearing aid (HA). Many studies were published showing a benefit for most bimodal listeners, when using both devices instead of just one in binaural performance (Ching et al., 2004, 2006; Hoppe et al., 2018; Sheffield et al., 2017) as well as in quality of life (Farinetti et al., 2015). Despite the reported benefits, Dorman and colleagues could show that in teRMS of sound source localization bimodal CI/HA users perform poorer than e.g. bilateral CI users or single sided deaf CI users (Dorman et al., 2016). One of many reasons for this poor performance is that both modalities are not synchronized in teRMS of processing delay. Zirn et al. (2015) showed that there can be a temporal asymmetry in the range of 3 to 10 milliseconds between the ear provided with a MED-EL CI and the ear provided with a HA.

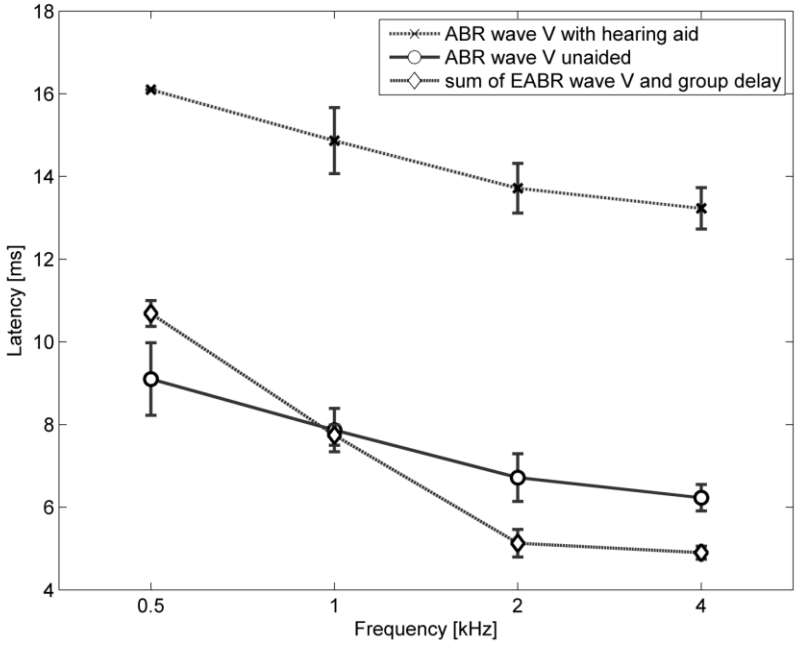


FIGURE 1: ABR and eABR wave V latencies for different frequencies. The curve with diamond markers stands for the sum of eABR wave V latencies and the group delay, i.e. processing delay, of the MED-EL MAESTRO CI system at the respective frequency. Figure adapted from Zirn et al. (2015)

This asymmetry is further referred to as device delay mismatch. In the latest study by Zirn and colleagues, a significant improvement in sound source localization ability could be reported for bimodal CI/HA users, when the device delay mismatch was reduced by delaying the CI stimulation (Zirn et al., 2019).
This paper reports a follow-up experiment to the Zirn et al.(2019) study, extending the A-B study design to an A-B-B-A design. Furthermore, not only the RMS error is analyzed but also the Signed Bias which delivers information about the direction of the localization error. Furthermore, spatial release of masking (SRM) of speech (S) in noise (N). S and M were either collocated at 0° or S was situated at 0° and N at 90° at the HA side. In none of the bimodal listeners an improvement of SRM by the temporal adjustment of modalities was observed.

METHODS

9 adult bimodal listeners (age: 61.1 ± 6.9 years, min. 47 years, max. 71 years) participated in the study. Details are listed in Tables 1 and 2.
To delay sound signals with sufficient temporal resolution in bimodal listeners a programmable DL based on the Arduino Due microcontroller (µC) board with a built-in Atmel SAM3X8E ARM Cortex-M3 CPU was used. The DL corresponds exactly to the one used and described in detail in (Zirn et al., 2019). It provides the possibility to delay signals by integer multiples of $20.8\mu s$ while ensuring a very low minimum delay of only $50\ \mu s$. The delay is based on a ring buffer and affects all frequencies in the same way, it is thus an across-frequency delay.

All tests were administered in the same audiometric booth as for the earlier study published in 2019. 7 loudspeakers (type Genelec 8030C) were located at intervals of 30 degrees from -90 degrees (loudspeaker #1) to 90 degrees (loudspeaker #7) in a frontal semicircle 1 m in diameter in a horizontal plane at the subject's head level. The loudspeakers carried number plates corresponding to the numbers 1 to 7. The subjects were not allowed to search the presenting speaker by moving their heads during the stimulus presentations.
Instead of a single 200 ms noise burst, as used in the previous study, a stimulus consisting of multiple noise bursts was used in the current study. For this stimulus wide band Gaussian white noise (125 Hz to 20 kHz) was divided

into five 70 ms pulses with 30 ms pauses between the pulses. Further details are described in (Angermeier et al., 2021)

RESULTS

In Figure 2 results of the A-B-B-A test procedure are shown. At the first A test the delay line was active but adjusted

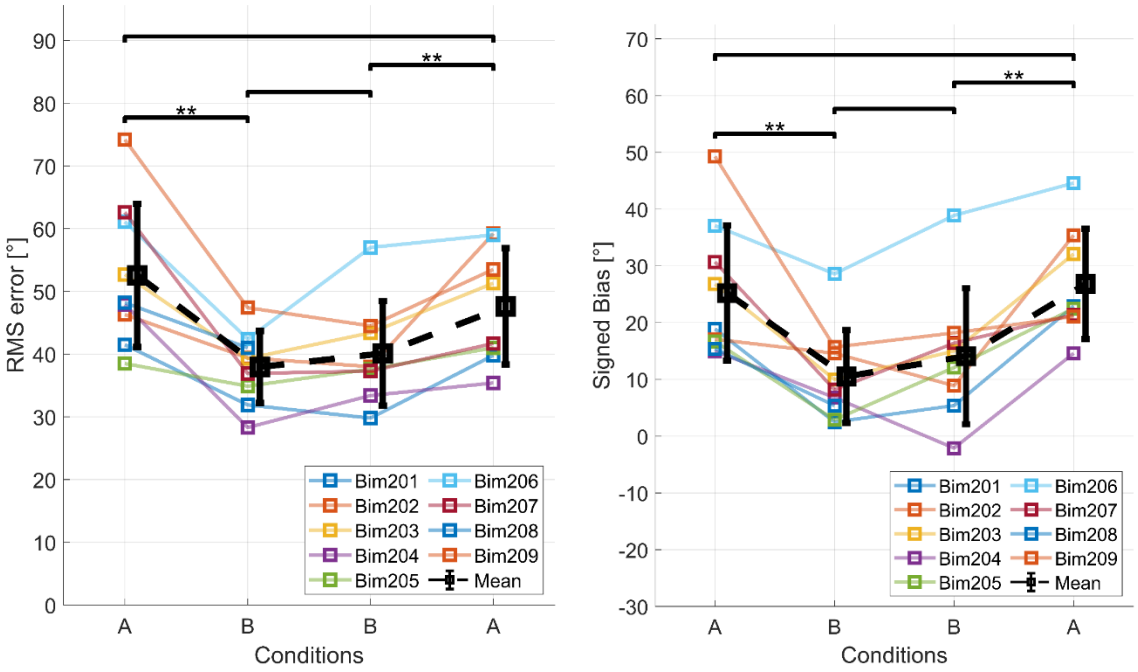


FIGURE 2: left: RMS errors in the sound localization test. Right: Signed Bias is shown normalized to the CI side (positive values mean localization towards the CI side). The colored markers and lines stand for individual results. The black markers and the dashed black line stand for mean ± standard deviation. Adapted from Angermeier et al. (2021).

to 0 ms delay. After the sound localization test the delay line was programmed with the measured processing delay of the individually worn HA at the contralateral ear. Interestingly, all participants tended to localize toward the CI side in the initial condition, as indicated by the signed bias. The first B test was then conducted acutely. A highly significant improvement for both, RMS error and Signed Bias occurred.

After this test, a familiarization period of 1 hour followed. The bimodal listeners could walk around to familiarize to the delayed CI stimulation. After this familiarization period a second sound localization test followed, the second B. No significant difference between the first and the second B test occurred neither for the RMS nor for the Signed Bias. Now, the CI delay was again programmed to 0 ms and the localization test was repeated acutely without familiarization. A highly significant increase of RMS error and the Signed Bias occurred.

Overall RMS error and Signed Bias improved by 14.7° in average by reducing the device delay mismatch.

Figure 3 shows results of the spatial release of masking of speech in noise tests. The results were acquired in a similar mode as the sound localization results (A-B-B-A paradigm) but in a later part of the experiment with another group of bimodal listeners (in this case 10 subjects).

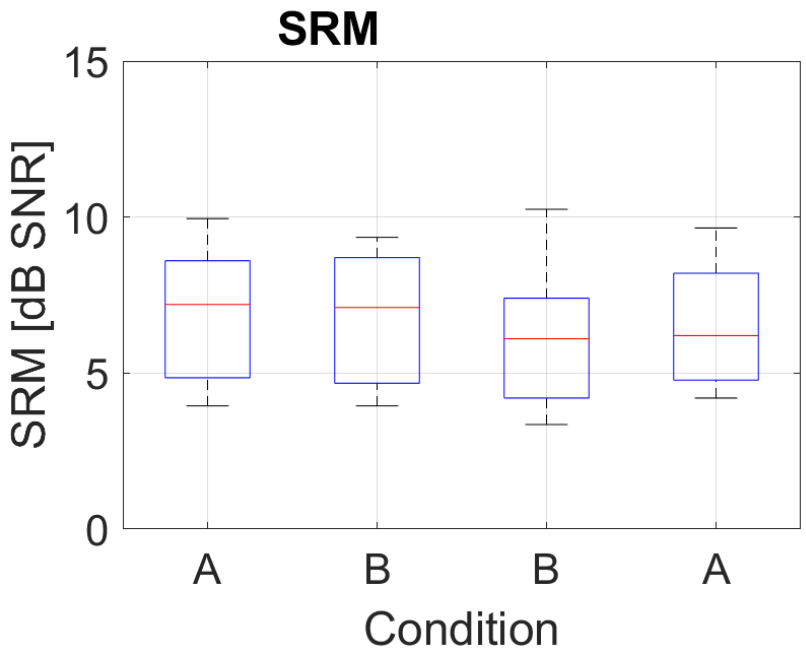


FIGURE 3: Spatial release of masking of speech in noise measured in 10 bimodal listeners.

No significant difference between the SRM across conditions occurred.

DISCUSSION

The results reveal that sound localization in bimodal listeners improves from a reduction of the device delay mismatch between CI and HA. With this form of temporal adjustment of modalities, spatial re-centering seems possible which can be seen in the Signed Bias in Figure 2 right where values close to 0° could be achieved in some subjects.
However, transmission of temporal information via the CI and thus binaural processing seems to be not sufficient for improved spatial unmasking of speech in noise. But regarding the SRM results from a positive side, reducing the device delay mismatch can be done without impairing speech reception.

CONCLUSION

Reducing the device delay mismatch is a valid procedure to increase sound localization accuracy in bimodal listeners. The

ACKNOWLEDGMENTS

We want to thank all bimodal listeners which participated in our studies. Furthermore, we want to thank MED-EL for funding.

REFERENCES

Angermeier, J., Hemmert, W., & Zirn, S. (2021). Sound Localization Bias and Error in Bimodal Listeners Improve Instantaneously When the Device Delay Mismatch Is Reduced. *Trends in Hearing*, *25*, 233121652110161. https://doi.org/10.1177/23312165211016165

Ching, T. Y. C., Incerti, P., & Hill, M. (2004). Binaural Benefits for Adults Who Use Hearing Aids and Cochlear Implants in Opposite Ears: *Ear and Hearing*, *25*(1), 9–21. https://doi.org/10.1097/01.AUD.0000111261.84611.C8

Ching, T. Y. C., Incerti, P., Hill, M., & van Wanrooy, E. (2006). An overview of binaural advantages for children and adults who use binaural/bimodal hearing devices. *Audiology & Neuro-Otology*, *11 Suppl 1*, 6–11. https://doi.org/10.1159/000095607

Dorman, M. F., Loiselle, L. H., Cook, S. J., Yost, W. A., & Gifford, R. H. (2016). Sound Source Localization by Normal-Hearing Listeners, Hearing-Impaired Listeners and Cochlear Implant Listeners. *Audiology and Neurotology*, *21*(3), 127–131. https://doi.org/10.1159/000444740

Farinetti, A., Roman, S., Mancini, J., Baumstarck-Barrau, K., Meller, R., Lavieille, J. P., & Triglia, J. M. (2015). Quality of life in bimodal hearing users (unilateral cochlear implants and contralateral hearing aids). *European Archives of Oto-Rhino-Laryngology*, *272*(11), 3209–3215. https://doi.org/10.1007/s00405-014-3377-8

Hoppe, U., Hocke, T., & Digeser, F. (2018). Bimodal benefit for cochlear implant listeners with different grades of hearing loss in the opposite ear. *Acta Oto-Laryngologica*, 1–9. https://doi.org/10.1080/00016489.2018.1444281

Sheffield, B. M., Schuchman, G., & Bernstein, J. G. W. (2017). Pre- and Postoperative Binaural Unmasking for Bimodal Cochlear Implant Listeners. *Ear and Hearing*, *38*(5), 554–567. https://doi.org/10.1097/AUD.0000000000000420

Zirn, S., Angermeier, J., Arndt, S., Aschendorff, A., & Wesarg, T. (2019). Reducing the Device Delay Mismatch Can Improve Sound Localization in Bimodal Cochlear Implant/Hearing-Aid Users. *Trends in Hearing*, *23*, 233121651984387. https://doi.org/10.1177/2331216519843876

Zirn, S., Arndt, S., Aschendorff, A., & Wesarg, T. (2015). Interaural stimulation timing in single sided deaf cochlear implant users. *Hearing Research*, *328*, 148–156. https://doi.org/10.1016/j.heares.2015.08.010

Two-Tone Suppression and Power Balance in a 2D Nonlinear Cochlear Model

Aleš Vetešník,[1, a)] Ondřej Klimeš,[2, b)] and Václav Vencovský[2, c)]

[1)] *Department of Nuclear Chemistry, Czech Technical University in Prague, Czech Republic.*
[2)] *Department of Radioelectronics, Czech Technical University in Prague, Czech Republic.*

[a)] *Corresponding author: ales.vetesnik@fjfi.cvut.cz*
[b)] *Electronic mail: klimeon5@fel.cvut.cz*
[c)] *Electronic mail: vaclav.vencovsky@gmail.com*

Abstract. Two-tone suppression (TTS) is the nonlinear phenomenon in which cochlear responses in the region most sensitive to a probe tone are reduced if a second tone (suppressor) of different frequency is simultaneously presented. TTS occurs due to saturation of mechanoelectrical transducers (MET) in the outer hair cells. The nonlinear dependence of MET channel conductance is the main cause of nonlinearity in cochlear amplification. Although the cochlear amplifier primarily acts locally, it influences distant fluid coupled parts of the basilar membrane (BM). The question therefore arises, how the fluid coupling contributes to the TTS. To answer this question, TTS was analysed using 2D nonlinear cochlear model in two cases: for a suppressor of higher frequency and for a suppressor of lower frequency than the probe tone. It was shown that the BM-BM hydrodynamical coupling redistributed the excess of the power generated by the OHC electromechanical feedback force. It was also shown that the short-range part of the BM-BM coupling plays important role in power balance along the BM. Moreover, it turned out from the model simulations that the suppression by a low-frequency suppressor causes a basal shift of the amplitude maximum and affects the phase of the BM stationary response. In contrast, the suppression by a high-frequency suppressor does not significantly affect the BM response phase. It flattens the amplitude of the BM stationary response, which leads to an apical shift of the amplitude maximum.

INTRODUCTION

Two-tone suppression (TTS) is the nonlinear phenomenon in which cochlear responses in the region most sensitive to a probe tone are reduced if a second tone (suppressor) of different frequency is simultaneously presented. TTS occurs due to saturation of mechanoelectrical transducers (MET) in the outer hair cells [1]. The nonlinear dependence of MET channel conductance is the main cause of nonlinearity in cochlear amplification. Although the cochlear amplifier primarily acts locally it influences distant fluid coupled parts of the basilar membrane (BM) [2]. The question therefore arises, how the fluid coupling contributes to the TTS. To answer this question, we used a nonlinear 2D cochlear model to simulate and analyse TTS. TTS is in the model simulated due to the change in the nonlinear force elicited by a suppressor.

We briefly introduce the used nonlinear cochlear model, i.e. a nonlinear integrodifferential BM motion equation together with a differential motion equation of the outer hair cell (OHC) stereocilia displacement, which is coupled to the BM motion equation via the OHC electromechanical feedback force. These motion equations are formulated in the time domain and in the frequency domain. The nonlinear force is calculated in the frequency domain as a deviation of the nonlinear feedback undamping force from the undamping force in the linear regime. In the end of the theoretical part of the current paper, a power balance equation for the BM motion is presented. The results then present TTS simulations and power balance analysis.

THEORETICAL ANALYSIS

The nonlinear 2D cochlea model

This study uses a nonlinear hydrodynamic cochlea model introduced by Nobili and Mammano [3]. The model's dynamics are described by combining two equations: one integro-differential for the transverse BM displacement, $\xi(x,t)$, and the other differential for the OHCs stereocilia displacement $\eta(x,t)$, both of which are functions of time t and the BM longitudinal coordinate x.

The BM is modeled as a continuous array of harmonic oscillators of mass, $m(x)$, damping, $h(x)$, and stiffness, $k(x)$, per unit BM length, longitudinally coupled by the shearing viscosity term, $\partial_x s(x)\partial_x$, and the hydrodynamic force

represented by the Green's functions $G(x,\bar{x})$ and $G_S(x)$, amplified by the OHC electromechanical feedback force term, $U(x,t)$ and externally driven by the stapes acceleration $\partial_t^2\sigma(t)$

$$m(x)\partial_t{}^2\xi(x,t)+h(x)\partial_t\xi(x,t)-[\partial_x s(x)\partial_x]\partial_t\xi(x,t)+k(x)\xi(x,t)+\int_0^L G(x,\bar{x})\partial_t{}^2\xi(\bar{x},t)d\bar{x} \\ =-G_S(x)\partial_t^2\sigma(t)-U(x,t), \quad (1)$$

where ∂_t and ∂_x denote partial derivates with respect to t and x, respectively. The Green's functions $G_S(x)$ and $G(x,\bar{x})$ represent the stapes-BM hydrodynamic coupling and the BM-BM hydrodynamics coupling, respectively. $G(x,\bar{x})$ can be decomposed into a long-range part, $G_0(x,\bar{x})$, representing the fluid coupling inherent in a 1D cochlea model, and a short-range part, $G_P(x,\bar{x})$, that accounts for the model extension to 2D.

The model assumes direct proportionality between $U(x,t)$ and the OHC mechanoeletrical transducer current elicited by $\eta(x,t)$; it is given by

$$U(x,t)=u(x)S[a\eta(x,t)], \quad (2)$$

where $u(x)$ is a suitable spatial function controlling the degree of amplification along the BM and $S[\cdot]$ is a sigmoidal function with parameters defined in [3].

Stereocilia radial deflection, $\eta(x,t)$, is calculated as the displacement of damped harmonic oscillator

$$\partial_t^2\eta(x,t)+\gamma_{\mathrm{TM}}(x)\partial_t\eta(x,t)+\omega_{\mathrm{TM}}^2(x)\eta(x,t)=-\partial_t^2\xi(x,t), \quad (3)$$

forced by the negative BM acceleration $\partial_t^2\xi(x,t)$; $\gamma_{\mathrm{TM}}(x)$ is the damping resulting from the viscosity of the subtectorial space and the TM viscoelasticity and $\omega_{\mathrm{TM}}(x)$ is the TM resonance frequency at x.

Modeling two-tone suppression

We now focus on a steady-state analysis of two-tone suppression. To this end, we transform Eqs. (1, 2, and 3) to the frequency domain. Similarly to the study [4], we assume that the radial frequencies $\omega_i=n_i\omega_0$ of the input tones are integer multiples of the fundamental frequency ω_0. Then, the stationary solutions of Eq. (1) can be approximated by the truncated Fourier series [4]:

$$\xi(x,t)\approx\sum_{n=-N}^{N}\hat{\xi}_n(x)e^{in\omega_0 t}, \quad (4)$$

where a circumflex denotes a phasor and the Fourier displacement amplitude $\hat{\xi}_n(x)$ satisfies the linear integro-differential equation which can be obtained by substituting Eq. (4) into Eq. (1). It reads

$$\left[-(n\omega_0)^2m(x)+in\omega_0\{h(x)-\partial_x s(x)\partial_x\}+k(x)+u(x)\hat{T}_n^{\mathrm{TM}}(x)\right]\hat{\xi}_n(x)-(n\omega_0)^2m(x)\int_0^L G(x,\bar{x})\partial_t{}^2\hat{\xi}_n(\bar{x},t)d\bar{x} \\ =(n\omega_0)^2G_S(x)\hat{\sigma}_n-\hat{U_n}^{\mathrm{NL}}(x), \quad (5)$$

where $\hat{U_n}^{\mathrm{NL}}$ is the nonlinear part of the Fourier amplitude of the OHC feedback force defined as

$$\hat{U}_n^{\mathrm{NL}}(x)\equiv u(x)\frac{1}{T_0}\int_0^{T_0}S\left[\sum_{n=-N}^{N}\hat{T}_n^{\mathrm{TM}}(x)\hat{\xi}_n(x)e^{in\omega_0 t}\right]e^{-in\omega_0 t}dt-u(x)\hat{T}_n^{\mathrm{TM}}(x)\hat{\xi}_n(x). \quad (6)$$

$\hat{T}_n^{\mathrm{TM}}(x)$ is transfer function derived from the Fourier transform of Eq. (3).

We denote the solution of Eq. (5), and corresponding $\hat{U}_n^{\mathrm{NL}}(x)$ and $\hat{\sigma}_n$, for a single tone of the angular frequency $n\omega_0$ and the input level L as $\hat{\xi}_n^0(x)$, $\hat{U}_n^{\mathrm{NL},0}(x)$ and $\hat{\sigma}_n^0$, respectively. The presence of the second tone, here referred as the suppressor, of the angular frequency $n_S\omega_0\neq n\omega_0$ influences in Eq. (5) primary the nonlinear force $\hat{U}_n^{\mathrm{NL,S}}(x)=$

$\hat{U}_n^{\mathrm{NL},0}(x)+\Delta\hat{U}_n^{\mathrm{NL}}(x)$, and subsequently the BM transverse displacement $\hat{\xi}_n^S(x)=\hat{\xi}_n^0(x)+\Delta\hat{\xi}_n(x)$, and finally the stapes displacement $\hat{\sigma}_n^S=\hat{\sigma}_n^0+\Delta\hat{\sigma}_n$. These changes satisfy Eq. (5), i.e.

$$\left[-(n\omega_0)^2 m(x)+i\omega_0\{h(x)-\partial_x s(x)\partial_x\}+k(x)+u(x)\hat{T}_n^{\mathrm{TM}}(x)\right]\Delta\hat{\xi}_n(x)-(n\omega_0)^2 m(x)\int_0^L G(x,\bar{x})\partial_t{}^2\Delta\hat{\xi}_n(\bar{x})d\bar{x}$$
$$=(n\omega_0)^2 G_S(x)\Delta\hat{\sigma}_n-\Delta\hat{U}_n{}^{\mathrm{NL}}(x). \quad (7)$$

Eq. (7) can be solved by means of the Green's function $\hat{H}_n(x,\bar{x})$ of the linear integro-differential operator of the left side of Eq. (5). The solution can be written as

$$\Delta\hat{\xi}_n(x)=(n\omega_0)^2\Delta\hat{\sigma}_n\int_0^L\hat{H}(x,\bar{x})G_S(\bar{x})d\bar{x}-\int_0^L\hat{H}(x,\bar{x})\Delta\hat{U}_n{}^{\mathrm{NL}}d\bar{x}. \quad (8)$$

Power balance equation for the BM motion

In the steady state and from the point of view of the global BM power balance, the power produced by the stapes-BM hydrodynamic force and the OHC electromechanical feedback force should balance power dissipated by the BM damping and the shearing viscosity force terms over a whole cycle. The BM-BM hydrodynamic force contributes only to the local power balance, it satisfies the equation

$$\int_0^L P_n^G(x)dx=\int_0^L P_n^{G_o}(x)dx+\int_0^L P_n^{G_P}(x)dx=0, \quad (9)$$

where $P_n^G(x)$ represents the power generated by the BM-BM hydrodynamic coupling force term, $P_n^{G_o}(x)$ is power generated by long range part of BM-BM fluid coupling and $P_n^{G_P}(x)$ power generated by short range part of BM-BM fluid coupling. We represent $P_n^G(x)$ in the local power balance equation together with the power generated by the stapes-BM hydrodynamic coupling force, $P_n^S(x)$. The local power balance equation can be obtained by multiplying Eq. (5) by the complex conjugate of the BM transverse velocity and then taking its real part [5]. For the analysis we express the local power balance equation in the form

$$P_n^{D_a}(x)=P_n^{G+S}(x)+P_n^{U^{NL}}(x), \quad (10)$$

where $P_n^D(x)$ includes both, the dissipated power by the BM damping and the shearing viscosity force term; and the power delivered by the linear part of OHC electromechanical feedback force. We included the linear part of feedback force into $P_n^D(x)$ because TTS is manifested in Eq. (5) only by a nonlinear part of $\hat{U}_n$. $P_n^{G+S}(x)$ denotes the power delivered by by the stapes-BM hydrodynamic coupling force and distributed by the BM-BM hydrodynamic force.

RESULTS

We present simulations for two cases of TTS: the first case is for the suppressor of higher angular frequency, i.e. $n_S > n$; the second case is for the suppressor of lower angular frequency, i.e. $n_S < n$. This choice is motivated by the assumption, that in the first case the suppressor acts only basally to the characteristic frequency site of the probe tone and in the second case the suppressor acts all along the probe tone response. In particular, we stimulate the model with input tones of frequencies $f_1 = 2.2$ kHz and $f_2 = 2.4$ kHz. In the first case, we present the f_1 tone at a constant level of $L_1 = 50$ dB SPL and change the level of the f_2 tone: $L_2 = 30, 40, 50, 60$ and 70 dB SPL. In addition, we calculate the model response to the f_1 tone for $L_1 = 50$ dB SPL without the f_2 tone. In the second case, we present the f_2 at a constant level of $L_2 = 50$ dB SPL and change the level of the f_1 tone: $L_1 = 30, 40, 50, 60$ and 70 dB SPL. We also obtain the model response to the f_2 tone presented at $L_2 = 50$ dB SPL without the f_1 tone.

Suppression by the suppressor of higher frequency

Figure 1 illustrates TTS due to a suppressor of higher frequency (f_2) than the probe tone frequency (f_1); $f_2/f_1 \approx 1.1$. Left hand side panels depict amplitudes (A-1) and phases (A-2) of the BM stationary response at f_1 frequency

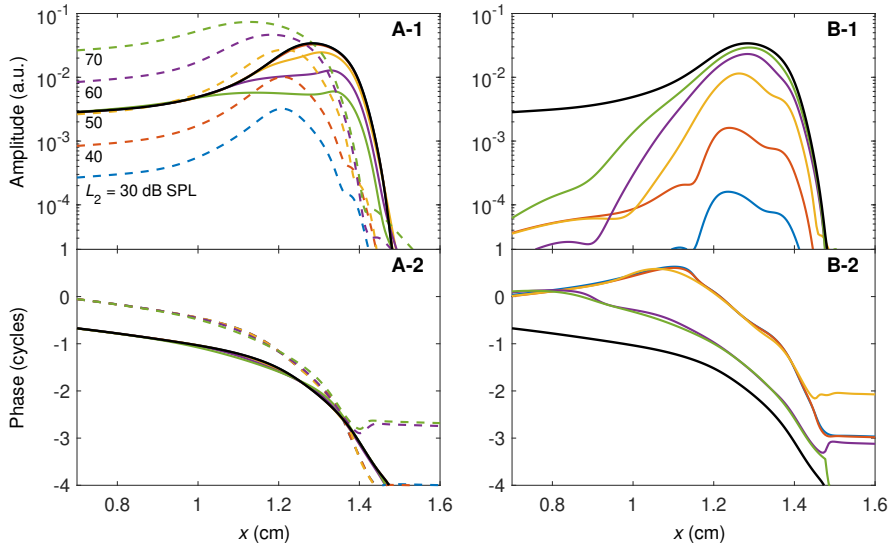


FIGURE 1. Suppression of the stationary displacement response of the model BM to $f_1 = 2.2\,\text{kHz}$, $L_1 = 50\,\text{dB SPL}$ (coloured solid lines) by the higher-frequency suppressor of $f_2 = 2.4\,\text{kHz}$, and $L_2 = 30, 40, 50, 60$, and $70\,\text{dB SPL}$ (dashed coloured lines). The responses for the same stimulus settings are depicted with the same colors. Panel A-1 depicts amplitudes and panel A-2 depicts phases. Amplitudes (B-1) and phases (B-2) of $\Delta\hat{\xi}_n$ at f_2. Black solid lines in panels A-1,2 and B-1,2 depict the BM response for the f_1 tone without the suppressor (f_2 tone).

(coloured solid lines), and at f_2 frequency (coloured dashed lines). The black thick line belongs to the BM stationary response to the single tone of frequency f_1, i.e. without the suppressor. It is obvious that as the f_2 tone level increases, the phase of f_1 traveling wave (TW) remains almost unchanged but the amplitude undergoes significant changes (is suppressed) for $L_2 > 40\,\text{dB SPL}$. For the suppressor levels of $L_2 = 30$ and $40\,\text{dB SPL}$, there is no significant suppression of f_1 TW. For $L_2 = 50\,\text{dB SPL}$, the f_2 TW (suppressor BM response depicted with the yellow dashed line) becomes comparable to that of the f_1 TW (yellow solid line). As a consequence, a significant TTS occurs. The amplitude decreases and flattens around the maximum. The amplitude flatting causes that the position of the TW maximum shifts apically as the suppressor level increases. Despite the amplitude flattening of the supressed TWs, the TW phase seems to be almost invariant (panel A-2).

Figure 1, B-1 and B-2 show changes in the f_1 TWs, $\Delta\hat{\xi}_n(x)$, at f_1 frequency given by Eq. (8). As well as in Fig. 1, the black solid line depicts the BM response to a single f_1 tone at $L_1 = 50\,\text{dB SPL}$. The amplitude of $\Delta\hat{\xi}_n(x)$ increases with increasing level. For $L_2 = 70\,\text{dB SPL}$ $\Delta\hat{\xi}_n(x)$ is almost equivalent (near the best frequency place) to the f_1 TW evoked without the suppressor (black solid line). The phases of profiles of $\Delta\hat{\xi}_n(x)$ firstly increase (phase leading profile) and then decrease (phase lagging profile) along the BM length x. The place x where the phase slope changes shifts basally as the suppressor level increases. This basal shift indicates that the suppressor affect broader portion of the f_1 TW, i.e., $\Delta\hat{U}_n{}^{\text{NL}}(x)$ broadens. This is because $\Delta\hat{U}_n{}^{\text{NL}}(x)$ represents in Eq. (8) source, and the phase slope of the Green's function $\hat{H}_n(x,\bar{x})$ is positive for $x < \bar{x}$ and negative for $x > \bar{x}$ [4].

Suppression by the suppressor of lower frequency

Figure 2 depicts TTS due to a suppressor of lower frequency (f_1) than the probe tone frequency f_2. Figure description is otherwise the same as for Fig. 1. Panels A-1 and A-2 show amplitudes and phases for the BM stationary responses at f_1 and f_2 frequencies, panels B-1 and B-2 the amplitudes and phases of $\Delta\hat{\xi}_n$ at f_2 frequency. The black thick line represents the BM stationary response to the single tone of f_2 frequency. As well as in Fig. 1, the suppressor levels $L_1 = 30$ and $40\,\text{dB SPL}$ do not cause significant TTS. Significant TTS occurs for $L_1 > 40\,\text{dB SPL}$ and is mostly visible in the TW amplitude which is suppressed and its peak is shifted towards the BM base (opposite direction than in Fig. 1). In contrast to Fig. 1, as the suppressor level increases the TW phase depicted in the panel A-2 has smaller phase slope than in the case without the suppressor.

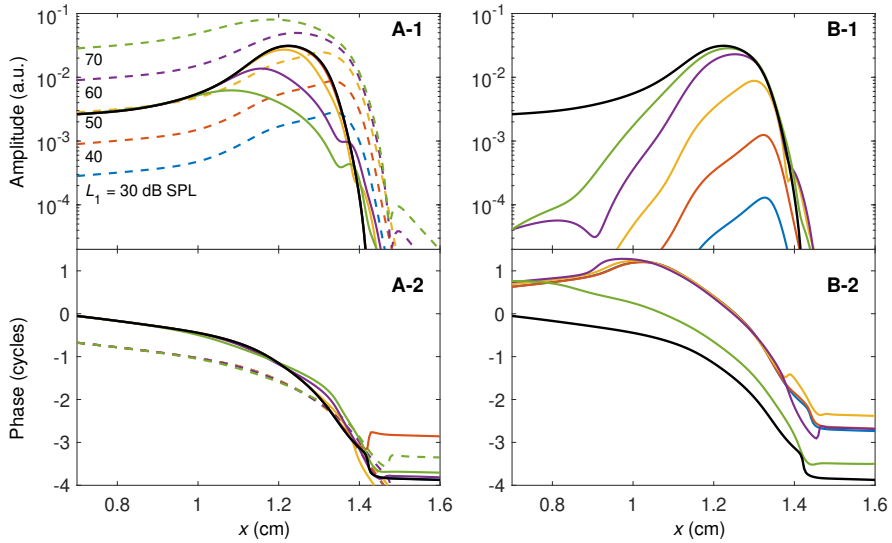


FIGURE 2. Suppression of the stationary displacement response of the model BM to $f_2 = 2.4\,\text{kHz}$, $L_2 = 50\,\text{dB}$ SPL (coloured solid lines) by the lower frequency suppressor of $f_1 = 2.2\,\text{kHz}$, and $L_1 = 30, 40, 50, 60$, and $70\,\text{dB}$ SPL (dashed coloured lines). The responses for the same stimulus settings are depicted with the same colors. Panel A-1 depicts amplitudes and panel A-2 depicts phases. Amplitudes (B-1) and phases (B-2) of $\Delta\hat{\xi}_n$ at f_2. Black solid lines in panels A-1,2 and B-1,2 depict the BM response for the f_2 tone without the suppressor (f_1 tone).

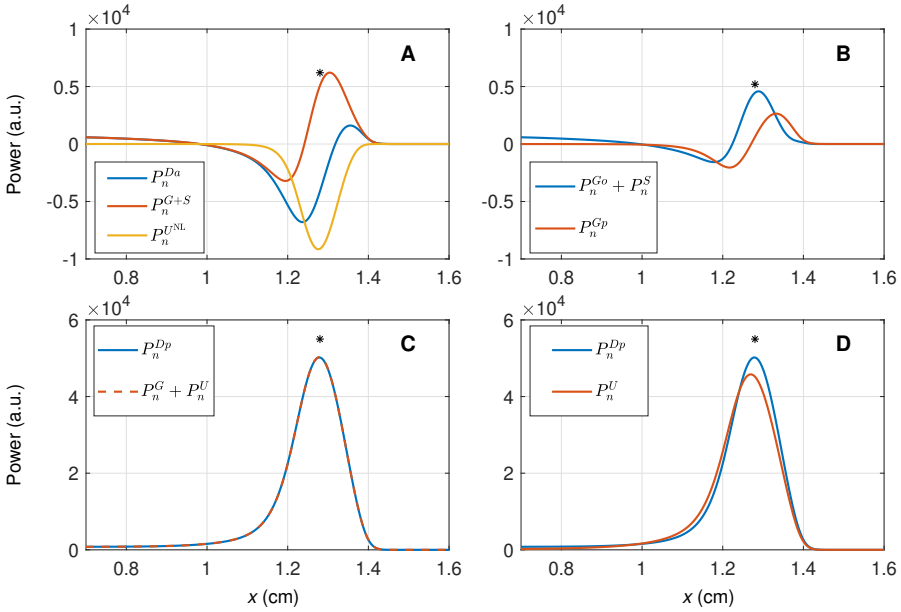


FIGURE 3. Power balance along the BM for a single $f_1 = 2.2\,\text{kHz}$ tone at $L_1 = 50\,\text{dB}$ SPL. The black asterisks in the all panels indicate x position of the maximum of the TW on the BM. A: P_n^{Da} denotes a sum of the power dissipated by the damping and the shearing viscosity force and the power delivered by the linear part of the undamping force $\hat{U}_n$, P_n^{G+S} is the power generated by the stapes-BM hydrodynamic coupling force and distributed along the BM by the BM-BM hydrodynamic force, $P_n^{U^{NL}}$ is the power yielded by the nonlinear part of $\hat{U}_n$, which is negative along the BM length. B: P_n^{Go} is the power yielded by the long-range part of BM-BM fluid coupling and P_n^{Gp} is the power yielded by the short-range part of BM-BM fluid coupling. C: P_n^{Dp} is the power dissipated by the damping and the shearing viscosity force, P_n^U is the total power provided by $\hat{U}_n$, and P_n^G is the power distributed along the BM by the BM-BM hydrodynamic force and stapes-BM hydrodynamic force. D: Compares P_n^U and P_n^{Dp}. The difference is due to the power yielded by BM-BM and stapes-BM hydrodynamic force.

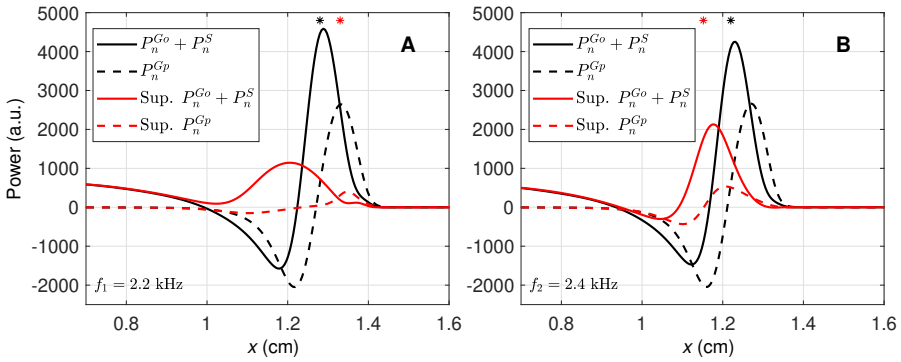


FIGURE 4. Suppression of power due to fluid coupling between stapes and BM, $G_S(x)$, and between each BM segment ($G(x,\bar{x})$. $P_n^{Go} + P_n^S$ represents the power from the stapes and the long range part of $G(x,\bar{x})$. P_n^{Gp} represents the power from the short range part of $G(x,\bar{x})$. A: High-frequency suppressor $f_2 = 2.4$ kHz, $L_2 = 60$ dB SPL, probe tone $f_1 = 2.2$ kHz, and $L_1 = 50$ dB SPL (red lines). Black lines depict powers for the probe tone without the suppressor. Red and black dots indicate position of maximum in f_1 TW amplitudes (Fig. 1). B: Low-frequency suppressor $f_1 = 2.2$ kHz, $L_1 = 60$ dB SPL, probe tone $f_2 = 2.4$ kHz, and $L_2 = 50$ dB SPL (red lines). Description is otherwise the same as for the panel A.

Power balance along the BM

Figure 3 illustrates a power balance along the BM for a single tone of $f_1 = 2.2$ kHz. Panel A compares the terms of the local power balance equation, Eq. (10). At the basal part of the BM ($x < 1$ cm), dissipation of P_n^{Da} due to the damping and the shearing viscosity force prevail (positive value). Then the power delivered by the linear part of $\hat{U}_n$ exceeds dissipation (negative value), and finally, dissipation again prevails. P_n^{G+S}, which is the power yielded by stapes-BM fluid coupling, $G_s(x)$, and BM-BM fluid coupling, $G(x,\bar{x})$, partially balances P_n^{Da}. In the BM parts where P_n^{Da} delivers the power, P_n^{G+S} absorbs excess of the power (negative value) and thus stabilizes the system. The power corresponding to the nonlinear part of $\hat{U}_n$, $P_n^{U^{\mathrm{NL}}}$, decreases the power provided by its linear part. Notice that the power is negative along the x axis. Panel B compares the power, $P_n^{G_0+S}$, produced by the stapes and distributed by the long-range part of the BM-BM fluid coupling and power, $P_n^{G_p}$, yielded by the short-range part of the BM-BM fluid coupling. The letter power is negative in the BM basal to the TW maximum (absorbs power) and positive apically to the TW maximum (delivers power). Panel C verifies a local stability of the model. The power dissipated by the damping and the shearing viscosity force, P_n^{Dp}, is compensated by the power provided by the stapes and distributed by the BM-BM fluid coupling, P_n^{G+S}, and by the power produced by the OHC electromechanical feedback force, P_n^U. Panel D compares P_n^{Dp} and P_n^U. The major part of dissipated energy is compensated by $P_n^{U^{\mathrm{NL}}}$.

Figure 4 shows the suppression of power delivered by the stapes and distributed by the BM-BM fluid coupling. The figure compares the power without the suppressor (black lines) and with suppressor (red lines). In addition, the figure decomposes the long-range and short-range parts of the BM-BM fluid coupling. Notice that high-frequency suppressor substantially influences power distribution due to fluid coupling (panel A). $P_n^{Go} + P_n^S$ is everywhere positive and P_n^{Gp} has distorted shape with flattened negative part. Low-frequency suppressor (panel B) decreases the fluid coupling power and shifts their minimum and maximum basally.

CONCLUSION

TTS was analysed using 2D nonlinear cochlear model in two cases, for suppressor of higher frequency and for suppressor of lower frequency. It was shown that the suppression by high-frequency suppressor leads to flatting of the amplitude of the BM stationary response, apical shift of the amplitude maximum and does not strongly influence the BM response phase. The suppression by lower frequency suppressor does not flatten the amplitude profile, causes basal shift of the amplitude maximum and does influence the phase of the BM stationary response.

In addition, it was shown that the BM-BM hydrodynamical coupling redistributed the excess of the power generated by the OHC electromechanical feedback force. It was also shown that the short range part of the BM-BM coupling plays an important role in the power balance along the BM. This term is missing in 1D cochlear models. Although

it can be formally represented in a 1D model by means of a transfer function, this approach may lead to erroneous conclusions, especially in the case of time domain simulations. Such transfer function depends on level and frequency of suppressor.

ACKNOWLEDGMENTS

Supported by an internal grant at the Czech Technical University in Prague SGS20/180/OHK3/3T/13, by the European Regional Development Fund-Project "Center for Advanced Applied Science" (Grant No. CZ.02.1.01/0.0/0.0/16_019/0000778). Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum provided under the programme "Projects of Large Research, Development, and Innovations Infrastructures" (CESNET LM2015042), is greatly appreciated.

REFERENCES

1. C. Geisler, G. K. Yates, R. B. Patuzzi, and B. M. Johnstone, "Saturation of outer hair cell receptor currents causes two-tone suppression," Hearing Research **44**, 241–256 (1990).
2. R. Sisto, D. Belardinelli, and A. Moleti, "Fluid focusing and viscosity allow high gain and stability of the cochlear response," J. Acosut. Soc. Am. **150**, 4283–4296 (2021).
3. R. Nobili and F. Mammano, "Biophysics of the cochlea. II: Stationary nonlinear phenomenology," J. Acoust. Soc. Am. **99**, 2244–2255 (1996).
4. A. Vetešník and A. W. Gummer, "Transmission of cochlear distortion products as slow waves: A comparison of experimental and model data," J. Acoust. Soc. Am. **131**, 3914–3934 (2012).
5. A. Vetešník, R. Nobili, and A. W. Gummer, "Physical mechanisms of oae generation and propagation: The hydrodynamic approach," in *Auditory Mechanisms: Processes and Models* (World Scientific Publishing, Singapore, 2006) pp. 315–321, A. L. Nuttall, T. Ren, P. Gillespie, K. Grosh, and E. de Boer ed.

Investigating the Effect of Change in Cochlear Micromechanics and Activity Levels on Stimulus Frequency Otoacoustic Emissions Phase-gradient Delay

Yiwei Xia,[a] George Samaras,[b] and Julien Meaud[c]

George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, 771 Ferst Drive, Atlanta, Georgia 30332, USA).

[a]*Electronic mail: yxia63@gatech.edu*
[b]*Electronic mail: gsamaras3@gatech.edu*
[c]*Corresponding author: julien.meaud@me.gatech.edu*

Abstract. Stimulus frequency otoacoustic emissions (SFOAEs), which are sounds emitted by the cochlea at the frequency of the stimulus, have been used as a noninvasive measure of cochlear function. The gradient of the SFOAE phases characterizes the latency of emission and is associated with the frequency selectivity and sharpness of tuning of the mammalian cochlea. However, whether the phase-gradient delay of SFOAE can be used as a indicator of cochlear tuning and sensitivity reliably when the properties of the cochlea change remains unclear. The objective of this study is to address this question by varying cochlear model activity, tectorial membrane (TM) properties and organ of Corti (OoC) micromechanical properties to change cochlear tuning. In this work, a three-dimensional gerbil cochlear model that couples mechanical, electrical and acoustic domains with cochlear roughness has been used. The roughness is implemented on outer hair cells (OHCs) force acting on the basilar membrane (BM). Parameters that control the activity levels, TM longitudinal coupling and OoC impedance are varied. The results show that changes in sharpness of tuning due to reduction in cochlear activity and TM longitudinal coupling can be detected by using SFOAE phase-gradient delay. However, changes in cochlear tuning due to changes in OoC impedance are not necessarily reflected by corresponding changes in SFOAE phase-gradient delay.

INTRODUCTION

Otoacoustic emissions (OAEs) are sounds generated inside of cochlea, which are caused by the active feedback by outer hair cells (OHCs) and can be measured at the ear canal (EC). OAEs have been used as a simple, efficient and noninvasive measure of cochlear function in both research and clinical practice [1]. Based on coherent reflection theory, a linear reflection mechanism due to impedance perturbations gives rise to reflection OAEs [2, 3]. Rapidly rotating phase has been found to be one of the key characteristics of reflection OAEs [3]. This study focuses on one type of reflection OAEs called stimulus frequency OAEs (SFOAEs), which are sounds emitted by the cochlea at the frequency of the external stimulus. The gradient of the SFOAE phase characterizes the latency of emission and is associated with frequency selectivity and sharpness of tuning of the mammalian cochlea. The phase-gradient delay of SFOAEs has been proposed to estimate the quality factor of the basilar membrane (BM) response [4, 5]. A more recent study has found the tuning ratio (ratio of tuning sharpness to SFOAE phase-gradient delay expressed in periods) to be approximately invariant among cat, guinea pig, and chinchilla, opening the possibility of determining cochlear tuning from SFOAE phase-gradient delay [6]. However, whether the phase-gradient delay of SFOAE can be used as an indicator of cochlear tuning and sensitivity reliably when the properties of the cochlea change remains unclear. The objective of this study is to address this question by varying cochlear model activity levels, tectorial membrane (TM) properties and organ of corti (OoC) impedance that can affect the cochlear tuning and sensitivity. Varying TM properties in the case study is inspired by studies that have shown that changing tectorial membrane (TM) longitudinal coupling affects cochlear tuning [7, 8, 9], spontaneous OAEs (SOAEs) [10] and SFOAEs [11].

METHOD

In this work, a three-dimensional gerbil cochlear model that couples mechanical, electrical and acoustic domains with cochlear roughness is used, which has been described in detail in previous works [12, 13]. The parameters of the baseline model can be found in [12]. The smooth version of this cochlear model has been calibrated based on in vivo recent experiments, with both mechanical and electrical data. In order to simulate reflection OAEs, cochlear roughness on the OHCs electromechanical coupling coefficient (which relates the electromotile force applied on the

BM to the transmembrane potential) is introduced as in our previous work [13]. ΔR is the standard deviation of the perturbations from the smooth case that quantifies the roughness. A random number seed (RNS) is used in the model to initialize the random number generator, which enables the possibility of simulating a family of various roughness profiles using different RNS values, which mimics the measurement of OAEs in multiple cochleae. $\Delta R = 0.1\%$ has been used throughout this work for the rough models. As in Ref. [13], calculation of the SFOAE is based on the reflection component in EC pressure, which is the vector difference between the steady-state pure tone response of the model with roughness and without roughness (smooth model).

To get the phase-gradient delay of SFOAE at 20 kHz, two methods are implemented. The first method (phase-fitting method) applies a quadratic regression to the phase of SFOAE between 10 kHz to 30 kHz to extract the trend line of the phase. The slope of the quadratic fit at stimulus frequency $f = 20$ kHz is used to determine the SFOAE phase-gradient delay (τ_{SFOAE}), and the equivalent number of stimulus periods, $N_{\mathrm{SFOAE}}(f) = \tau_{\mathrm{SFOAE}}(f) \cdot f$. The second method implements the peak-picking algorithm [14], which considers only those values of $N_{\mathrm{SFOAE}}(f)$ that occur near frequencies corresponding to peaks in SFOAE level. Here the peak-picking selection includes the peak itself and two points on either side of the peak. The peak-picked phase-gradient delays of SFOAE are then smoothed with a LOESS fit to find the value at $f = 20$ kHz. The theory of reflection emissions implies that τ_{SFOAE} is around twice the phase-gradient delay of BM response (τ_{BM}) evaluated at the location where CF is equal to the stimulus frequency [15]. To facilitate the comparison between BM and SFOAE phase-gradient delay, normalized values $\frac{1}{2}N_{\mathrm{SFOAE}}$ and $N_{\mathrm{BM}} = \tau_{\mathrm{BM}} \cdot \mathrm{CF}$ are analyzed in the RESULTS section. To compare the cochlear tuning sharpness, parameter-free measure $Q_{\mathrm{ERB}}(f_{\mathrm{CF}}) = f_{\mathrm{CF}}/\mathrm{ERB}(f_{\mathrm{CF}})$ is computed [16]. Here, $\mathrm{ERB}(f_{\mathrm{CF}})$ is the equivalent rectangular bandwidth at the local characteristic frequency, f_{CF}.

Three model parameters that can change cochlear tuning have been varied. The first parameter is activity level (α_{ActLev}), which is defined as a factor constant that scales the saturating mechano-electrical transduction (MET) conductance G_{hb}^{max} from the baseline value used in [12]. Activity level of 1 corresponds to a fully active model, and 0 corresponds to a passive model. In this study, the activity level is varied between 0.4 to 1. The second parameter is TM longitudinal coupling ratio (R_{TMLC}) that scales the value of longitudinal stiffness and viscosity of the TM [9] from its baseline value [12]. The range of interest for R_{TMLC} is between 0.5 and 1 here, where 1 corresponds to the normal cochlea and a value of 0.5 is representative of the TM coupling observed in mutant mice with reduced longitudinal coupling [7]. The third model parameter that has been varied is OoC impedance ratio (R_{OOC}). It is a factor constant that scales the stiffness of OHCs, reticular lamina (RL) and hair bundles (HBs), as well as the bending stiffness, mass and damping of TM from the corresponding baseline values [12]. $R_{\mathrm{OOC}} = 1$ corresponds to the baseline model which was calibrated based on the vivo data. In addition, the state-space formulation of the model is ran in frequency domain and the linear stability of each case considered has been checked [17, 18].

RESULTS

Change in BM responses and SFOAE due to reduction in cochlear activity level

Figure 1 shows the effect of change in activity levels on BM responses to pure tones and SFOAE. Decreasing the activity levels broadens the tuning of the responses of the BM velocity and decreases their sensitivity (panel a). The sharpness of tuning, indicated by Q_{ERB}, decreases as the activity level decreases (panel f). The slope of the phase of V_{BM} (panel d) decreases near the best frequency (BF) when the activity levels decreases. The model predicts the fine structure of SFOAE, including the relatively broad peaks separated by deep notches in the amplitude plot (panel b). The SFOAE level tends to decrease as the activity level decreases, and the deep notches in the fine structure shift from high frequencies to lower frequencies. In parallel with the changes in SFOAE amplitude and fine structure, the slope of the SFOAE phase becomes shallower at lower activity levels, indicating a shortening of the phase-gradient delay. Panel c compares $\frac{1}{2}N_{\mathrm{SFOAE}}$ obtained by two different methods to N_{BM} for various activity levels. Variance of the normalized SFOAE phase-gradient delay exists between different RNS values; however, at the population level, by taking the average of those individual values $\left(\frac{1}{2}\bar{N}_{\mathrm{SFOAE}}\right)$, N_{SFOAE} remains approximately equal to twice N_{BM} as the activity level is reduced. Comparing panels c and f shows that the changes in $\bar{N}_{\mathrm{SFOAE}}$ and Q_{ERB} are correlated.

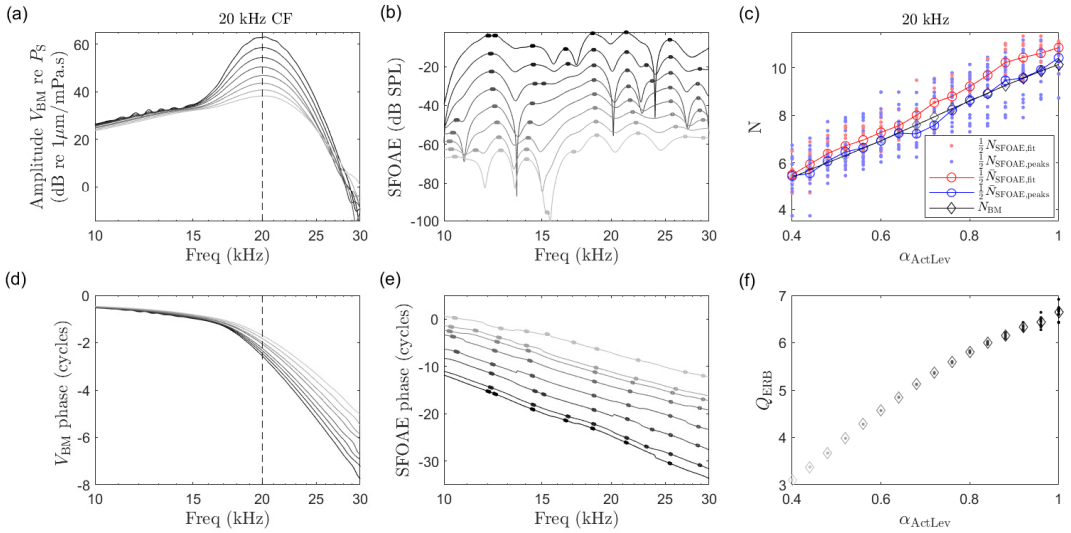


FIGURE 1: Effect of change in activity levels on BM responses and SFOAE. Amplitude (a) and phase (d) of BM velocity response to a pure tone for various reduced activity levels (from 100% to 40% activity level with respect to baseline). Amplitude (b) and phase (e) of SFOAE with reduced activity levels. Dots mark data points used in peak-picking algorithm to determine the phase-gradient delay values. All intracochlear responses are at the location tuned to 20 kHz, i.e., the 20 kHz best place (BP). The roughness parameter is set to be $\Delta R = 0.1\%$. Only RNS$= 0$ case is shown here among 10 different random seed numbers. Phases are offset by integer numbers of cycles to distinguish them in panel e. (c) Normalized phase-gradient delay with respect to activity levels, α_{ActLev} at 20 kHz. Dots indicate normalized SFOAE phase-gradient delays obtained by phase-fitting method (red) and peak-picking method (blue) for RNS$= 0-9$; circles represents the mean values of normalized SFOAE phase-gradient delays obtained by phase-fitting method (red) and peak-picking method (blue) among 10 different random seed numbers; black diamonds show the normalized phase-gradient delays of BM response. (f) Tuning sharpness of BM response as a function of activity levels at 20 kHz. Dots represent the individual values of $Q_{\text{ERB}}(20\text{ kHz})$ for RNS$= 0-9$; diamond markers represent the mean values of $Q_{\text{ERB}}(20\text{ kHz})$ among 10 different seed numbers. The color gradient is corresponding to change in activity levels: from black to light gray reflects the baseline case ($\alpha_{\text{ActLev}} = 1$) to lowest reduced activity level case ($\alpha_{\text{ActLev}} = 0.4$) accordingly.

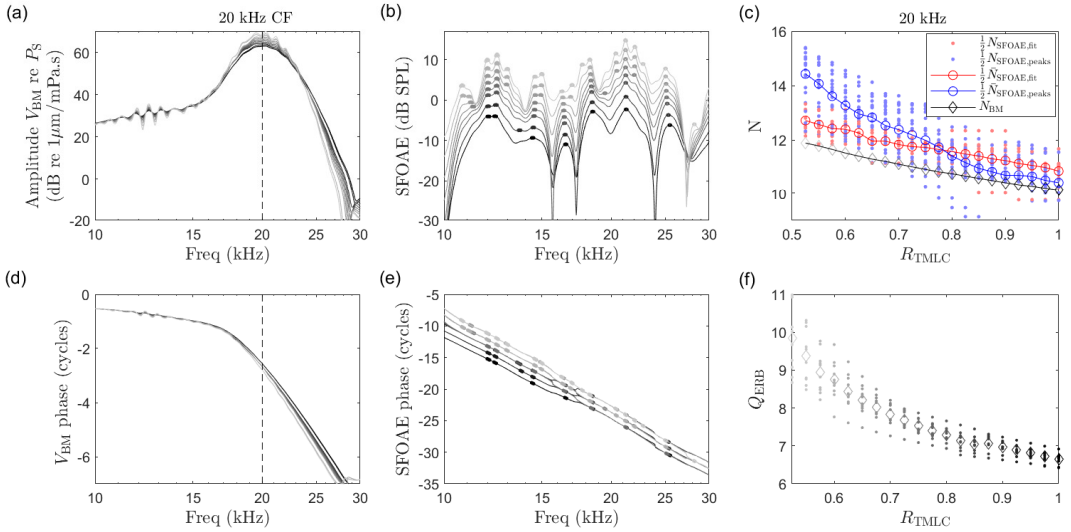


FIGURE 2: Same as Figure 1 but for the effect of change in TM longitudinal coupling.

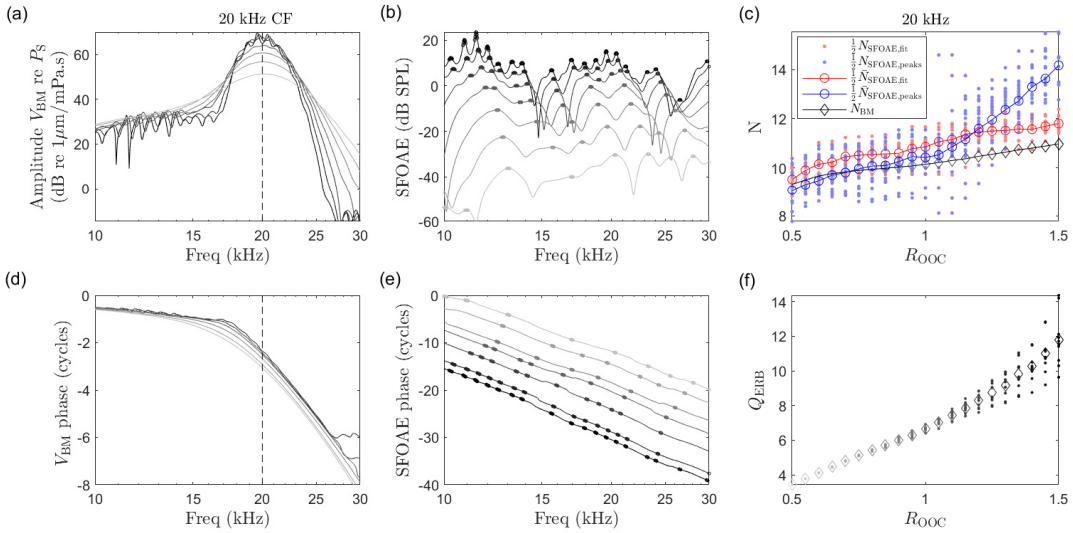


FIGURE 3: Same as Figure 1 but for the effect of change in OoC impedance.

Change in BM responses and SFOAE due to reduction in TM longitudinal coupling

Figure 2 shows the effect of changes in TM longitudinal coupling on BM responses and SFOAE. Response of the BM velocity (panel a) at the CF increases as the R_{TMLC} decreases; the bandwidth of the responses becomes slightly narrower as discussed in [9]. As a consequence, the sharpness of the tuning increases as the TM longitudinal coupling decreases (panel f). The slope of the phase of V_{BM} (panel d) steepens slightly near the BF when R_{TMLC} decreases. The SFOAE level grows as R_{TMLC} decreases, and more peaks are formed in the fine structure and the deep notches become less pronounced. The phase-gradient delay increases since the slope of the SFOAE phases becomes slightly steeper when TM longitudinal coupling is weaker. As shown in panel c, $\frac{1}{2}\bar{N}_{\mathrm{SFOAE}}$ decreases as R_{TMLC} increases. Even though the values of $\frac{1}{2}\bar{N}_{\mathrm{SFOAE,fit}}$ obtained from phase-fitting method are biased (higher than N_{BM}), the trend of $\frac{1}{2}\bar{N}_{\mathrm{SFOAE,fit}}$ is more consistent with the change in N_{BM} as compared to the trend of $\frac{1}{2}\bar{N}_{\mathrm{SFOAE,peaks}}$ when TM longitudinal coupling varies. The peak-picking method correctly determines the phase-gradient delay value at specific frequency (i.e., 20 kHz) when R_{TMLC} is close to 1; however, it starts to overestimate the phase gradient delay when the SFOAE responses have more peaks caused by multiple internal reflections [14] as R_{TMLC} decreases.

Change in BM responses and SFOAE due to variation in OoC impedance

The effect of change in OoC impedance on BM responses and SFOAE is demonstrated in Figure 3. Panel a shows that decreasing the OoC impedance reduces the peak values and expands the bandwidth of BM velocity responses. Subsequently, the tuning sharpness drops as the OoC impedance decreases as shown in panel f. There exists a shift in the BM velocity phases as OoC impedance changes; however, the slope of the phase barely changes as the traveling wave approaching BF when the OoC impedance varies. The SFOAE level tends to decrease as the OoC impedance decreases, and the fine structure becomes less prominent with less pronounced peaks. The SFOAE phase-gradient delay changes slightly since the slope of the SFOAE phases hardly decreases as the OoC impedance decreases. Panel c compares the change in $\frac{1}{2}\bar{N}_{\mathrm{SFOAE}}$ obtained by two different methods and N_{BM} as R_{OOC} varies. N_{BM} grows as R_{OOC} increases with almost a constant rate; on the other hand, the relationship between $\frac{1}{2}\bar{N}_{\mathrm{SFOAE,fit}}$ obtained from phase-fitting method and R_{OOC} is monotonic with a varying rate. The slope of $\frac{1}{2}\bar{N}_{\mathrm{SFOAE,fit}}$ turns to almost zero as the R_{OOC} varies around 0.75 and 1.25. In terms of phase-gradient delays obtained by peak-picking method, $\frac{1}{2}\bar{N}_{\mathrm{SFOAE,peaks}}$ follows the trend of N_{BM} when $R_{\mathrm{OOC}} < 1$ (when SFOAE has less prominent fine structures with less pronounced peaks), being unable to reflect the change in OoC impedance well when the impedance values are smaller than the baseline value. When the SFOAE responses include more prominent fine structures with more peaks as R_{OOC} increases

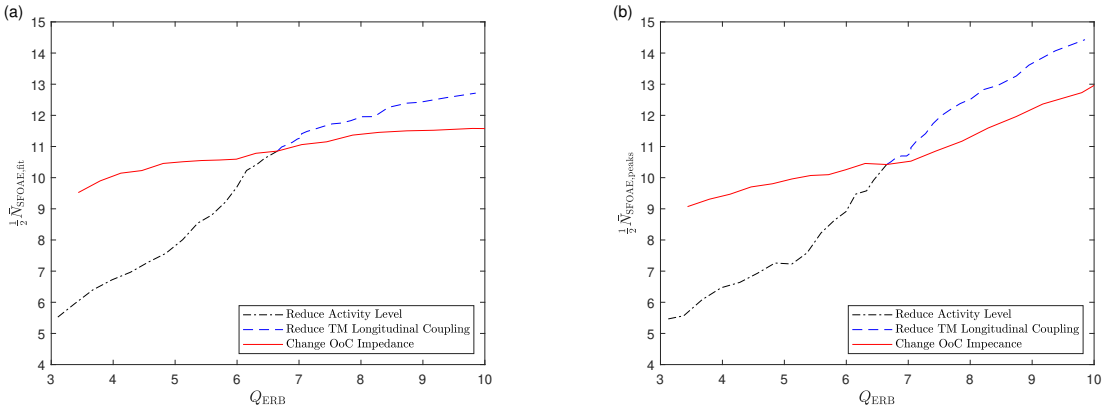


FIGURE 4: Comparison of normalized phase-gradient delay of SFOAE as a function of quality factor of the response between three considered scenarios based on (a) phase-fitting method, (b) peak-picking method.

($R_{\mathrm{OOC}} > 1$), $\frac{1}{2}\bar{N}_{\mathrm{SFOAE,peaks}}$ deviates from N_{BM}, not reflecting the BM phase-gradient delay correctly. Since the change in $\frac{1}{2}\bar{N}_{\mathrm{SFOAE}}$ can be hard to notice at a certain range where the R_{OOC} varies, the phase-gradient delay of SFOAE cannot indicate the change in OoC impedance as well as the BM phase-gradient delay.

SFOAE phase-gradient delay can detect changes in tuning sharpness due to reduction in cochlear activity level, decrease in TM longitudinal coupling, but not necessarily variation in OoC impedance

Figure 4 compares the normalized phase-gradient delay of SFOAE with respect to Q_{ERB} of the BM responses based on phase-fitting method (a) and peak-picking method (b) to evaluate whether SFOAE phase-gradient delay can detect change of cochlear tuning. As shown in both panel a and b, there exists a strong positive correlation between $\frac{1}{2}\bar{N}_{\mathrm{SFOAE}}$ and Q_{ERB} for the sharpness change due to reduction in cochlear activity and TM longitudinal coupling. However, the positive correlation between $\frac{1}{2}\bar{N}_{\mathrm{SFOAE}}$ and Q_{ERB} for the varied OoC impedance case is much weaker compared to the previous two cases ($\frac{1}{2}\bar{N}_{\mathrm{SFOAE}}$ varies only by around 20% when Q_{ERB} varies nearly by 300% in Fig. 4a for example), indicated by the shallower slope among all Q_{ERB} values shown in panel a, and when $Q_{\mathrm{ERB}} < 6.5$ in panel b. Both methods confirm that changes in tuning sharpness due to reduction in cochlear activity and decrease in TM longitudinal coupling can be detected by measuring the changes in SFOAE phase-gradient delays. However, changes in the tuning sharpness due to variation in OoC impedance are not necessarily reflected by changes in SFOAE phase-gradient delays, especially when the OoC impedance decreases from baseline value.

DISCUSSION

This work utilizes a physiologically-based model of gerbil cochlea to investigate the influence of changes in OoC micromechanics and activity levels on SFOAE phase-gradient delay. Two different methods are used to obtain the phase-gradient delay of SFOAE at specific stimulus frequency. The phase-fitting method provides better trend of the phase-gradient delay, including models that have sharp peaks in SFOAE. However, this does not compute the physical meaningful phase-gradient delay, i.e., the red curve is significantly above the black curve in panel c in Figs.$1-3$. The peak-picking method is better at getting the correct phase-gradient delay when the SFOAE peaks are not too sharp.

Both methods lead to a similar conclusion. At the individual levels, SFOAE phase-gradient delay cannot be used to indicate the cochlear tuning and sensitivity since there is no clear correlation between N_{SFOAE} and Q_{ERB} of the BM response for different RNS values. At the population level, the results show that changes in cochlear tuning due to variations in cochlear activity and TM longitudinal coupling can be detected by using average SFOAE phase-gradient delay among various RNS values. However, changes in cochlear tuning as a result of changes in OoC impedance are

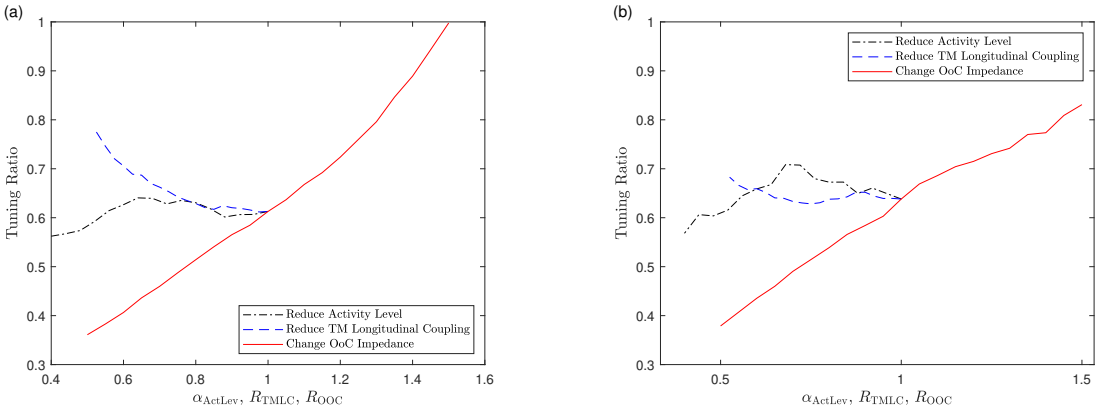


FIGURE 5: Comparison of tuning ratio with respect to variables that control the three considered parameters according to (a) phase-fitting method, (b) peak-picking method.

not necessarily reflected by corresponding changes in mean SFOAE phase-gradient delay. When R_{OOC} is varied, the shape of BM responses varies, such that it affects the relation between Q_{ERB} and N_{BM}, as well as the relation between Q_{ERB} and N_{SFOAE} since N_{SFOAE} and N_{BM} are correlated. This implies that changes in cochlear micromechanics (such as varying OoC impedance) can affect the tuning ratio [6], which is the ratio of tuning sharpness to SFOAE phase-gradient delay in periods (i.e., $Q_{\mathrm{ERB}}/\frac{1}{2}\bar{N}_{\mathrm{SFOAE}}$). Figure 5 compares the tuning ratio with respect to variables that control activity level, TM longitudinal coupling and OoC impedance respectively using (a) phase-fitting method and (b) peak-picking method. The possibility that the tuning ratio may change significantly needs to be kept in mind when SFOAE phase-gradient delays are used to estimate cochlear tuning in multiple species, or in animals with altered properties (such as transgenic mice with altered TM expressions [11]).

An alternative method to estimate cochlear tuning using SFOAE is by measuring the quality factor of SFOAE suppression tuning curves (SFOAE-STC) [19]. Future work might explore the influence of cochlear micromechanics and activity levels on SFOAE-STC.

ACKNOWLEDGMENTS

This research is funded by NIH Grants R01 DC016114.

REFERENCES

1. D. T. Kemp, "Otoacoustic emissions, their origin in cochlear function, and use," British medical bulletin **63**, 223–241 (2002).
2. G. Zweig and C. A. Shera, "The origin of periodicity in the spectrum of evoked otoacoustic emissions," The Journal of the Acoustical Society of America **98**, 2018–2047 (1995).
3. C. A. Shera and J. J. Guinan Jr, "Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian oaes," The Journal of the Acoustical Society of America **105**, 782–798 (1999).
4. C. A. Shera, J. J. Guinan, and A. J. Oxenham, "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements," Proceedings of the National Academy of Sciences **99**, 3318–3323 (2002).
5. C. A. Shera, "Mammalian spontaneous otoacoustic emissions are amplitude-stabilized cochlear standing waves," The Journal of the Acoustical Society of America **114**, 244–262 (2003).
6. C. A. Shera, J. J. Guinan, and A. J. Oxenham, "Otoacoustic estimation of cochlear tuning: validation in the chinchilla," Journal of the Association for Research in Otolaryngology **11**, 343–365 (2010).
7. I. J. Russell, P. K. Legan, V. A. Lukashkina, A. N. Lukashkin, R. J. Goodyear, and G. P. Richardson, "Sharpened cochlear tuning in a mouse with a genetically modified tectorial membrane," Nature neuroscience **10**, 215–223 (2007).
8. R. Ghaffari, A. J. Aranyosi, and D. M. Freeman, "Longitudinally propagating traveling waves of the mammalian tectorial membrane," Proceedings of the National Academy of Sciences **104**, 16510–16515 (2007).
9. J. Meaud and K. Grosh, "The effect of tectorial membrane and basilar membrane longitudinal coupling in cochlear mechanics," The Journal of the Acoustical Society of America **127**, 1411–1421 (2010).
10. T. Bowling, C. Lemons, and J. Meaud, "Reducing tectorial membrane viscoelasticity enhances spontaneous otoacoustic emissions and compromises the detection of low level sound," Scientific reports **9**, 1–11 (2019).

11. M. A. Cheatham, "Comparing spontaneous and stimulus frequency otoacoustic emissions in mice with tectorial membrane defects," Hearing research **400**, 108143 (2021).
12. T. Bowling, H. Wen, S. W. Meenderink, W. Dong, and J. Meaud, "Intracochlear distortion products are broadly generated by outer hair cells but their contributions to otoacoustic emissions are spatially restricted," Scientific Reports **11**, 1–14 (2021).
13. H. Wen and J. Meaud, "Link between stimulus otoacoustic emissions fine structure peaks and standing wave resonances in a cochlear model," The Journal of the Acoustical Society of America **151**, 1875–1894 (2022).
14. C. A. Shera and C. Bergevin, "Obtaining reliable phase-gradient delays from otoacoustic emission data," The Journal of the Acoustical Society of America **132**, 927–943 (2012).
15. C. A. Shera and J. J. Guinan, "Mechanisms of mammalian otoacoustic emission," in *Active processes and otoacoustic emissions in hearing* (Springer, 2008) pp. 305–342.
16. C. A. Shera and J. J. Guinan Jr, "Stimulus-frequency-emission group delay: A test of coherent reflection filtering and a window on cochlear tuning," The Journal of the Acoustical Society of America **113**, 2762–2772 (2003).
17. S. J. Elliott, E. M. Ku, and B. Lineton, "A state space model for cochlear mechanics," The Journal of the Acoustical Society of America **122**, 2759–2771 (2007).
18. J. Meaud and C. Lemons, "Nonlinear response to a click in a time-domain model of the mammalian ear," The Journal of the Acoustical Society of America **138**, 193–207 (2015).
19. K. K. Charaziak, P. Souza, and J. H. Siegel, "Stimulus-frequency otoacoustic emission suppression tuning in humans: comparison to behavioral tuning," Journal of the Association for Research in Otolaryngology **14**, 843–862 (2013).

Numerical Analysis of Nonlinearity of Outer Hair Cells Based on Comparison with Measurements of DPOAEs

Sinyoung Lee[1, a)] and Takuji Koike[1, b)]

[1] *Department of Mechanical and Intelligent Systems Engineering*
Graduate School of Informatics and Engineering
The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo, Japan

[a)] Corresponding author: sylee@bio.mce.uec.ac.jp
[b)]koike@mce.uec.ac.jp

Abstract. Distortion-product otoacoustic emissions (DPOAEs) are considered to be generated by nonlinear force of the outer hair cells (OHC), and those frequencies and levels depend on the function curve of the force. Each order of distortion component is generally generated by different degree term of a nonlinear function which is approximated by a power series. In this study, the nonlinear force of OHC was formulated based on the distortion components detected in measurements of DPOAEs. The input/output (I/O) functions of DPOAEs were measured in normal-hearing human ears. The I/O functions of each component showed different nonlinear curve and varied by stimuli conditions. Odd-order of distortion components (e.g. $2f_1$-f_2 or $3f_2$-$2f_1$) were generated in all the ears, on the other hand, even-order of distortion components (e.g. f_2-f_1 or $3f_1$-f_2) were seldom detected. Thus, the force of the OHC, P_{OHC}, was formulated as a function including both of saturability and symmetry around operating point, because even-order of distortion components were barely detected. The I/O functions of DPOAEs obtained from numerical analysis using the finite-element model of the human cochlea which includes the force were compared with the measurements in normal-hearing human ears, and validity of the force was evaluated. As the result, the I/O functions of each component obtained from the displacement of the stapes footplate of the model showed different nonlinear curve. If the sound pressure gains of the middle ear are assumed to be constant at a certain frequency regardless of sound pressure level, the ease of saturation of the I/O function obtained from the model is considered to be comparable to that obtained from the measurements in the human ear. The curve of the I/O functions of each component obtained from the model approximately matched that obtained from the measurements.

INTRODUCTION

The outer hair cells (OHCs) show somatic motility which non-linearly amplifies the vibration of the basilar membrane (BM). Distortion-product otoacoustic emissions (DPOAEs) are considered as intermodulation distortion components ($mf_1 \pm nf_2$, here, m and n are integers) produced by the nonlinear activities of the OHCs when the cochlea was simultaneously stimulated by two pure tones with different frequencies (f_1 and f_2). Intermodulation distortion is generated when a complex signal is applied in a nonlinear system and is a general acoustic phenomenon. Each order of distortion component is generally generated by different degree term of a nonlinear function which is approximated by a power series, i.e. N-order (here, N is an integer number) of distortion components are generated by N-degree term of the function. The coefficient of each distortion component can be calculated by plugging an expression composed of sum of sinusoidal wave of f_1, f_2 into a variable of the power series. For example, distortion component ($f_1 + f_2$) and harmonics ($2f_1$, $2f_2$) are generated when expand and convert the square of the expression. The coefficient of each components are represented by polynomial expression of coefficients of the sinusoidal waves. The non-linear activity of the mammalian OHC has been investigated experimentally, e.g., its transfer function between the current and stereocilia bundle deflection is shown as a sigmoid-shaped Boltzmann function[1]. The frequencies and levels of DPOAEs depend on the shape of the nonlinearity curve[2], and the optimum frequency ratio f_2/f_1 of the stimuli varies for each component of DPOAE[3], e.g. $2f_1$-f_2 and $2f_2$-f_1. Therefore, a nonlinear function of the OHCs can be formulated based on the levels of each distortion component obtained from measurement of the DPOAEs. In this study, we aimed

to formulate the nonlinear activity of the OHC based on the distortion components detected in measurements of DPOAEs with varying the stimuli condition. The input/output (I/O) functions of DPOAEs were measured in normal-hearing human ears. The force of the OHC, P_{OHC}, was formulated as a function of velocity of the BM was introduced on the surface of the BM of a finite-element model of human cochlea. DPOAEs were simulated by applying sound pressure composed of frequencies f_1 and f_2 to the stapes head of the model. Each distortion component of the vibration obtained at the stapes-footplate and the BM were calculated. The I/O functions of DPOAEs obtained from the numerical analysis were compared with those obtained from the measurements to evaluate the validity of P_{OHC}.

MEASUREMENT OF DISTORTION PRODUCT OTOACOUSTIC EMISSIONS

Methods

The measurement system consists of a probe composed of two transducers and one microphone (ER-10C, Etymotic Research), an AD/DA converter (USB-6251, National Instruments) and a PC. The software created in LabVIEW (Ver. 2018, National Instruments) was used for both measurements and data analyses. The recorded data was averaged in the time domain by considering the frequency stability of the DP components. In addition, the number of averaging, time window length and frequency resolution were set to detect the exact frequency of each DP component and reduce the noise level. DPOAEs were measured in 4 normal-hearing subjects, i.e., 2 females and 2 males in their 20s. Not only the third-order DP component, which is usually investigated in clinical situations, but also second-order, fourth-order and fifth-order components were investigated. The second-order (f_2-f_1) and the fourth-order ($3f_1$-f_2) DP components were selected for comparison between even-order and odd-order DP. Both lower sidebands ($2f_1$-f_2) and the upper sideband ($2f_2$-f_1) of the third-order DP component were measured. Those DP levels were measured by varying the frequency f_2/f_1, because several studies[2, 3] showed that the optimum frequency ratio f_2/f_1 of the stimuli varied for each DP component. DPOAEs were measured by sweeping frequency f_1 with ranging from 400 Hz to 7200 Hz in 400 Hz steps. The frequency f_2 was varied corresponding to a frequency ratio f_2/f_1 of 1.1 or 1.2. The sound pressure level for f_1 and f_2 were set to the same value, and the both levels were changed from 10 dB to 65 dB SPL in 5 dB steps to measure the input/output (I/O) function. Each measurement was done 2 times for assessment of reproducibility. The number of averaging was set to 40 in all measurements. Window length was set to 0.5 s, i.e., frequency resolution of 2 Hz, and the sampling rate was set to 1 MHz in the all measurements. The DP component produced by the system itself was evaluated by measuring DPOAEs by inserting the probe into a 2 cc cylinder with one end closed.

Results

Figure 1 shows the levels of the frequency components obtained from a subject when complex tone (f_1=2000 Hz, f_2=2400 Hz) was applied. The combination tone of the stimulus frequencies (f_1+f_2) and the harmonic tones of stimulus frequencies were obtained from the measurement using the cylinder, and the levels of the other frequency components were smaller than -20 dB SPL. Therefore, those components were considered to be produced by the measurement system itself. A baseline for estimating whether the level of each DP component was sufficient for detection was determined as -20 dB SPL. The third-order DP component $2f_1$-f_2 was detected in the subject regardless of the frequency ratio f_2/f_1. The fifth-order distortion component ($3f_1$-$2f_2$) was detected when the frequency ratio f_2/f_1 was 1.1 or 1.2. The other DP components, i.e., $2f_2$-f_1 and $4f_1$-$3f_2$, were generated when the frequency ratio f_2/f_1 was 1.1. By contrast, even-order distortion components (i.e., f_2-f_1, $3f_1$-f_2) were barely detected regardless of the frequency ratio f_2/f_1. The I/O functions, i.e., change the DP level with specific frequency are shown in Figure 2. The DP levels became larger with an increase in the sound pressure level of the stimulus and saturated. The I/O functions of the $2f_1$-f_2 component measured in all ears were included in the normal range of the previous study[4]. The DP levels of the $2f_1$-f_2 component were more detectable when the frequency ratio f_2/f_1 was 1.2, and the other components were more detectable when the frequency ratio f_2/f_1 was 1.1. The I/O functions of each DP component showed different saturation tendencies.

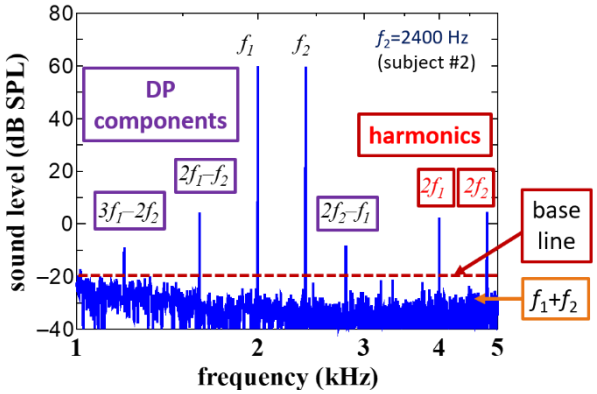
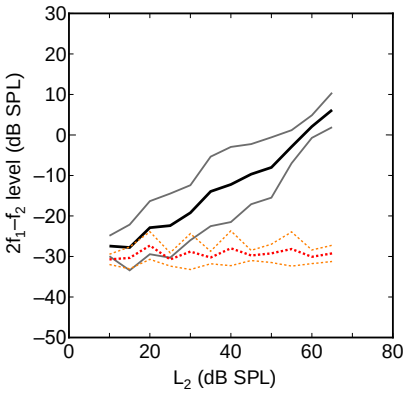


FIGURE 1. Levels of the frequency components obtained by DPOAEs measurement (f_2=2400 Hz, 60 dB SPL).

(a)



(b)

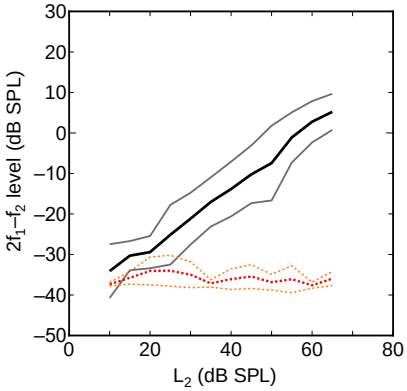


FIGURE 2. Means and standard deviation of the I/O function. (a) f_1=1600 Hz, f_2=1920 Hz, (b) f_1=3200 Hz, f_2=3840 Hz were applied. The dotted line shows the noise floor (average of DP component plus/minus 100 Hz area).

FORMULATION OF THE EXCITATION FORCE OF THE OUTER HAIR CELL

Each order of the DP component is generally generated by a different degree term of a nonlinear function which is approximated by a power series. Although the force curve of the OHC is assumed to be a function of the movement of the stereocilia with saturation, its symmetry around the operating point has not been clarified[5]. A nonlinear function including both saturability and symmetry around the operating point, e.g., the arctangent function, is approximated with only odd-degree terms (Fig. 3 (a)). An arctangent function represents asymmetry by shifting the operating point (Fig. 3 (b)). One of the non-linear activities measured in the mammalian OHC was shown as a sigmoid-shaped Boltzmann function[1]. A sigmoid function has both saturability and asymmetry around the operating point and is approximated with even-degree terms and odd-degree terms by Maclaurin expansion around the operating point (Fig. 3 (b)). According to the measurement results, the odd-order of DP components were generated and the even-order of DP were barely detected. Thus, the excitation force of the OHC, P_{OHC}, was formulated as a function of velocity of the BM, V_{BM} with using the arctangent function in Eq. (1) below:

$$P_{OHC} = a(y)b(y)[\tan^{-1}\{c(y)V_{BM}\}] \quad (1)$$

where,

$$a(y) = 13.5 - 180y, \quad (2)$$

$$b(y) = \cos(0.436 + 10.9y)\ , \text{and} \quad (3)$$

$$c(y) = 7200 - 20400y\ . \quad (4)$$

Here, $a(y)$, $b(y)$, $c(y)$ are functions of the distance from the base of the BM, y ($0 \leq y \leq 32$). Each function represents different characteristics of the OHCs at each location of the BM and is determined based on the measurement of a guinea pig's isolated OHC. The $a(y)$ shows that the generative force of the OHC at the basal side is smaller than the apical side[6, 7]. The $b(y)$ shows that the angle between the BM and the OHC is 25° at the base of BM and 45° at the apex of BM[8], and the $c(y)$ shows the difference in the saturation curve of the generative force according to the position on the BM[9]. Here, the generative force at the basal side is assumed to be easier to saturate than the apical side. Figure 4 shows an example of P_{OHC} at different places on the BM. The generative force and the ease of saturation are different at each place on the BM. P_{OHC} was applied to the portion of the surface of the BM (Fig. 5).

ASSESSMENTS OF THE FORMULATION

Methods

The human cochlear finite element model (Fig. 5) used in this study was constructed following the model of Koike *et al.*[10] The cochlear model consists of structural parts and fluid parts. This model was simplified to avoid complications of calculation as follows: the BM of this model has a linear shape and the width and thickness of the BM change linearly from the base to apex. The scala vestibule (SV) and the scala media (SM) were unified and Reissner's membrane (RM) was ignored. The effect of the straightening on cochlear vibration was slight and was corrected by adjusting the mechanical property of the BM[10]. Each parameter of the model was determined based on the reported values. Validity of the model was verified by compared with some measurements, e.g. Characteristic frequency (CF) distribution map by Greenwood[11]. The lymph was considered incompressible viscous fluid. The BM was assumed to be an orthotropic material because the collagen fibers run in the direction of the width of the BM (in the direction of the X-axis).

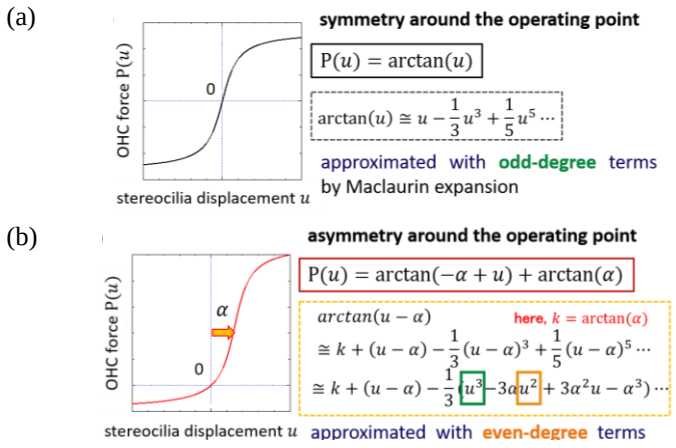


FIGURE 3. Approximation of the nonlinear functions (a) symmetry function around the operating point, (b) asymmetry function around the operating point.

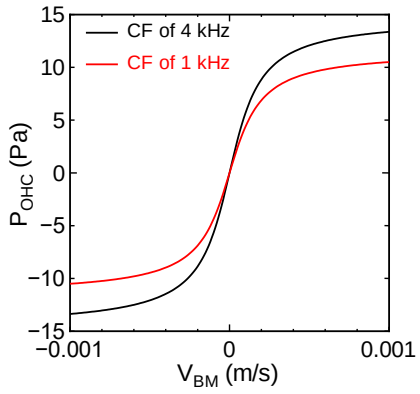


FIGURE 4. Examples of P_{OHC} at CF location of 1 kHz and 4 kHz.

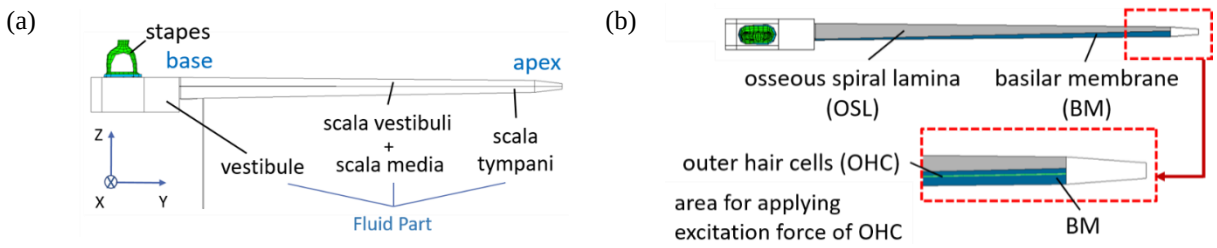


FIGURE 5. Finite element model of the human stapes and cochlea. (a) Front view. (b) Top view.

DPOAEs obtained from numerical analysis using the finite-element model of the human cochlea including P_{OHC} were compared with the measurements of DPOAEs and the validity of the P_{OHC} was evaluated. The optimal frequency ratio f_2/f_1 of each DP component and features of the I/O function (e.g. ease of saturation) were compared. DPOAEs were simulated by applying a complex tone ($L_1\sin(2\pi f_1 t) + L_2\sin(2\pi f_2 t)$) to the stapes head of the model. The DP components are generated by the vibration of the BM introduced the P_{OHC}, and are considered to be transmitted to the stapes via lymph flow. Thus, the components are obtained on the BM and the stapes footplate of the model. The frequency f_2 of 2 kHz, 4 kHz and the frequency ratio f_2/f_1 of 1.2 were used. The sound pressure levels for f_1 and f_2 were set to the same. The sound pressure level was set to range from 40 dB to 120 dB SPL in 10 dB steps to measure the input/output (I/O) function.

Results

Figure 6 shows I/O function obtained at the stapes footplate of the model ((a)~(c)) and in the human ear ((d)~(f)). If the anterograde and retrograde pressure gains of the middle ear are assumed to be constant at a certain frequency regardless of sound pressure, features of the I/O function (e.g., ease of saturation) obtained at the stapes footplate of the model are considered to be comparable to that measured in the human ear. The pressure transmission gain of the middle ear is approximately 45 dB considering the area ratio of the tympanic membrane to the stapes-footplate. The reverse middle-ear transfer functions in guinea pig[12] and gerbil[13] have been measured, and the reverse gain of the intracochlear DP component to the DP component at the ear canal was assumed to be approximately -35 dB. In addition, the I/O function of P_{OHC} obtained at the BM of the model where P_{OHC} showed maximum value was similar to that obtained from the stapes footplate. Therefore, rigorous formulation of P_{OHC} by comparing the I/O function of each distortion component between the measurement and numerical analysis is possible. Those transmission gains should be considered for comparison between the I/O function obtained from measurements and the stapes-footplate of the model. When the I/O function obtained from the model is translated, the I/O function could approximately match with the I/O function of the measurement.

DISCUSSION

I/O functions obtained from the measurements show individual difference of features. The validity of the P_{OHC} was evaluated approximately by comparing the ease of saturation of the I/O functions of DPOAEs between numerical analysis and measurements (Fig. 6). However, the horizontal axis, i.e. input sound pressure levels, between the numerical analysis and measurements was not matched. A rigorous formulation of P_{OHC} by comparing the I/O function of each distortion component between the measurement and numerical analysis is possible when the anterograde and retrograde pressure gains of the middle ear is considered. Future work also entails modifying P_{OHC} as below: modifying $a(y)$ for generating the distortion component when low-level of stimulus was applied, and modifying $c(y)$ for saturating the component when low-level of stimulus was applied.

ACKNOWLEDGMENTS

This study was supported by JSPS KAKENHI (Grants No. 19J13684) and Grant-in-Aid from Hirose Foundation.

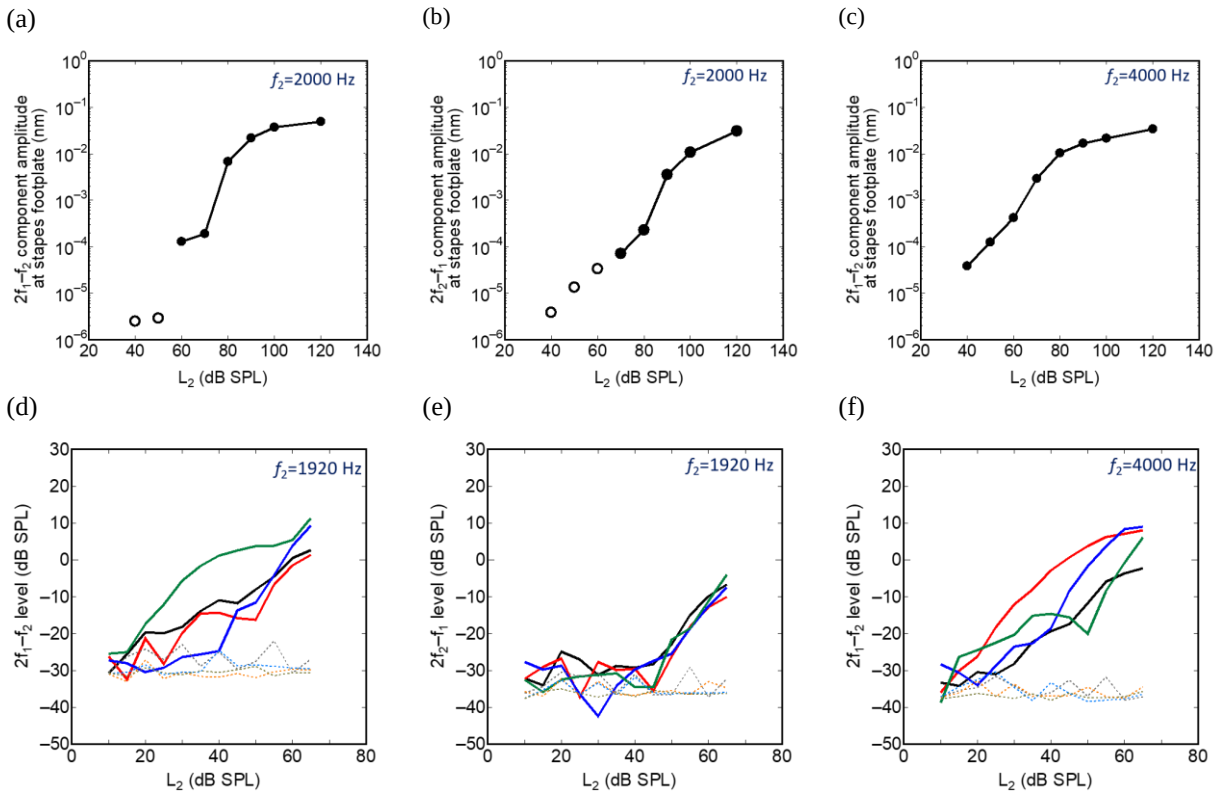


FIGURE 6. I/O function obtained at the stapes footplate of the model ((a)~(c)) and in the human ear ((d)~(f)). (a), (c), (d) and (f) show I/O function of $2f_1$-f_2 component. (b) and (e) show I/O function of $2f_2$-f_1 component. The opened circle, ○ represent that the level of the distortion component was no more than 5 dB larger than an average level of plus/minus one frequency bin. Each colored line in (d), (e) and (f) shows I/O function of each subject. The dotted line shows the noise floor.

REFERENCES

1. P. Avan, B. Buki and C. Petit, Physiol. Rev., **93**, 1563-1619 (2013).
2. N. A. Lukahikin, A. V. Lukashkina and J. I. Russell, J. Acoust. Soc. Am., **111**(6), 2740-2748 (2002).
3. A. Moulin, J. Acoust. Soc. Am., **107**(3), 1471-1486 (2000).
4. P. A. Dorn, D. Konrad-Martin, S. T. Neely, D. H. Keefe, E. Cyr and M. P. Gorga, J. Acoust. Soc. Am., **110** (6), 3119-3131 (2001).
5. C. A. Shera and C. Abdala, "5. Otoacoustic emissions - Mechanisms and applications," in *Translational perspectives in auditory neuroscience: Hearing across the life span - Assessment and disorders* (Plural publishing, 2012), pp. 123-159.
6. E. Chan and M. Ulfendahl, Acta. Physiol., **161,** 533-539 (1997).
7. Santos-sacchi, S. Kakehata, T. Kikuchi, Y. Katori, and T. Takasaka, Neurosci., **256,** 155-158 (1998).
8. S. Ramamoorthy, N. V. Deo and K. Grosh, J. Acoust. Soc. Am., **121**(5), 2758-2773 (2007).
9. C. Zinn, H. Maier, H. P. Zenner and W. A. Gummer, Hear. Res., **142**, 159-183 (2000).
10. T. Koike, C. Sakamoto, T. Sakashita, K. Hayashi and S. Kanzaki, Hear. Res., **283**, 117-125 (2012).
11. D. D. Greenwood, J. Acoust. Soc. Am., **87**, 2592-2605 (1990).
12. P. Magnan, P. Avan, A. Dancer, J. Smurzynski and R. Probst, Hear. Res., **107**, 41-45 (1997).
13. W. Dong and S. E. Olson, J. Neuro-physiol., **95**, 2951-2961 (2006)

Speech-DPOAEs for probing speech processing in the inner ear

Marina Saiz-Alia,[1, a)] Peter Miller,[1, b)] and Tobias Reichenbach[2, 1, c)]
[1)]*Department of Bioengineering,*
Imperial College London,
South Kensington Campus,
SW7 2BX,
U.K.
[2)]*Department of Artificial Intelligence in Biomedical Engineering,*
Friedrich-Alexander-Universität Erlangen-Nürnberg,
Konrad-Zuse-Strasse 3/5,
91052 Erlangen,
Germany

(Dated: 29 April 2022)

Abstract. Speech is a complex real-world stimulus that consists of time-varying contributions in many frequency bands. The inner ear, or cochlea, not only transduces the acoustic vibrations into electrical signals, but also spatially segregates the acoustic waveform into different frequency components, as well as mechanically amplifies the vibrations in frequency bands where the amplitude is low. This processing of the speech signal can potentially be controlled through efferent nerve fibers that extend from higher processing centers in the brain to the cochlea. However, we are currently lacking methodological tools to non-invasively investigate the putative role of top-down feedback on speech processing in humans. Here we develop such a method that builds on distortion-product otoacoustic emissions (DPOAEs) related to the temporal fine structure of the voiced parts of speech (speech-DPOAEs).

INTRODUCTION

Understanding speech is crucial for human interaction and communication [1]. Humans are indeed experts at understanding speech even in adverse listening conditions such as loud background noise in a noisy restaurant, coffee shop or in traffic.

Our ability to understand speech in noise is, however, vulnerable. Hearing impairments affect more than 16% of the adult population in the E.U., and more than 5% in children [1, 2, 3, 4, 5]. Moreover, hearing impairments progress with age and are hence a particular problem in the E.U.'s aging society: 40% of people above age 50 and 70% of those above age 70 have some form of hearing disorder. Such hearing impairment impacts first and foremost the ability to understand speech in noisy backgrounds, and current hearing aids are unfortunately of little help [6]. The resulting difficulties in communicating in many social settings have major effects on an afflicted person's personal, social and economic development [4, 7].

A better diagnosis of hearing impairment as well as the development of corresponding treatments require a better understanding of the neural pathways of hearing, and in particular of the neural mechanisms that allow a healthy person to understand speech in noise. This issue has traditionally been tackled by investigating the underlying processes in the ear and brain through simple acoustic signals such as pure tones or clicks. As an example, pure-tone audiometry uses pure tones of different frequencies to assess the hearing threshold of a subject at those frequencies. As another example, click-evoked auditory brainstem responses record the electrical activity of the brainstem at different latencies when a subject hears a series of short clicks. These methods are highly useful for assessing the basic functioning of different parts of the auditory system, such as the inner ear and the auditory brainstem. However, they do not fully reveal the neural mechanisms that allow us to process more complex signals such as speech in background noise.

Recent research has therefore sought to measure neural responses to natural speech, and therefrom learn more about the neural mechanisms of speech processing. An important example concerns the neural response to the temporal fine structure of speech. Many parts of speech are voiced, originating from vibrations of the vocal fold. The vibration occurs at the so-called fundamental frequency, typically between 100 - 300 Hz. This frequency as well as its many

[a)]Electronic mail: marina.saiz-alia15@imperial.ac.uk
[b)]Electronic mail: peter.miller15@imperial.ac.uk
[c)]Corresponding author: tobias.j.reichenbach@fau.de

higher harmonics constitute the so-called temporal fine structure and carry most of the acoustic energy of a voiced speech segment. When measuring the neural response to running speech through electroencephalography (EEG) or magnetoencephalography (MEG), a strong signal can be found at the fundamental frequency, as well as, to a much lower degree, at the higher harmonics. This response presumably arises predominantly from the auditory brainstem, although recent work has also identified cortical contributions as well [8, 9, 10, 11].

The neural response at the fundamental frequency of speech has traditionally been studied using many repeated presentations of the same short speech token, such as a single syllable [12]. We as well as others, however, recently developed the statistical methodology to measure this neural response in running speech [13, 14]. To deal with the time-varying nature of the fundamental frequency in natural speech, we first extracted a waveform that, at each time point, oscillates at the fundamental frequency. In a second step, we then related that waveform to the EEG recordings obtained when a person listened to the corresponding speech. The relation allowed to infer both magnitude and latency of the neural response [13, 15].

Importantly, our work on the neural response at the fundamental frequency of speech showed that this neural signal is consistently modulated by selective attention [16]. Because the neural response to the fundamental frequency of speech arises predominantly in subcortical structures such as the inferior colliculus, the modulation of this response through selective attention likely involves efferent feedback from the auditory cortex to these subcortical areas [17, 18].

An important part of the auditory system for which responses to natural speech have not yet been measured is the inner ear. The inner ear houses the mechanosensitive hair cells that transduce the mechanical sound vibration into electrical signals. Strikingly, through a so-called active process, these hair cells also mechanically amplify weak sound stimulation [19, 20]. Otoacoustic emissions (OAEs) are generated as a byproduct of the amplification [21]. As an example, when stimulated with two frequencies, so-called distortion-product otoacoustic emissions (DPOAEs) emerge at other frequencies that correspond to linear combinations of the primary ones. These sounds can be measured with a sensitive microphone in the ear canal.

DPOAEs have so far almost exclusively been elicited by pure tones, such that the emissions are pure tones as well. These measurements have established DPOAEs as a highly useful diagnostic test of the inner ear's functioning that is widely applied to assess hearing in newborns [22].

Importantly, DPOAEs as well as other types of OAEs have also been employed to show that the inner ear's activity can be influenced by efferent feedback from higher neural processing centers. In particular, the inner ear's outer hair cells that provide mechanical feedback and generate the DPOAEs are innervated by the medial olivocochlear (MOC) bundle. Activation of this bundle, for instance through ipsilateral or contralateral noise, leads to a reduction of mechanical activity of the outer hair cells and thereby to a reduction of DPOAEs (MOC reflex) [23, 24].

Computational studies have shown that the regulation of cochlear activity through the efferent feedback can aid speech processing [25]. Such efferent feedback can, for instance, help to filter out unwanted background noise by reducing the amplification at cochlear locations that lie inbetween locations at which resolved harmonics of the target speech are detected. The efferent feedback to the cochlea could therefore contribute significantly to speech-in-noise comprehension, and could explain a portion of the attentional modulation of the neural response to the temporal fine structure of speech that we have measured earlier [13, 15]. On the other hand, impairment with this pathway could constitute one form of speech-in-noise difficulty.

However, due to the complexity of speech, the contribution of the active feedback in the inner ear to speech processing has not yet been investigated.

Methods and Results

Our approach to measure speech-DPOAEs, that is, DPOAEs related to the temporal fine structure of speech, was as follows (Figure 1A). We first extracted two waveforms w_{P1} and w_{P2} that corresponded to two nearby harmonics of the fundamental frequency $f(t)$ of a speech signal, say to $a(f(t)$ and $bf(t)$ with a and b two close integers. The two waveforms were obtained by bandpass-filtering the speech signal in a narrow region of the frequency range of the corresponding harmonics.

If the speech signal was monotone, that is, if the fundamental frequency $f(t)$ was constant at a value f, the waveforms w_{P1} and w_{P2} would simply be sinusoidal oscillations at the constant frequencies af and bf of the corresponding harmonics. In natural speech, however, the fundamental frequency varies over time, such that the frequencies of the harmonics vary as well. The waveforms w_{P1} and w_{P2} are therefore not sinusoidal, although, on a short temporal scale, they appear similar to a sinusoidal oscillation at the waveform's instantaneous frequency $af(t)$ or $bf(t)$.

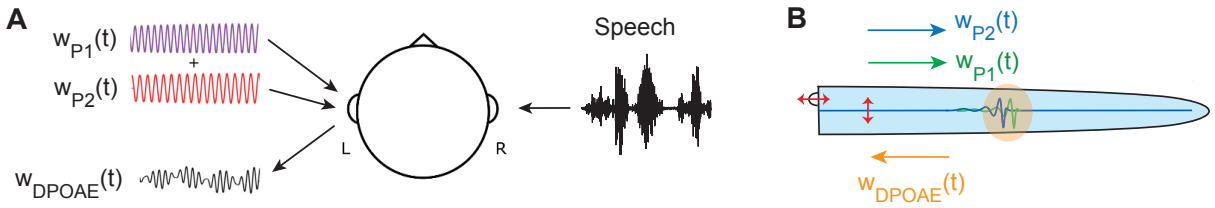


FIGURE 1. Principle of speech-DPOAEs. **A**, One ear is stimulated with speech. DPOAEs related to the harmonic structure of that speech signal can then be evoked and recorded from the contralateral ear. To this end, two waveforms w_{P1} and w_{P2} that correspond to two nearby harmonics of the fundamental frequency of the speech signal are delivered. The nonlinear activity in the cochlea leads to the generation of distortion products, and hence to the emission of a wave w_{DPOAE}. This emission can be measured by correlation the acoustic signal in the ear canal with the waveform w_{DPOAE}. **B**, The distortion product is generated by the cochlear nonlinearity, presumably in the overlap region of the peaks of the traveling waves elicited by w_{P1} and w_{P2}.

Because the waveforms w_{P1} and w_{P2} are narrow-band signals, the cochlear activity elicited by them is likely similar to that in response to a pure tone (Figure 1B). In particular, the two waveforms will elicit traveling waves on the basilar membrane that peak at near the characteristic place of the instantaneous frequencies $af(t)$ and $bf(t)$ of the two waveforms. In the overlap regions of the peak, the cochlear nonlinearity will then produce distortion, such as at the cubic distortion frequencies $2af(t)-bf(t)=(2a-b)f(t)$ and $2bf(t)-af(t)=(2b-a)f(t)$. These cubic distortion frequencies are therefore again harmonics of the speech signal's fundamental frequency $f(t)$.

In the case of monotone speech, the cubic distortion signals would be pure tones, at the constant frequencies $(2a-b)f$ and $(2b-a)f$. Both the stimulation signals w_{P1} and w_{P2} and the speech-DPOAEs could then be measured through the Fourier transform of a microphone recording in the ear canal. In particular, the stimulation signals would be evident through peaks at the frequencies af and bf, and the cubic DPOAEs through peaks at the frequencies $(2a-b)f$ and $(2b-a)f$.

Because the fundamental frequency of natural speech varies, neither the stimulation waveforms nor the resulting speech-DPOAEs will be apparent in the Fourier transform of the microphone recording. To assess these emissions, we instead employed a technique that we recently used to measure the neural response at the fundamental frequency [13]. There, we computed the cross-correlation of the expected waveform w_{DPOAE} with the microphone recording. If the microphone recording contains an emission at the time-varying distortion frequency, the cross-correlation will exhibit a peak at the delay of the emission. The peak will be the broader the more narrow-band the waveform w_{DPOAE} is. In particular, for monotone speech where the emission consists of a pure tone, the peak will be infinitely broad. For natural speech, however, where the emission varies across a certain frequency range, the peak in the cross-correlation will have a finite width, allowing to assess the presence of the emission as well as to obtain an estimate of its latency.

Of course, this technique can not only be employed to measure the presence of the emission waveform w_{DPOAE} in the microphone recording, but also those of the stimulation waveforms w_{P1} and w_{P2}.

To put these ideas to the test, we constructed waveforms that oscillated at three nearby harmonics of the fundamental frequency of a voiced speech signal, the 7th, the 9th, and the 11th (Figure 2). We stimulated one ear (contralateral to the ear to which the speech stimuli were presented) with two of these three harmonics, the 9th and the 11th. We measured the signal in the ear canal with a very sensitive microphone (ER10X, Etymotics).)We verified the presence of the stimulation waveforms w_9 and w_{11} in the ear canal through computing the cross-correlation of the waveforms with the microphone recording. Both cross-correlation showed pronounced peaks at the delay of 0 ms, evidencing that both waveforms were present instantaneously in the ear canal where they were generated by small speakers.

The inner ear's nonlinear distortion then created a signal, amongst others, at the 7th harmonic. We measured this signal by cross-correlating the microphone recording with the waveform w_7 that corresponded to the 7th harmonic. The cross-correlation showed a peak at a delay of about 2.2 ms, in line with estimates of the temporal delay that is required for the stimulation signals to propagate into the inner ear and for the distortion product to propagate back from its generation site to the middle ear and the ear canal. This showed that the distortion signal was indeed present in the ear canal.

We repeated the measurement, but this time positioning the DPOAE probe with the speakers and microphone outside the ear canal (Figure 2B, D, F). The stimulation waveforms w_9 and w_{11} were still present in the microphone recording, as evidence by pronounced peaks in the corresponding cross-correlation. The waveform w_7 that corresponded to a distortion product, however, did not yield a peak in the cross-correlation. The speech-DPOAE could therefore not be detected in the microphone recording. This verified that the speech-DPOAE observed above was indeed produced by the inner ear and not by the DPOAE probe itself.

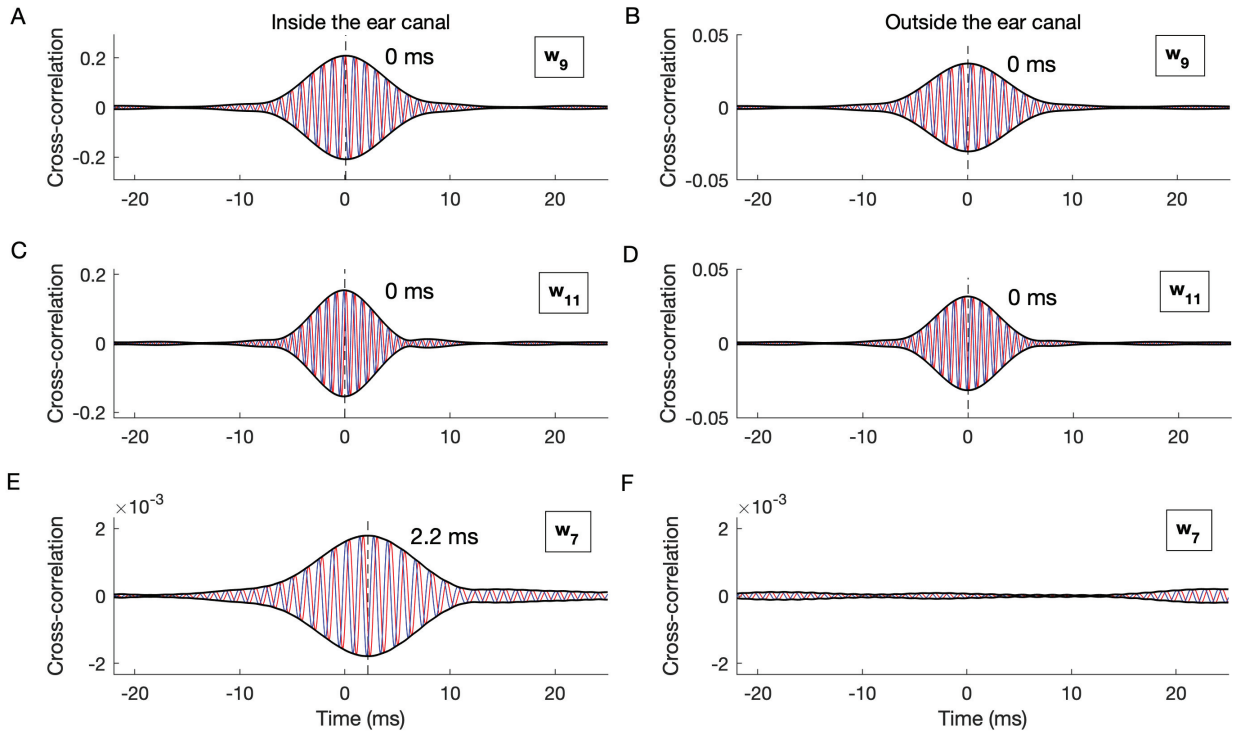


FIGURE 2. An example of recorded speech-DPOAEs. **A,C,E**, We presented a subject with the waveforms w_9 and w_{11}, corresponding to the 9th and 11th harmonics of a speech signal, in the contralateral ear, and recorded the sound with a microphone in the same ear. **A,C**, We show the correlation of the microphone recording with the waveforms w_9 and w_{11} as red lines. We also show the correlation with these waveforms shifted by 90 degrees (blue line) as well as the envelope of the cross-correlation (black line). The cross-correlations of the microphone exhibits a peak at 0 ms, confirming that the microphone is picking up the sound. **E**, The cross-correlation of the microphone recording with the waveform w_7 of the corresponding DPOAE shows a peak at a latency of 2.2 ms, evidencing an otoacoustic emission at this delay. **B, D, F**, When the OAE probe is placed outside the ear canal, the microphone in the probe still detects the signals that are played by the probe (**B,D**), but not the otoacoustic emission (**F**), confirming that the speech-DPOAE w_7 is generated by the inner ear and not by the measurement device.

Discussion

We have proposed a method to measure DPOAEs related to the temporal fine structure of natural speech. The method extracts waveforms of the harmonics of the time-varying fundamental frequency of a speech signal. It then employs two waveforms at nearby frequencies to stimulate the cochlea. The resulting speech-DPOAE emerges at another waveform and can be measured through cross-correlation of the waveform with a microphone recording from the ear canal.

In our work on verifying this methodology, we have stimulated one ear with the speech signal, and measured the speech-DPOAEs from the contralateral ear. Future work may extend this method to measure speech-DPOAEs from the ipsilateral ear as well.

Importantly, the speech-DPOAEs allow to assess cochlear responses to natural speech in humans in a non-invasive manner. Natural speech has recently received increased attention. In particular, we have shown that, through efferent feedback, the early neural response at the fundamental waveform is already modulated by selective attention to one of two competing talkers [13]. The attentional modulation can allow to decode the attentional focus of a listener from short segments of EEG recordings, suggesting potential applications in mind-controlled auditory prosthesis that decode a listener's attention from neural recordings and then use it to steer the prosthesis towards the target of attention [26]. In addition, and pointing towards clinical applications, we have shown that the strength of the attentional modulation relates to a subject's ability to understand speech in noise [27].

Because the efferent feedback extends to the inner ear, the latter might be involved in selective attention to speech as well. We have already started to investigate this issue through the use of our speech-DPOAEs [28]. Our results

provide tentative evidence that the speech-DPOAEs are larger when the corresponding speech signal is attended then-https://www.overleaf.com/project/620b719042daa8a0a93cd1ec when it is ignored. In the future, it will be important to further investigate such attentional effects, for instance regarding the issue of resolved versus unresolved harmonics of speech. Moreover, it will be interesting to see how such attentional effects link to deficits with speech processing, such as speech-in-noise impairments.

For potential medical applications of speech-DPOAEs, it is worth considering the clinical implications of a recent breakthrough in our understanding of neural speech processing, namely the tracking of speech rhythms in the auditory cortex. In this regard, the neural activity in the theta frequency band (4 - 8 Hz) has been shown by my research group and others to reflect largely the processing of lower-level acoustic aspects of speech, while the slower delta frequency band (1 - 4 Hz) contains information on linguistic aspects of speech processing, including on speech comprehension [29, 30, 31, 32]. The neural tracking of speech rhythms is further modulated by selective attention: when listening to one of two competing talkers, for instance, the cortical activity tracks the speech rhythms of the attended speaker more than those of the unattended talker [33, 34, 35].

These novel insights are now being employed towards a better understanding of neurological impairments. Both autism and dyslexia, for instance, have already been linked to impairments with the cortical tracking of speech rhythms [36, 37, 38]. Moreover, disorders of consciousness have been shown to be better assessible when including the neural tracking of speech in the diagnostic toolkit [39]. While such applications are still in their infancy, they demonstrate the potential of using running speech, rather than simpler speech tokens or yet simpler acoustic signals, for better understanding our hearing system as well as for diagnosing neurological impairments. Whether efferent feedback to the cochlea is involved as well, and if so, in which manner, is an exciting research topic that can now be studied.

ACKNOWLEDGMENTS

This research was supported by the Royal British Legion Centre for Blast Injury Studies, from the 'la Caixa' Foundation (LCF/BQ/EU15/10350044) as well as by EPSRC grants EP/M026728/1 and EP/R032602/1.

REFERENCES

1. C. Cherry, "On human communication," (1966).
2. A. C. DAvIs, "The prevalence of hearing impairment and reported hearing disability among adults in great britain," International journal of epidemiology **18**, 911–917 (1989).
3. A. Davis and S. Wood, "The epidemiology of childhood hearing impairment: factors relevant to planning of services," British Journal of Audiology **26**, 77–90 (1992).
4. D. R. Moore, S. Rosen, D.-E. Bamiou, N. G. Campbell, and T. Sirimanna, "Evolving concepts of developmental auditory processing disorder (apd): a british society of audiology apd special interest group 'white paper'," International Journal of Audiology **52**, 3–13 (2013).
5. A. C. Davis and H. J. Hoffman, "Hearing loss: rising prevalence and impact," Bulletin of the World Health Organization **97**, 646 (2019).
6. R. Plomp, "A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired," Journal of Speech, Language, and Hearing Research **29**, 146–154 (1986).
7. D. R. Moore, "Auditory processing disorder (apd): Definition, diagnosis, neural basis, and intervention," Audiological Medicine **4**, 4–11 (2006).
8. E. Skoe and N. Kraus, "Auditory brainstem response to complex sounds: a tutorial," Ear and Hearing **31**, 302 (2010).
9. E. B. Coffey, S. C. Herholz, A. M. Chepesiuk, S. Baillet, and R. J. Zatorre, "Cortical contributions to the auditory frequency-following response revealed by meg," Nature Communications **7**, 1–11 (2016).
10. E. B. Coffey, T. Nicol, T. White-Schwoch, B. Chandrasekaran, J. Krizman, E. Skoe, R. J. Zatorre, and N. Kraus, "Evolving perspectives on the sources of the frequency-following response," Nature Communications **10**, 1–10 (2019).
11. G. M. Bidelman, "Subcortical sources dominate the neuroelectric auditory frequency-following response to speech," Neuroimage **175**, 56–69 (2018).
12. N. Kraus and T. Nicol, "Brainstem encoding of speech and music sounds in humans," Oxford Handbook of the Auditory Brainstem. Oxford, UK: Oxford University Press. Advance online publication (2018).
13. A. E. Forte, O. Etard, and T. Reichenbach, "The human auditory brainstem response to running speech reveals a subcortical mechanism for selective attention," elife **6**, e27203 (2017).
14. R. K. Maddox and A. K. Lee, "Auditory brainstem responses to continuous natural speech in human listeners," Eneuro **5** (2018).
15. O. Etard, M. Kegler, C. Braiman, A. E. Forte, and T. Reichenbach, "Decoding of selective attention to continuous speech from the human auditory brainstem response," Neuroimage **200**, 1–11 (2019).
16. M. Kegler, O. Etard, A. Forte, and T. Reichenbach, "Pyhton code for the computation of complex trfs (ctrfs). github," (2018).
17. R. F. Huffman and O. Henson Jr, "The descending auditory pathway and acousticomotor systems: connections with the inferior colliculus," Brain Research Reviews **15**, 295–323 (1990).

18. M. Saiz-Alía and T. Reichenbach, "Computational modeling of the auditory brainstem response to continuous speech," Journal of Neural Engineering **17**, 036035 (2020).
19. T. Reichenbach and A. Hudspeth, "The physics of hearing: fluid mechanics and the active process of the inner ear," Reports on Progress in Physics **77**, 076601 (2014).
20. P. Dallos, "The active cochlea," Journal of Neuroscience **12**, 4575–4585 (1992).
21. R. Probst, B. L. Lonsbury-Martin, and G. K. Martin, "A review of otoacoustic emissions," The Journal of the Acoustical Society of America **89**, 2027–2067 (1991).
22. J. Pickles, *An introduction to the physiology of hearing* (Brill, 1998).
23. J. J. Guinan Jr, "Olivocochlear efferents: Their action, effects, measurement and uses, and the impact of the new conception of cochlear mechanical responses," Hearing Research **362**, 38–47 (2018).
24. J. J. Guinan Jr, "Olivocochlear efferents: anatomy, physiology, function, and the measurement of efferent effects in humans," Ear and Hearing **27**, 589–607 (2006).
25. N. R. Clark, G. J. Brown, T. Jürgens, and R. Meddis, "A frequency-selective feedback model of auditory efferent suppression and its implications for the recognition of speech in noise," The Journal of the Acoustical Society of America **132**, 1535–1541 (2012).
26. O. Etard, M. Kegler, C. Braiman, A. E. Forte, and T. Reichenbach, "Real-time decoding of selective attention from the human auditory brainstem response to continuous speech," bioRxiv , 259853 (2018).
27. M. Saiz-Alía, A. E. Forte, and T. Reichenbach, "Individual differences in the attentional modulation of the human auditory brainstem response to speech inform on speech-in-noise deficits," Scientific Reports **9**, 1–10 (2019).
28. M. Saiz-Alía, P. Miller, and T. Reichenbach, "Otoacoustic emissions evoked by the time-varying harmonic structure of speech," Eneuro **8** (2021).
29. A. Kösem and V. Van Wassenhove, "Distinct contributions of low-and high-frequency neural oscillations to speech comprehension," Language, Cognition and Neuroscience **32**, 536–544 (2017).
30. N. Ding, L. Melloni, H. Zhang, X. Tian, and D. Poeppel, "Cortical tracking of hierarchical linguistic structures in connected speech," Nature Neuroscience **19**, 158–164 (2016).
31. M. P. Broderick, A. J. Anderson, G. M. Di Liberto, M. J. Crosse, and E. C. Lalor, "Electrophysiological correlates of semantic dissimilarity reflect the comprehension of natural, narrative speech," Current Biology **28**, 803–809 (2018).
32. O. Etard and T. Reichenbach, "Neural speech tracking in the theta and in the delta frequency band differentially encode clarity and comprehension of speech in noise," Journal of Neuroscience **39**, 5750–5759 (2019).
33. N. Ding and J. Z. Simon, "Neural coding of continuous speech in auditory cortex during monaural and dichotic listening," Journal of Neurophysiology **107**, 78–89 (2012).
34. S. Srinivasan, A. Keil, K. Stratis, A. F. Osborne, C. Cerwonka, J. Wong, B. L. Rieger, V. Polcz, and D. W. Smith, "Interaural attention modulates outer hair cell function," European Journal of Neuroscience **40**, 3785–3792 (2014).
35. J. A. O'sullivan, A. J. Power, N. Mesgarani, S. Rajaram, J. J. Foxe, B. G. Shinn-Cunningham, M. Slaney, S. A. Shamma, and E. C. Lalor, "Attentional selection in a cocktail party environment can be decoded from single-trial eeg," Cerebral Cortex **25**, 1697–1706 (2015).
36. D. Jochaut, K. Lehongre, A. Saitovitch, A.-D. Devauchelle, I. Olasagasti, N. Chabane, M. Zilbovicius, and A.-L. Giraud, "Atypical coordination of cortical oscillations in response to speech in autism," Frontiers in Human Neuroscience **9**, 171 (2015).
37. M. Lizarazu, M. Lallier, M. Bourguignon, M. Carreiras, and N. Molinaro, "Impaired neural response to speech edges in dyslexia," Cortex **135**, 207–218 (2021).
38. G. M. Di Liberto, V. Peter, M. Kalashnikova, U. Goswami, D. Burnham, and E. C. Lalor, "Atypical cortical entrainment to speech in the right hemisphere underpins phonemic deficits in dyslexia," Neuroimage **175**, 70–79 (2018).
39. C. Braiman, E. A. Fridman, M. M. Conte, H. U. Voss, C. S. Reichenbach, T. Reichenbach, and N. D. Schiff, "Cortical response to the natural speech envelope correlates with neuroimaging evidence of cognition in severe brain injury," Current Biology **28**, 3833–3839 (2018).
40. A.-L. Giraud and D. Poeppel, "Cortical oscillations and speech processing: emerging computational principles and operations," Nature Neuroscience **15**, 511–517 (2012).
41. N. Ding and J. Z. Simon, "Cortical entrainment to continuous speech: functional roles and interpretations," Frontiers in Human Neuroscience **8**, 311 (2014).
42. J. E. Peelle, J. Gross, and M. H. Davis, "Phase-locked responses to speech in human auditory cortex are enhanced during comprehension," Cerebral Cortex **23**, 1378–1387 (2013).
43. G. M. Di Liberto, J. A. O'Sullivan, and E. C. Lalor, "Low-frequency cortical entrainment to speech reflects phoneme-level processing," Current Biology **25**, 2457–2465 (2015).
44. C. Micheyl and A. J. Oxenham, "Across-frequency pitch discrimination interference between complex tones containing resolved harmonics," The Journal of the Acoustical Society of America **121**, 1621–1631 (2007).
45. M. J. Polonenko and R. K. Maddox, "Exposing distinct subcortical components of the auditory brainstem response evoked by continuous naturalistic speech," Elife **10**, e62329 (2021).
46. T. Duke and F. Jülicher, "Active traveling wave in the cochlea," Physical Review Letters **90**, 158101 (2003).
47. F. Jülicher, D. Andor, and T. Duke, "Physical basis of two-tone interference in hearing," Proceedings of the National Academy of Sciences U.S.A. **98**, 9080–9085 (2001).
48. N. Ciganović, A. Wolde-Kidan, and T. Reichenbach, "Hair bundles of cochlear outer hair cells are shaped to minimize their fluid-dynamic resistance," Scientific Reports **7**, 1–9 (2017).
49. L. Sumner, J. Mestel, and T. Reichenbach, "Steady streaming as a method for drug delivery to the inner ear," Scientific Reports **11**, 1–12 (2021).
50. D. McFadden, "Sex differences in the auditory system," Developmental Neuropsychology **14**, 261–298 (1998).
51. D. McFadden, G. K. Martin, B. B. Stagner, and M. M. Maloney, "Sex differences in distortion-product and transient-evoked otoacoustic emissions compared," The Journal of the Acoustical Society of America **125**, 239–246 (2009).

Does Endolymphatic Hydrops Shift the Cochlear Tonotopic Map?

Samantha Stiepan[a)], Christopher A. Shera, and Carolina Abdala

Auditory Research Center, Caruso Department of Otolaryngology,
University of Southern California, 1640 Marengo St, Los Angeles, CA, 90033

[a)] Corresponding author: Samantha.Stiepan@usc.edu

Abstract. The cochlear tonotopic map determines where along the basilar membrane traveling waves of different frequencies peak. Endolymphatic hydrops has been hypothesized to shift the tonotopic map by altering the stiffness of the cochlear partition, especially in the apex. In this exploratory study performed in a handful of normal and hydropic ears, we report preliminary measurements of interaural differences assayed using behavioral pitch-matching supplemented by measurements of reflection otoacoustic-emission phase-gradient delays.

INTRODUCTION

The stiffness gradient along the basilar membrane determines where the traveling wave peaks. In ears with endolymphatic hydrops, the increased fluid volume is thought to stiffen the cochlear partition, particularly near the apex where the hydrops is more pronounced [1]. This increased stiffness alters the motion of the traveling wave along the cochlear partition by increasing the local characteristic frequency, thus altering the tonotopic map.

Binaural diplacusis is the perceptual anomaly where the perceived pitch of a tone is different between ears. Diplacusis has been observed sporadically and anecdotally in ears with endolymphatic hydrops and attributed to a change in the cochlear frequency-place map of the affected ear [1-4]. If the traveling wave peaks in a different physical place along the cochlear length in each ear, the frequency coded (and hence the pitch perceived) might be expected to differ, at least if the perceived pitch is dominated by the cochlear place code. In normal ears with a frequency range of ~10 octaves, interaural pitch differences are small, typically less than one semitone [5–7], indicating that the interaural pitch agreement is better than 1% of the audible frequency range. The theories of Tonndorf [1] predict that the stiffened basilar membrane shifts the site of the traveling-wave peak, producing diplacusis in ears with unilateral endolymphatic hydrops. Recent work in guinea pigs with induced endolymphatic hydrops suggests that a perceptual shift in pitch between ears could be produced by mechanical changes in the frequency-place map [8]. Guinan et al. [8] estimated apical shifts of 0.5 to 2 octaves in the affected ear of guinea pigs with induced endolymphatic hydrops.

To determine if there is a corresponding shift in the tonotopic organization of humans with endolymphatic hydrops, we conducted an interaural pitch-matching task for diplacusis in a small group of subjects diagnosed with endolymphatic hydrops. Figure 1 provides a schematic of inter-aural pitch matching in (A) one normal-hearing subject and (B) one subject with unilateral endolymphatic hydrops who shows a pitch mismatch between ears. Reflection-source otoacoustic emissions (i.e., stimulus-frequency otoacoustic emissions, or SFOAEs) were also measured in each subject, with the eventual goal of comparing perceptual and otoacoustic results in larger groups with and without endolymphatic hydrops. No such comparison was conducted in this preliminary study.

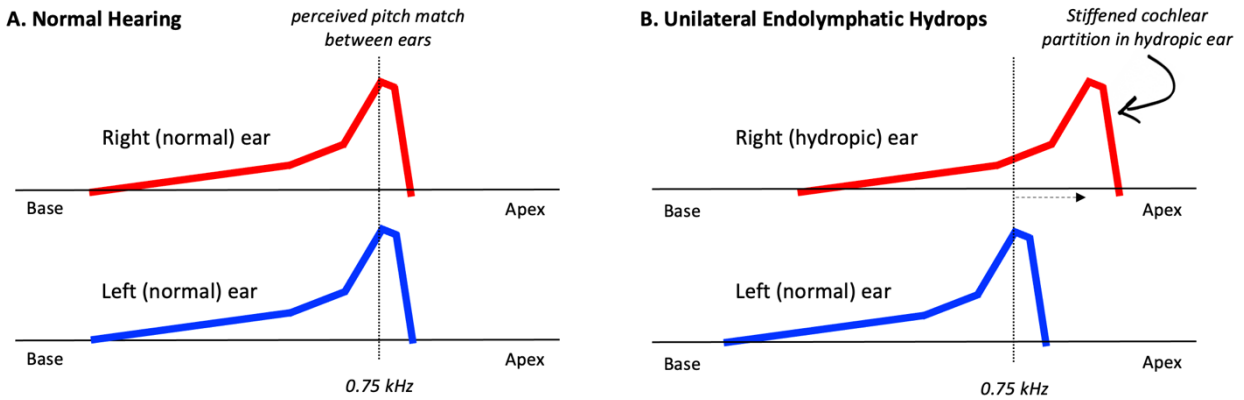


FIGURE 1. A schematic illustrating binaural diplacusis produced by an altered tonotopic map. An unspiraled basilar membrane and traveling wave is shown in (A) a normal hearer and (B) a subject with unilateral endolymphatic hydrops. The figure illustrates the hypothesized interaural mismatch in tonotopic location produced by unilateral endolymphatic hydrops. In a subject with normal hearing, a 0.75 kHz tone peaks at the same location along the basilar membrane in both ears. In a subject with unilateral endolymphatic hydrops, however, the stiffened cochlear partition increases the local characteristic frequency in the hydropic ear, shifting the tonotopic map apically along the basilar membrane. When the perceived pitch is determined primarily by the cochlear place code, the subject with endolymphatic hydrops will experience binaural diplacusis (an interaural mismatch in the perceived pitch of the tone). For subjects with endolymphatic hydrops, we expect the perceived pitch will be lower when the test tone is presented to the hydropic ear.

METHODS

Subjects were four adults with endolymphatic hydrops (three with unilateral endolymphatic hydrops and one with bilateral endolymphatic hydrops), ranging in age from 30 to 58 years (mean 46 years), and four normal-hearing adults, ranging in age from 27 to 36 years (mean 32 years). One of the subjects with unilateral endolymphatic hydrops was unable to perform the interaural pitch matching task and was eliminated from the study. Normal hearing was defined as audiometric thresholds ≤ 15 dB HL at octave and inter-octave frequencies from 0.5 to 8 kHz.

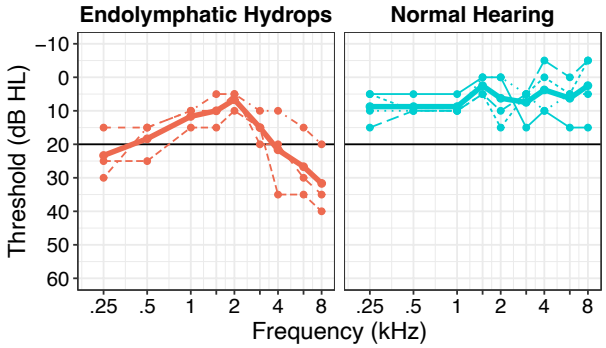


FIGURE 2. Audiometric thresholds for normal-hearing (cyan) and hydropic (coral) subjects. The thick solid line represents the mean for the group while thin lines show individual ears.

Audiometric thresholds were not controlled for in the group with endolymphatic hydrops; one of these individuals had hearing within normal limits while others showed a mild-to-moderate low-frequency sensorineural hearing loss (see Fig. 2). In the low frequencies (< 1.5 kHz), the mean audiometric threshold for the group with endolymphatic hydrops was 10 dB worse than that of the group with normal hearing. The only inclusion criterion was a medical diagnosis of either Ménière's disease or endolymphatic hydrops by an otologist. These diagnoses were made mostly from symptoms (e.g., vertigo, aural fullness, rushing/roaring tinnitus, fluctuating low-frequency hearing loss) and patient history.

For the interaural pitch comparison task, we followed the protocol of Colin et al. [7]. On each of the 85 trials, the participant heard two pure tones presented sequentially. The first tone was always the reference tone, which was presented to the healthy (non-hydropic) ear when possible. The reference tone was followed by a test tone of variable frequency presented to the affected ear. The subject was instructed to pick which of the two tones was higher in pitch. The tones were presented at 50 dB SL in each ear, which was intended to ensure audibility and compensate for differences in hearing threshold between the groups. The frequency of the reference tone was picked to coincide with peaks in SFOAE spectra which were recorded prior to the psychoacoustic test. The test tone was varied over a full half-octave range (+/- one quarter-octave re: reference tone) in $^{1}/_{32}$ octave steps. The responses to this task were used to create a logistic regression function and estimate the interaural pitch shift, as shown in Fig. 3 for one normal hearer and one individual with unilateral endolymphatic hydrops. The interaural pitch shift is defined as the frequency difference (test ear minus reference ear)

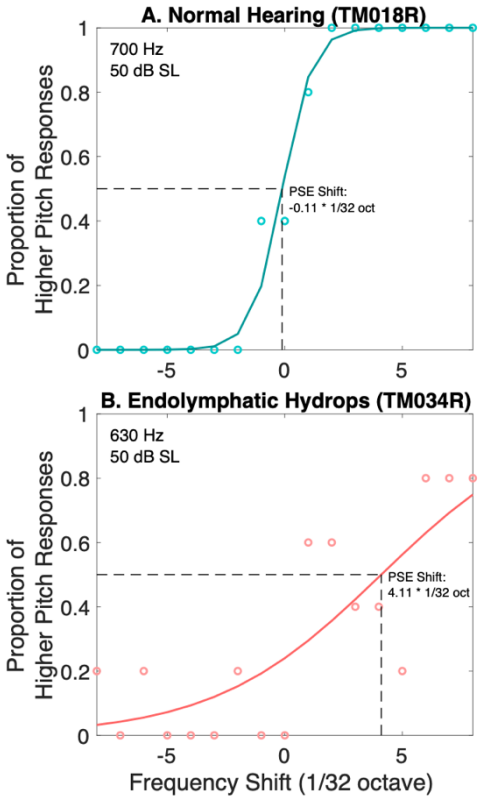
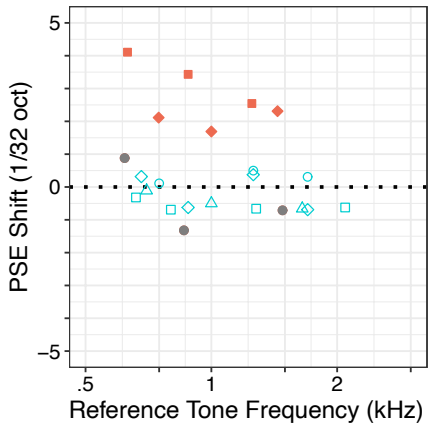


FIGURE 3. The psychometric function derived from the interaural pitch-comparison data allows one to identify the "point of subjective equality" (PSE) where the test and reference tones are perceived as having the same pitch. Panels A and B show psychometric functions from a normal-hearing and hydropic subject, respectively. Dashed lines indicate the interaural pitch shift at the PSE, where the probability is 0.5.



at the "point of subjective equality" (PSE) where the two tones are perceived as having the same pitch. Prior to testing, a brief training set was presented to teach each participant the task.

As stated above, an SFOAE protocol was conducted prior to pitch matching. SFOAEs were evoked with tones swept downward at 1 octave/second across a 4-octave range (0.5 to 8 kHz). Segments of the frequency span were stacked (with 0.1 octave overlap) and presented concurrently (as described in Abdala et al. [10]). SFOAEs were measured at three stimulus levels, 40, 50, and 60 dB FPL, using a modified, interleaved suppression paradigm [9-11]. SFOAE level and phase were derived from the recorded ear-canal signal using a least-squares fitting procedure [12-14]. The SFOAE was filtered with a continuous wavelet transform to eliminate short-latency energy associated with stimulus artifact and long-latency energy associated with the presence of multiple internal reflections [15]. SFOAE phase was converted to group (or 'phase-gradient') delay by calculating the negative of the phase slope in milliseconds. Delays were calculated across frequency and stimulus level and compared between groups with normal hearing and endolymphatic hydrops (for more details about the SFOAE paradigm, please see Abdala et al. [10]).

RESULTS

Figure 4 summarizes results for the interaural pitch-matching task displaying pitch shifts as a function of reference-tone frequency. Interaural pitch shifts for the subjects with unilateral endolymphatic hydrops (coral) ranged from 5% to 13% of an octave, while normal-hearing subjects (cyan) show interaural pitch shifts hovering around zero (-2% to 1.5% of an octave). Interaural pitch shifts for one subject with bilateral endolymphatic hydrops (in gray) were most similar those of the normal-hearing subjects as might be expected if the two ears were similarly affected by a hydropic state. The pitch shifts observed in the two subjects with unilateral endolymphatic hydrops are consistent with an apically shifted traveling wave in the hydropic ear.

Figure 5 displays loess trend lines fit to SFOAE phase-gradient delays as a function of frequency for the groups with endolymphatic hydrops and normal hearing. Line width denotes stimulus level: 40, 50, and 60 dB FPL, from thin to thick. Both normal-hearing individuals and those with endolymphatic hydrops showed the characteristic level dependence of reflection-emission delays: the lower the stimulus level, the steeper the phase-gradient and longer the delay [10,16-17].

FIGURE 4 (left). Pitch shift as a function of reference-tone frequency for four normal-hearing subjects (cyan) and two subjects with unilateral endolymphatic hydrops (coral). The dashed line indicates no difference in pitch was perceived between ears. A positive pitch shift is consistent with a traveling wave shifted apically (re: healthy ear). The gray symbols are from a subject with bilateral hydrops.

DISCUSSION

In this preliminary study, we tested a handful of ears with diagnosed endolymphatic hydrops and a comparable group of individuals with no ear disease, using an interaural pitch-matching task and SFOAEs. Endolymphatic hydrops produces an increased volume of endolymph in the scala media, which is hypothesized to stiffen the cochlear partition [4,18-19]. An increase in basilar-membrane stiffness would be expected to increase the local resonant frequency [20], altering the normal tonotopic map. The expected shift direction was consistent with how our subjects with endolymphatic hydrops matched two tones presented sequentially to the affected and healthy ears. These data are sparse and preliminary; they are but the beginnings of a study to define interaural pitch-matching in hydropic ears more thoroughly. Nevertheless, the initial results show support for our hypothesis.

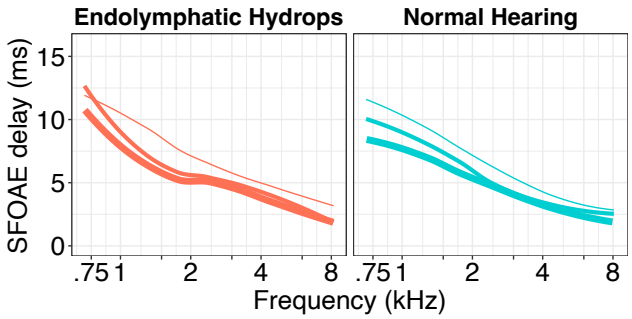


FIGURE 5. Loess trend lines fit to SFOAE phase-gradient delays from subjects with endolymphatic hydrops (coral) and normal hearing (cyan) at 40 (thinnest line), 50, and 60 dB FPL. Both groups show characteristic level-dependence of the SFOAE delay.

In individuals with unilateral endolymphatic hydrops, our perceptual findings are consistent with diplacusis produced by an altered tonotopic map (see Fig. 4). The direction of the interaural pitch shift was consistent with an apically-shifted traveling wave in the hydropic ear. Although endolymphatic hydrops typically presents unilaterally, especially in early disease stages, the disease often becomes bilateral as it progresses. Interaural pitch comparisons for the subject with bilateral endolymphatic hydrops were more similar to the normal-hearing subjects than to those with unilateral endolymphatic hydrops. This is expected since shifts in the location of the traveling-wave peak will cause a perceptual shift only if the two ears are different. This subject may have a similar degree of hydrops (and stiffening of the cochlear partition) in right and left ears. A more substantive study of diplacusis in a larger group of individuals with endolymphatic hydrops may establish the diagnostic value of evaluating patients using an interaural pitch-matching task to characterize the laterality of the disease.

Although SFOAE delays were measured in both hydropic and normal-hearing ears, our current subject pool is too small, and SFOAE delays too variable across individuals, to allow meaningful comparisons. Future work will require a larger study in order to reliably compare possible differences across groups (and between ears in the subjects with endolymphatic hydrops).

Other auditory pathologies may also produce cochlear frequency-place map alterations and diplacusis. Indeed, diplacusis has been reported in people with sensorineural hearing loss, but the tone pitch is most often perceived at a higher frequency in the ear with the poorest hearing [6-7,18,21]. This finding is consistent with the notion that higher sound levels (required for stimulation of the poorer ear) can cause a basalward shift in cochlear place cues.

In conclusion, our preliminary findings are consistent with the hypothesis that endolymphatic hydrops can cause an altered tonotopic map, which can lead to abnormal interaural pitch perception. Our results support the idea that cochlear tonotopic-map changes would be valuable to evaluate when assessing the perceptual consequence of hearing pathologies and considering treatments.

ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Health (DC03552, DC018307, and T32 DC009975). We thank Raymond Goldsworthy for help with the interaural pitch-matching task and John Oghalai for verifying the diagnoses of endolymphatic hydrops. We also thank Ping Luo for algorithm development and technical assistance.

REFERENCES

1. Tonndorf, J., 1976. Endolymphatic Hydrops: Mechanical Causes of Hearing Loss. *Archives of oto-rhino-laryngology*, *212*(4), pp.293-299.
2. Shambaugh, G.E., 1940. Diplacusis: A Localizing Symptom of Disease of the Organ of Corti: Theoretic Considerations, Clinical Observations and Practical Application. *Archives of Otolaryngology*, *31*(1), pp.160-184.
3. Tonndorf, J., 1957. LIX The Mechanism of Hearing Loss in Early Cases of Endolymphatic Hydrops. *Annals of Otology, Rhinology & Laryngology*, *66*(3), pp.766-784.
4. Gluth, M.B., 2020. On the Relationship Between Menière's Disease and Endolymphatic Hydrops. *Otology & Neurotology*, *41*(2), pp.242-249.
5. Albers, G.D. and Wilson, W.H., 1968. Diplacusis: I. Historical Review. *Archives of Otolaryngology*, *87*(6), pp.601-603.
6. Jonas Brännström, K. and Grenner, J., 2008. Long-Term Measurement of Binaural Intensity Matches and Pitch Matches. I. Normal Hearing. *International Journal of Audiology*, *47*(2), pp.59-66.
7. Colin, D., Micheyl, C., Girod, A., Truy, E. and Gallégo, S., 2016. Binaural Diplacusis and its Relationship with Hearing-Threshold Asymmetry. *PLoS One*, *11*(8), p.e0159975.
8. Guinan, J.J., Lefler, S.M., Buchman, C.A., Goodman, S.S. and Lichtenhan, J.T., 2021. Altered Mapping of Sound Frequency to Cochlear Place in Ears with Endolymphatic Hydrops Provide Insight into the Pitch Anomaly of Diplacusis. *Scientific reports*, *11*(1), pp.1-9.
9. Shera, C.A. and Guinan Jr, J.J., 1999. Evoked Otoacoustic Emissions Arise by Two Fundamentally Different Mechanisms: A Taxonomy for Mammalian OAEs. *The Journal of the Acoustical Society of America*, *105*(2), pp.782-798.
10. Abdala, C., Guardia, Y.C. and Shera, C.A., 2018. Swept-Tone Stimulus-Frequency Otoacoustic Emissions: Normative Data and Methodological Considerations. *The Journal of the Acoustical Society of America*, *143*(1), pp.181-192.
11. Abdala, C., Ortmann, A.J. and Shera, C.A., 2018. Reflection-and Distortion-Source Otoacoustic Emissions: Evidence for Increased Irregularity in the Human Cochlea During Aging. *Journal of the Association for Research in Otolaryngology*, *19*(5), pp.493-510.
12. Long, G.R., Talmadge, C.L. and Lee, J., 2008. Measuring Distortion Product Otoacoustic Emissions Using Continuously Sweeping Primaries. *The Journal of the Acoustical Society of America*, *124*(3), pp.1613-1626.
13. Kalluri, R. and Shera, C.A., 2013. Measuring Stimulus-Frequency Otoacoustic Emissions Using Swept Tones. *The Journal of the Acoustical Society of America*, *134*(1), pp.356-368.
14. Abdala, C., Luo, P. and Shera, C.A., 2015. Optimizing Swept-Tone Protocols for Recording Distortion-Product Otoacoustic Emissions in Adults and Newborns. *The Journal of the Acoustical Society of America*, *138*(6), pp.3785-3799.
15. Moleti, A., Longo, F. and Sisto, R., 2012. Time-Frequency Domain Filtering of Evoked Otoacoustic Emissions. *The Journal of the Acoustical Society of America*, *132*(4), pp.2455-2467.
16. Schairer, K.S., Ellison, J.C., Fitzpatrick, D. and Keefe, D.H., 2006. Use of Stimulus-Frequency Otoacoustic Emission Latency and Level to Investigate Cochlear Mechanics in Human Ears. *The Journal of the Acoustical Society of America*, *120*(2), pp.901-914.
17. Sisto, R. and Moleti, A., 2007. Transient Evoked Otoacoustic Emission Latency and Cochlear Tuning at Different Stimulus Levels. *The Journal of the Acoustical Society of America*, *122*(4), pp.2183-2190.
18. Brookes, G.B. and Parikh, A., 1995. Meniere's Disease: An Evaluation of Conventional Audiological Diagnostic Methods. In *Ménière's disease. Proceedings of the 16th Danavox Symposium, Sept. 19-22* (pp. 157-172).
19. Allen, G.W., 1983. Clinical Implications of Experiments on Alteration of the Labyrinthine Fluid Pressures. *Otolaryngologic Clinics of North America*, *16*(1), pp.3-19.
20. Robles, L. and Ruggero, M.A., 2001. Mechanics of the Mammalian Cochlea. *Physiological reviews*, *81*(3), pp.1305-1352.
21. Webster, J.C. and Schubert, E.D., 1954. Pitch Shifts Accompanying Certain Auditory Threshold Shifts. *The Journal of the Acoustical Society of America*, *26*(5), pp.754-758.

Transverse-Longitudinal Structure Registration and Vibration Measurement via Optical Coherence Tomography

Brian L. Frost,[1, a)] C. Elliott Strimbu,[2, b)] and Elizabeth S. Olson[2, 3, c)]

[1)] *Department of Electrical Engineering, Columbia University, New York, NY, USA*
[2)] *Department of Otolaryngology Head and Neck Surgery, Columbia University, New York, NY, USA*
[3)] *Department of Biomedical Engineering, Columbia University, New York, NY, USA*

[a)] *Corresponding author: b.frost@columbia.edu*
[b)] *Electronic mail: ces2243@cumc.columbia.edu*
[c)] *Electronic mail: eao2004@columbia.edu*

Abstract. Intra-organ of Corti displacement measurements made via optical coherence tomography have provided significant information about cochlear micromechanics in recent years. However, several ambiguities inherent to this modality have complicated interpretation of these measurements. For one, optical coherence tomography measures the one-dimensional projection onto the optical axis of a three-dimensional motion. Also, the optical axis may make a substantial angle with the basilar membrane normal, meaning that structures along the optic axis, measured in a single measurement, may lie in different tonotopic cross-sections. We have developed a method that accounts for both of these ambiguities, to reconstruct the two-dimensional longitudinal-transverse components of displacements of structures within the organ of Corti. This is performed by taking data at multiple longitudinal positions at two viewing angles, without any *a priori* knowledge of the measurement locations or viewing angles. We present a sample data set in which we have applied this program to reconstruct the transverse-longitudinal motion of the base of the outer hair cells in the base of the sensitive gerbil cochlea. The results reinforce the importance of accounting for viewing angle when analyzing and reporting vibration results.

INTRODUCTION

Historically, the *in vivo* study of basal cochlear mechanics was limited to measurements of the basilar membrane (BM). The advent of optical coherence tomography (OCT) in the last decade has allowed for vibrometry at a depth, facilitating the study of intra-organ of Corti complex (OCC) motions. Of particular interest is the motion of the electromotile outer hair cells (OHCs), which play an important role in amplifying vibration responses and improving the range of sound-pressure levels (SPLs) over which hearing operates.

Several issues complicate the interpretation of OCT measurements. Firstly, the motion of the OHC region is not uniform. The OHCs are 40 μm long and 10 μm wide, and come in rows of three per longitudinal cross-section. The apical surface of the OHCs, called the reticular lamina (RL), moves differently from the basal surface attached to the Deiters' cells as the OHC compresses and expands due to electromotility [1]. Moreover, the OHCs within a row move differently from each other [2].

A second complication is that measurements are generally taken at an angle with respect to the transverse anatomical orientation of the cochlea, and this angle is not known *a priori*. This introduces two ambiguities: (1) displacements measured via OCT are projections of the three-dimensional motion onto an unknown axis, and (2) measured points at the OHC and BM within a single optical axis measurement (A-scan) will in general lie in different longitudinal cross-sections. The first of these ambiguities was discussed by Cooper et al [3]. Their measurements showed that the phase difference between OHC and BM motion relied heavily on the viewing angle. The second of these ambiguities was discussed by Frost et al [4], wherein we developed a program that measured the relative anatomical distances between structures imaged with OCT. We used this program to measure displacement of the OHC and BM at the same cross-section, and showed that this correction revealed up to 1/4 cycle difference in OHC-BM displacement phase when displacement-accounted (single cross-section) results were compared to single-measurement (two different cross-section) results.

With these complications in mind, it is difficult to fully interpret OCT measurements of OHC displacement that are reported without the viewing angle specified. The group of Ren has achieved BM and OHC-region displacement measurements taken at a purely transverse angle, thus measuring purely transverse motion in a single tonotopic cross section, both in gerbil and mouse [5, 6]. The instrument they use is similar to

OCT, but does not provide imaging. The OHC-region measurement was reasoned to be the RL, but this identification is not ironclad. They found that RL displacement phase led BM at low frequencies, and this lead decreased monotonically until the structures were in phase at about 0.8 of the best frequency (BF). After this zero crossing, the RL re BM phase continued its monotonic decrease, with the RL lagging the BM at near- and supra-BF frequencies. Our displacement-accounted data of OHC-region and BM in the same tonotopic cross-section taken at a viewing angle with a significant longitudinal component found a phase difference of a different character. OHC led BM across frequency, including at high frequencies – where Ren et al saw an 80°*lag*, we saw a 90°*lead*. Given the ambiguities and technical variations noted above, the difference in observations is not surprising.

In order to address the experimental variations and ambiguities, and thereby enhance the value of intra-OCC motion measurements, we have developed a method to isolate the transverse and longitudinal components of motion of structures within the OCC, wherein we account both for the longitudinal distance between structures and the projection incurred by the viewing angle. Similar to the work of Lee et al, we do so by taking measurements at multiple viewing angles [7]. Our method requires no *a priori* knowledge of either measurement angle, or of the positions of the structures being measured. We determine a linear approximation of the longitudinal direction at acquisition time, followed by *post hoc* registration of structures measured at several longitudinal locations and two viewing angles. We then analyze the measured displacements to achieve a reconstructed longitudinal-transverse profile of a structure's motion.

We present the method, an analysis of its fidelity with respect to viewing angle, and *in vivo* data from the base of the gerbil cochlea in which the longitudinal-transverse motion at the basal region of several OHCs has been reconstructed. In this sample data set, the phase of our reconstructed transverse displacements were similar to the pure-transverse RL displacements measured by Ren. Also, the magnitude of longitudinal (base-to-apex) OHC displacement was generally smaller than that of transverse OHC displacement and $\sim$ 180°out of phase with transverse (towards scala vestibuli) OHC displacement, across frequency.

METHODS

Acquisition

The acquisition process follows a few steps: (1) prior to acquisition, ensure the BM looks as horizontal as possible in two orienting B-Scans, (2) use these B-scans to determine the approximate longitudinal direction, (3) decide on a structure-category of interest (for example, base of the OHC) and take A-scan displacement measurements including that structure-category at multiple longitudinal locations, (4) rotate the preparation and again find two new orienting B-scans with the BM appearing horizontal, (5) again, find the approximate longitudinal direction and take A-scan measurements of the same structure-category of interest at multiple longitudinal locations.

Orienting the preparation so that the BM appears horizontal in each cross-section ensures that the radial component of the optic axis is approximately zero. This simplifies the problem to two dimensions – transverse and longitudinal.

Determining the approximate longitudinal direction works via a linear approximation of the cochlea's anatomical coordinates, similar to the planar approximation we employed in [4]. As we only need the longitudinal direction for acquisition, this is a simpler process – we use ThorImage (the imaging program of the Thorlabs Telesto OCT) and locate a landmark in our orienting B-scans; in the example in Fig. 1 we use the thickest part of the BM. We record the optical coordinates at this position in the two B-scans, $p_1 = (x_1, y_1, z_1)$ and $p_2 = (x_2, y_2, z_2)$. The difference between these points is a linear approximation of a vector in the longitudinal direction, so the unit longitudinal vector is

$$l = \frac{p_2 - p_1}{|p_2 - p_1|}. \tag{1}$$

This process is illustrated in panels A and B of Fig. 1.

We can find points of measurement for some structure of interest, say the base of the OHC, at fixed increments along the longitudinal direction. If we start at a point q_0 containing that structure, and want to

measure N points over longitudinal distance L, we measure at

$$q_n = q_0 + \frac{nL}{N} l,\ n = 0, 1, \ldots, N-1. \tag{2}$$

This is illustrated in panel C of Fig. 1. Note that q_0 need not be one of p_1 or p_2; in this example p_1 and p_2 are within the BM and the q positions are at the base of the OHCs.

We take volume scans after each run so that outside of the time pressure of an experiment, we can apply our more complex orientation program [4]. In doing so, we can assess the accuracy of our assumption that the radial component of our measurement axis is 0, as well as the direction of the longitudinal vector.

Registration

As we have made sure to remove the radial component of motion from our measurements, we can assume the optical z axis is comprised of only longitudinal (l) and transverse (t) components. We write the optical axis' unit vector as a two-dimensional vector in anatomical coordinates, $z = (z_l, z_t)$.

Eqn. 1 is used in the acquisition step to find the longitudinal vector l, which has optical x, y and z components. The z component of this vector represents the amount of longitudinal motion that is projected onto the optical axis, z_l. To find the t component of z, we need only to recall that z is a unit vector, so that $z_l^2 + z_t^2 = 1$. Using the notation from Eqn. 1, we have

$$z_l = \frac{z_2 - z_1}{|p_2 - p_1|}, \tag{3}$$

$$z_t = \sqrt{1 - z_l^2}. \tag{4}$$

This is illustrated in panel C of Fig. 1.

Knowing this, we can relate the OHC and BM longitudinal locations within a single measurement. If the structures along a single measurement axis are spaced Δz apart, then the OHCs lie $\Delta l = z_l \Delta z$ apical of the BM. The measured BM Δl apical of this measurement is thereby in the same longitudinal cross-section as the OHC in the first measurement. We call these BM and OHC measurements *aligned* to one another. At each measurement angle, we compose a list of all aligned OHC and BM measurements. This process is illustrated in panels D and E of Fig. 1.

To register points to one another between viewing angles, we use the phase of BM motion. By matching the phase responses of BM measurements taken at different viewing angles, we can register the cross-sections in which these BM motions were measured. This operates under the assumption that BM motion is entirely transverse, so that its phase response at each cross-section will be the same at each viewing angle. Having registered BM points to one another, we can then consult our list of aligned BM-OHC pairs. The OHCs aligned to registered BM positions are also registered, allowing us to isolate the same OHCs at different viewing angles.

Reconstruction

Each OCT measurement is a projection of a true 3-D motion onto the optical z axis. In the current context, we have made efforts to eliminate the representation of radial motion in our projection, so that the problem can be framed as the projection of a 2-D longitudinal-transverse *true motion* d onto a 2-D transverse-longitudinal z axis, forming the *projected motion* δ.

Above we described the method by which we determine the z_l and z_t components of the optical axis in each experiment. The projection onto this axis is given by the dot product

$$\delta = z \cdot d = \begin{pmatrix} z_l & z_t \end{pmatrix} \begin{pmatrix} d_l \\ d_t \end{pmatrix}. \tag{5}$$

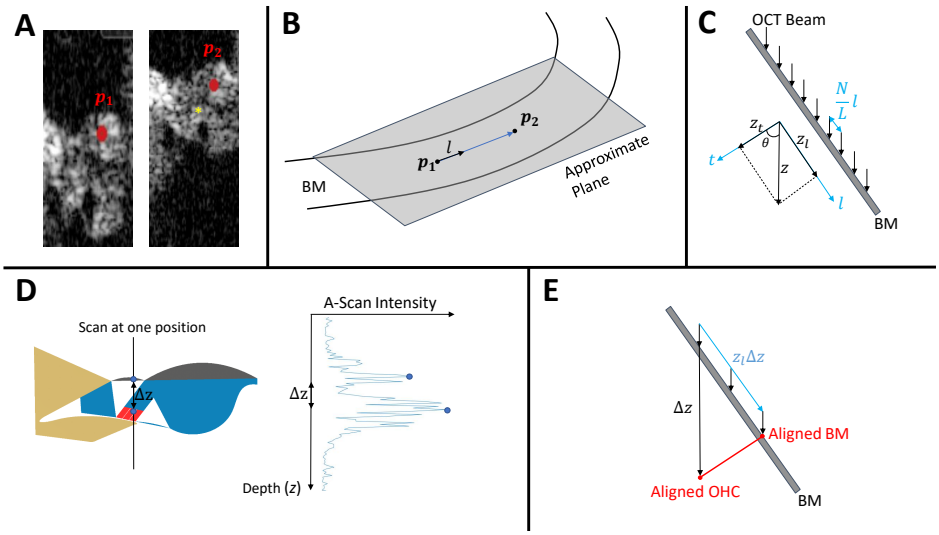


FIGURE 1. A – Two parallel B-Scans from a single volume, about 50 μm apart, on which we have marked points p_1 and p_2 used to determine the longitudinal direction l. We have selected the landmark to be the BM at its widest point. We have ensured that the BM appears approximately horizontal in each B-Scan, so as to remove radial displacement contributions. Included also is a yellow star at the OHC base, which could be used as q_0 to measure OHC base at multiple longituinal cross-sections once the l direction has been determined. **B** – A cartoon of the BM, with true longitudinal axis varying in space. We have approximated this axis by a line passing through both p_1 and p_2. **C** – A cartoon of the BM projected onto the longitudinal-transverse plane, in which our A-Scans lie. We take A-Scans along the z axis at multiple evenly spaced longitudinal positions. We can determine the anatomical components of the z vector via geometry, knowing l and knowing that z has unit length. **D** – A cartoon of the organ of Corti complex and an experimentally acquired A-Scan. We measure the distance between OHC base and BM in the A-Scan, Δz, to determine the longitudinal displacement between these structures. Gray is the BM, with the fluid compartment scala tympani above. Blue is the organ of Corti with red the OHCs. Tan is the spiral limbus and tectorial membrane. Scala vestibuli is the fluid compartment below to OCC. **E** – A cartoon of the BM in the longitudinal-transverse plane similar to **C**. The OHC at a measurement position is $z_l\Delta z$ apical of the BM at that same position. This OHC's *aligned* BM point is that which is measured at this same longitudinal position.

At each angle, if we are truly measuring at one position, d will remain the same and z will change. If we take measurements at two angles, we form the system of equations:

$$\begin{pmatrix}\delta_1\\ \delta_2\end{pmatrix} = \begin{pmatrix}l_1 & t_1\\ l_2 & t_2\end{pmatrix}\begin{pmatrix}d_l\\ d_t\end{pmatrix}, \tag{6}$$

where the rows of the matrix are the z axes corresponding to each angle, and δ_i is the projection measured at the i^{th} angle.

We measure δ_i and determine the l and t components of the z axis as described above. We want to reconstruct d, which can be done by inverting the matrix in Equation 6. This is possible if and only if the rows of the matrix are linearly independent, i.e. if the measurement axes are not colinear. As long as we measure at two sufficiently distinct (to be quantified shortly) angles, we can reconstruct d by performing the matrix inverse:

$$\begin{pmatrix}d_l\\ d_t\end{pmatrix} = \frac{1}{l_1t_2 - l_2t_1}\begin{pmatrix}t_2 & -t_1\\ -l_2 & l_1\end{pmatrix}\begin{pmatrix}\delta_1\\ \delta_2\end{pmatrix}. \tag{7}$$

Achieving measurements of the same structures at different angles is constrained by the preparation. In practice, a 15 degree rotation is tractable in our preparation, but significantly larger angles are not consistently achievable. The precision of our reconstruction can be found via the *condition number* κ of our projection matrix in Eqn. 6. The condition number is defined as

$$\kappa(A) = \frac{|\sigma_{max}|}{|\sigma_{min}|}, \tag{8}$$

Angular deviation	Condition number κ
20°	5.67
15°	7.60
10°	11.43
5°	22.90
1°	114.59

TABLE I. A table of condition numbers for the projection matrix at possible measurement angular deviations.

where σ_{max} and σ_{min} are the maximum and minimum singular values (similar to eigenvalues) of a given matrix A. The condition number of a matrix represents how "well-posed" a system of equations is, i.e. how close to singular the matrix is. A matrix with a large condition number will amplify noise significantly more than a matrix with a small condition number. The "rule of thumb" is that SNR is reduced by a factor of about κ. Note that a matrix and its inverse have the same condition number.

Table I shows condition numbers for some possible angular deviations. Condition number does not depend on the absolute angle, but only the absolute difference between the two measurement angles. Our usual precision is on the order of 0.1 nm, and our signal at higher dB SPL is in the range of 1-10 nm. The realized 15-degree angle will lead to an eight-fold increase in the noise level, which means our results are accurate to about 0.8 nm.

RESULTS

We have employed the method described above in the base of the gerbil cochlea, with measurements taken through the round window membrane at angles of approximately 45° and 60° relative to the transverse direction. Going from one to the other angle was accomplished with a goniometer that serves as the head-holder. The base of the OHC was used as the structure of interest. Using the ThorImage positioning system and the method outlined above, we took data at 11 positions spaced 15 μm apart longitudinally. We stimulated the ear with one-second, 15-frequency multitone stimuli at 80 dB SPL. We did this at both angles.

As we stepped longitudinally along the cochlea, we expected to see the phase response at the BM vary along with the tonotopic cross-section. This behavior is displayed in Fig. 2, providing evidence that our acquisition method is truly stepping along the tonotopic axis.

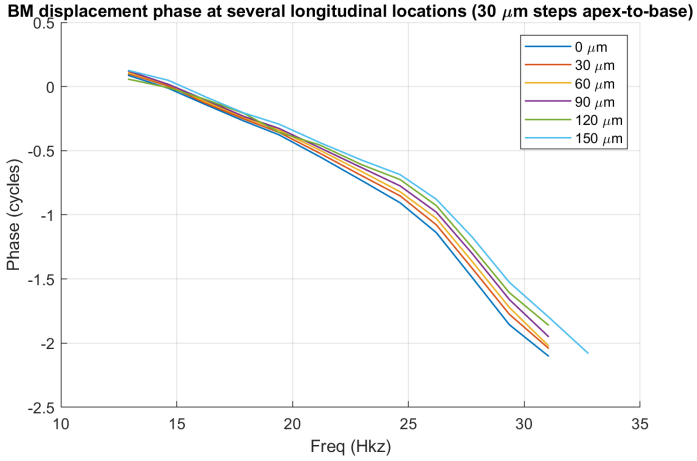


FIGURE 2. Phase of BM displacement measured in response to 80 dB SPL 15-frequency multitone stimulus, taken at several points along the longitudinal axis of the cochlea. The OCT beam is oriented at 45° to the BM, and we have stepped 15 μm at a time along the cochlea over a 150 μm range. We display every other step here to make clear the travelling wave pattern of the displacement as we step from apex to base.

As for the registration process, Fig. 3 shows matched BM phases taken at two different angles. We did not know the overlap of the measured regions at each angle *a priori* and we found a 60 μm region of longitudinal overlap *post hoc*, five positions spaced by 15 μm. Two of these matched positions are shown; these two were spaced by 45 μm.

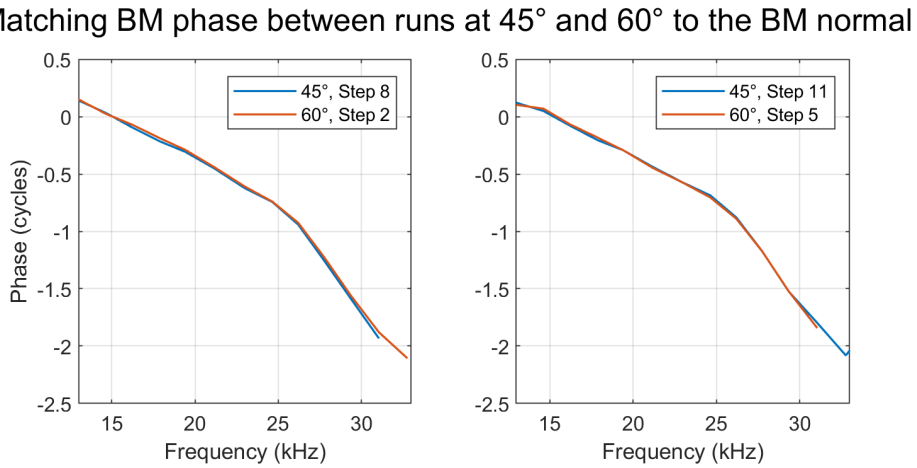


FIGURE 3. Matched BM phase plots used for registering locations between measurement angles. Phase was matched at five positions; two are shown. Each "step" was measured 15 μm basal of the previous step, and these two subplots show positions 45 μm apart from one another, with the panel on the right basal of that on the left.

Applying our reconstruction method to the associated aligned OHCs we arrive at the two-dimensional frequency responses presented in Fig. 4. At all five registered locations (two are shown), transverse OHC motion was significantly larger than longitudinal OHC motion across the bulk of the frequency range at 80 dB SPL. Reconstructed transverse OHC displacement phase underwent the linear lead-to-lag transition measured by Ren and colleagues, with zero-crossing near 0.8BF [5, 6].

The longitudinal displacement phase led transverse displacement phase by ∼180°. (Recall the transverse axis points toward scala vestibuli, the longitudinal axis points toward the base.) With respect to our coordinate system, this means that towards-scala tympani transverse displacement is approximately in phase with basal-to-apical longitudinal displacement.

CONCLUSION

We have developed and presented the application of a method by which longitudinal-transverse 2-D intra-OCC displacements can be reconstructed from 1-D OCT measurements taken at two measurement angles, without any *a priori* knowledge of the measurement angles or measurement locations. We have presented the application of this method to reconstructing the motion of the base of OHCs across a 45 μm longitudinal region of the base of the cochlea. This method can be used to reconstruct the 2-D motion of any structure in the OCC. Measurements of this type help to resolve apparent discrepancies between groups, contextualize 1-D measurements and serve as a step towards the goal of reconstructing total 3-D intra-OCC displacements.

A 2-D understanding of intra-OCC motions provides significantly more information about the mechanical operations than single-dimensional measurements. In particular, 1-D measurements of relative BM-OHC displacement phase taken with a large longitudinal component, even when the longitudinal distance between the structures was accounted for, did not show the ∼0.8 BF lead-to-lag transition of transverse OHC phase re BM that was observed when measuring purely transverse motion, as seen both in our reconstruction and in [5]. As 0.8 BF is approximately the onset frequency of BM displacement nonlinearity, this transition may be key for understanding the operation of the cochlear amplifier. This is one example of the significant features of motion that are revealed by multi-dimensional measurements of OCC vibration.

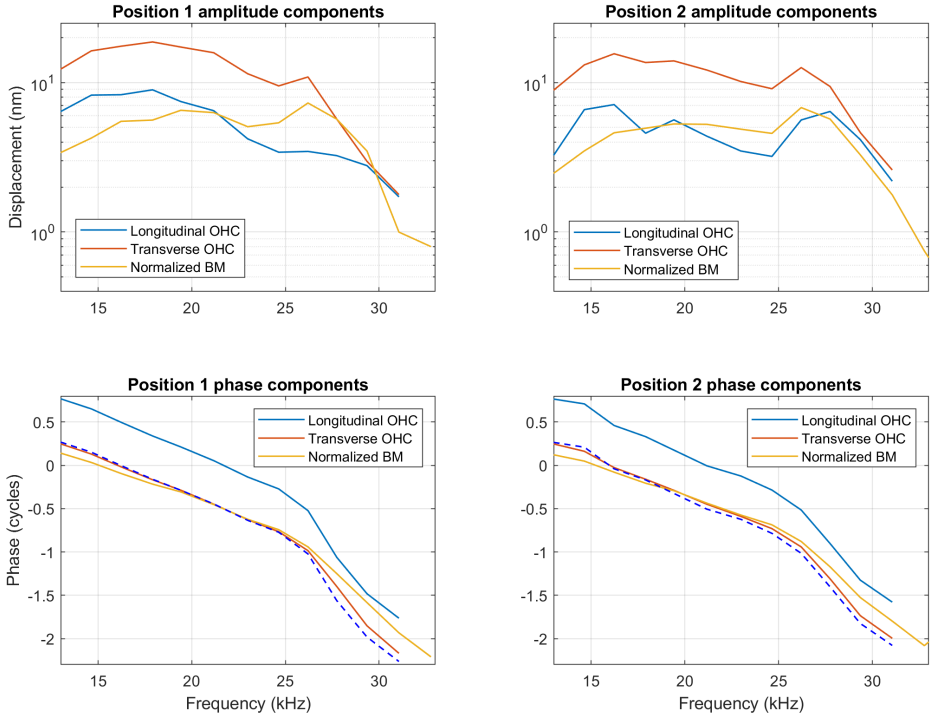


FIGURE 4. Amplitude and phase response of transverse and longitudinal OHC motion, reconstructed at two different positions spaced 45 μm apart longitudinally. Presented also is the BM displacement at this position, taken from the aligned BM displacement with 45° measurement angle. We have normalized this BM displacement value by multiplying by $\sqrt{2}$ – the theoretical geometric loss incurred by measuring a purely transverse BM motion at a 45° angle. The dashed blue phase curve is the longitudinal OHC phase curve subtracted by 180°, i.e. the apex-to-base phase.

REFERENCES

1. J. B. Dewey, A. Altoe, C. A. Shera, B. E. Applegate, and J. S. Oghalai, "Cochlear outer hair cell electromotility enhances organ of corti motion on a cycle-by-cycle basis at high frequencies in vivo," Proceedings of the National Academy of Sciences of the United States of America **118**, 37625 (2021).
2. N. H. Cho and S. Puria, "Motion of the cochlear reticular lamina varies radially across outer-hair-cell rows," bioRxiv (2022), 10.1101/2022.03.01.482580, https://www.biorxiv.org/content/early/2022/03/04/2022.03.01.482580.full.pdf.
3. N. P. Cooper, A. Vavakou, and M. van der Heijden, "Vibration hotspots reveal longitudinal funneling of sound-evoked motion in the mammalian cochlea," Nature Communications **9**, 3054 (2018).
4. B. L. Frost, C. E. Strimbu, and E. S. Olson, "Using volumetric optical coherence tomography to achieve spatially resolved organ of corti vibration measurements," Journal of the Acoustical Society of America **151**, 1115–1125 (2022).
5. W. He and T. Ren, "Transverse vibrations of the reticular lamina and basilar membrane in the basal turn of gerbil cochleae," (2022), association for Research in Otolaryngology Midwinter Meeting.
6. T. Ren, W. He, and D. Kemp, "Reticular lamina and basilar membrane vibrations in living mouse cochleae," Proceedings of the National Academy of Sciences of the United States of America **113**, 9910–9915 (2016).
7. H. Y. Lee, P. D. Raphael, A. Xia, J. Kim, N. Grillet, B. E. Applegate, A. K. E. Bowden, and J. S. Oghalai, "Two-dimensional cochlear micromechanics measured in vivo demonstrate radial tuning within the mouse organ of corti," Journal of Neuroscience **36**, 8160–8173 (2016).

Otoacoustic Emissions in a Deep-Neural-Network Model of Cochlear Mechanics

Arthur Van Den Broucke,[a] Fotios Drakopoulus, Deepak Baby, and Sarah Verhulst[b]

Hearing Technology @ WAVES, Dept. Information Technology, Ghent University, Belgium.

[a] *Electronic mail: Arthur.VanDenBroucke@UGent.be*
[b] *Electronic mail: Sarah.Verhulst@UGent.be*

Abstract. Otoacoustic emission (OAE) generation depends on global cochlear mechanics as well as on the properties of local cochlear nonlinearities associated with outer-hair-cells. A popular method to simulate OAE generation is via nonlinear transmission-line models of cochlear processing that incorporate the coherent reflection theory and cochlear distortion-products. In this study, we investigate whether another type of cochlear models can also be used to simulate realistic OAE properties. In particular, we consider a deep-neural-network approximation of a cochlear transmission-line model that accounts for nonlinear, dispersion and tuning properties of human basilar-membrane displacement (CoNNear). Because convolutional-neural-network (CNN) models have a fundamentally different architecture than that of a longitudinal coupled filterbank, it is not trivial how reverse traveling waves would emerge from such models. We investigated the properties of simulated click-evoked and distortion-product OAEs in this study, and conclude that trained CNN architectures of cochlear models are able to simulate properties associated with local OAE generation, but fail to account for the global phenomenon of reverse traveling waves. While not all aspects of OAE generation are captured in CoNNear, the method holds promise because CNN architectures can be used in connection with measured DPOAEs in closed-loop back-propagation systems to design new stimuli for hearing diagnostics or auditory research experiments.

INTRODUCTION

Otoacoustic emission can be simulated using 1-D transmission-line (TL) models of cochlear processing that represent basilar-membrane (BM) vibration as a cascade of fluid-coupled bandpass filters (e.g. [1, 2]). When introducing small impedance mismatches along the simulated cochlear transmission-line, energy in the forward traveling wave can back-propagate to the cochlear base and generate reflection-source emissions [3]. At the same time, the characteristics of the cochlear nonlinearity will determine the compressive growth features of basilar-membrane processing, and correspondingly the frequencies and strength of simulated distortion-product OAEs. Over the years, both cochlear model descriptions (e.g. [4]) and OAE generation theories [3] have been refined, but from an application perspective, we cannot wait until the theory of hearing is complete. In fact, (imperfect) models of OAE generation have a long history in explaining the mechanisms underlying experimental observations, generating hypotheses for new experiments, or in the design of auditory stimuli that result in OAEs that are quantify frequency-specific hearing deficits. For example, recorded OAEs and auditory evoked potentials can be used to individualize the hearing-impairment parameters in cochlear models [5, 6], that subsequently can be used to design of hearing-aid algorithms (e.g. [7]).

However, the full potential of *in-silico* approaches for hearing diagnostics and treatment can presently not be fully exploited. In particular, the numerical optimization of a stimulus or hearing-aid algorithm using a closed-loop system requires *reverse engineering* of, or *back-propagation* through, the state-of-the-art nonlinear analytical cochlear model inside the loop (Fig.1b). This operation is non-trivial for state-of-the-art auditory models with coupled nonlinear differential equations. To enable back-propagation and speed up the computations, we recently developed a neural-network framework that uses convolutions and differentiable operations to represent cochlear, inner-hair-cell and auditory-nerve processing, i.e. CoNNear (Fig.1a). With an appropriate set of hyperparameters, we showed that a CNN architecture can simulate (end-to-end) time-domain BM displacement responses and properties (Fig.1c-f). Simulations were fairly accurate and the model architecture was shown to generalize well to stimuli and auditory response features it did not see during training [8, 9]. For example, even though the model was never trained to simulate distortion-product otoacoustic emissions (DPOAEs), they simply emerged from the solution. This emerging model property holds promise, because if CNN versions of cochlear models can capture the dominant generators of OAEs, we could embed them into closed-loop systems.

In this paper, we further investigate to which extent CNN models of cochlear processing *learnt* key features of OAE generation as part of their parameter constraining procedure. On the one hand, we study whether local OAE phenomena are captured when CoNNear training was based on 1-D TL model simulations of time-domain BM responses to a

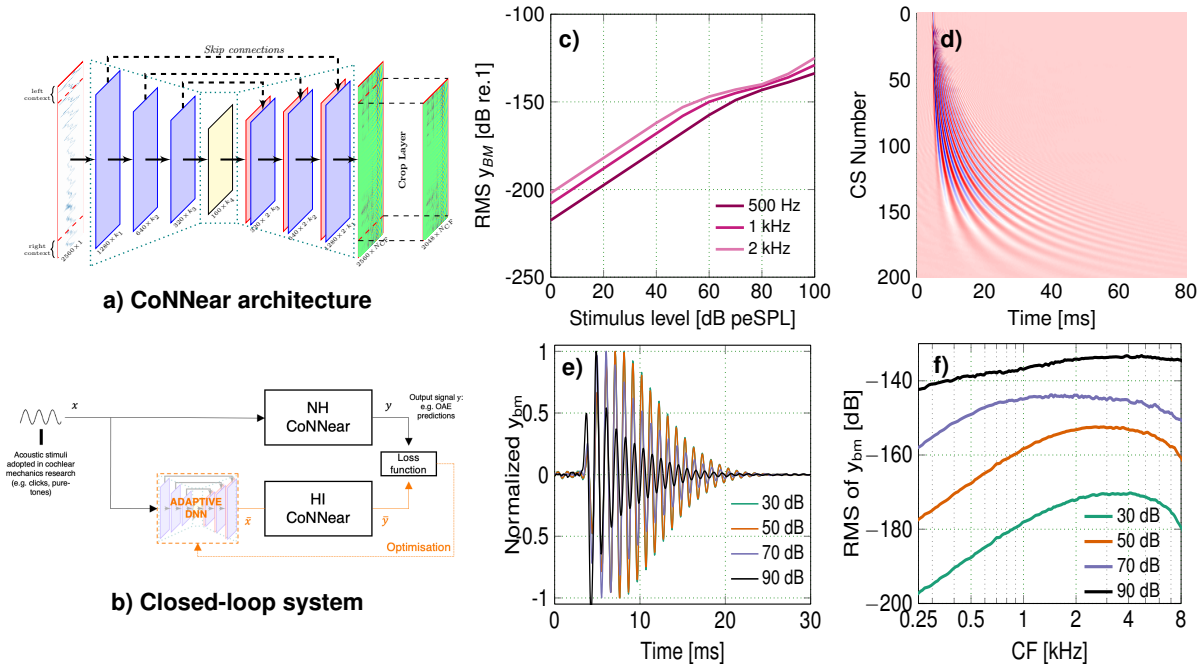


FIGURE 1. Overview of the CoNNear cochlear model: (a) The CNN architecture of CoNNear consists of 4 encoder and 4 decoder layers, uses strides and skip connections to map 2048 audio samples to BM displacement waveforms of shape (201 N_{CF} x 2048) using 1-D (de)convolution layers and tanh-activation functions. (b) Example of a closed-loop system where optimisation via backpropagation is applied to an adaptive DNN to map e.g. hearing-impaired (HI) CoNNear predictions of OAEs to normal hearing (NH) simulations. (c)-(f): Cochlear mechanics features derived from simulated CoNNear outputs to stimuli unseen during training: (c) I/O functions calculated as input-click stimulus level in dB peSPL versus BM displacement evaluated at the 500 Hz, 1kHz and 2kHz CF-channels. (d) Instantaneous BM displacement intensities for CFs (channel numbers, CS) between 100 Hz (channel 201) and 12 kHz (channel 1). The colour scale ranges between -15 µm (blue) and 15 µm. (e) Normalized impulse response (IR) at the 1 kHz channel for 100-μs click stimuli of different levels. The IRs were normalized by the maximal pressure at the stapes of the cochlea such that compression is observed as a reduction of the IR amplitude. (f) Simulated excitation patterns of BM displacement (rms of BM IR waveform at each CF) for clicks between 30 and 90 dB SPL.

speech corpus presented at 70 dB SPL. We focus on DP generation and local BM reflection sources (RS) associated with the generators of reverse traveling waves. On the other, we evaluate to which extent CoNNear is able to capture global cochlear phenomena associated with reverse traveling waves, or reflection source OAEs. We use a proxy measure of simulated ear-canal pressure for this purpose and evaluate whether the level-dependence and frequency characteristics of click-evoked otoacoustic emissions (CEOAEs) are present. Following the coherent reflection filtering theory, CEOAEs crucially rely on the longitudinal coupling of cochlear structures and the presence of impedance mismatches. It is unclear whether a CNN-based cochlear model can be trained to account for these global cochlear phenomena.

METHODS

We define the following variables to compare simulated CoNNear BM displacement waveforms with reference simulations of a 1-D nonlinear TL model of cochlear mechanics that captures aspects of human cochlear tuning, level-dependence and OAE generation [10, 11]. CoNNear is a CNN-based model of cochlear processing [8] and simulates forward middle-ear transduction, BM displacement, but not reverse middle-ear transduction. Global OAE generator mechanisms are therefore evaluated near the cochlear base of CoNNear (12-kHz CF channel) that could serve as input to a reverse middle-ear filter to yield the ear-canal-pressure, or OAE, as we would measure it in a test subject.

- OAE_{RS}_TL: The reflection source OAE, computed as the difference signal between the simulated ear-canal-pressure of a TL model with cochlear irregularities and a smooth TL model.

- DPOAE: $2f_1$-f_2 component of the 12-kHz CF channel BM response simulated either in the TL or CoNNear model to allow for direct comparison between model implementations.

- OAE_{RS}: The global reflection source component of the 12-kHz CF channel BM response computed as the difference signal between the simulated BM IR of a TL/CoNNear model with cochlear irregularities and a smooth TL/CoNNear model.

- $BM_{Xk,RS}$: The local reflection source component of the X-kHz CF channel BM response computed as the difference signal between the simulated BM IR of a TL/CoNNear model with cochlear irregularities and a smooth TL/CoNNear model.

Transmission-line and CoNNear models

The parameters in the reference TL model are derived from a cochlear input impedance description [2] and TLs are mechanistically well-suited to match the cochlear input impedance, the fluid-membrane coupling along the cochlear BM, and forward and reverse traveling waves. Nonlinear TL-models can account for distortion- as well as reflection-source OAEs as described in [3]. Specifically, the reference TL cochlear model includes a random-across-CF maximal 5% fluctuation around the double pole of the cochlear admittance that can be seen as a *cochlear roughness* pattern that generates reflection sources that cause reverse traveling waves along the TL and OAEs in the simulated ear-canal pressure [10]. Distortion-source emission generators are a natural consequence of the instantaneous nonlinearity included in the reference TL model that maximally preserves the phase of the BM zero-crossing as stimulus level increases [11, 12].
The CoNNear model was designed to offer a faster ($>$ 2000x speed-up) and differentiable version of the reference TL-model [8]. Its hyperparameters were initially determined using a principled approach that took into account the L1-loss between the CoNNear BM displacement and that of the reference TL-model at 201 CF sections spaced along the Greenwood map. The Timit speech corpus (at 70 dB SPL) was used for this purpose, and afterwards the models' performance was tested on six cochlear mechanics experiments using stimuli that were not part of the training set to evaluate and fine-tune the parameters related to model layer depth and activation function. CoNNear was developed using the Keras machine learning library [13] with a Tensorflow [14] back-end, and can be downloaded from github.com/hearingtechnology. Along with the simulations presented in [8], Fig. 1 shows that the final CoNNear architecture (8-layer, CNN encoder-decoder network with *tanh*-nonlinearity) is able to capture a variety of cochlear mechanics features related to cochlear dispersion and compression, in the same way as the original TL model did. Unlike the TL-model, the CNN architecture is not mechanistically representing coupled filters with impedance gradients that facilitate traveling wave propagation in either direction. With this in mind, we cannot call CoNNear a cochlear mechanics model *as such*, and neither can we expect that it would simulate OAEs, given that it was trained to map stimulus to BM-displacement-across-CF only.

To investigate whether the CoNNear model simulates reflection-source emissions, we isolated the local and global cochlear reflection sources from the BM and ear-canal-pressure waveforms by subtracting the simulations from a *smooth* CoNNear model from a CoNNear model that was trained using a TL-model that included impedance mismatches (i.e., a *rough* CoNNear model). This subtraction procedure corresponds to isolating the RS-OAE component in TL-model simulations [15] and can be performed in-silico, but not in-vivo using ear-canal recordings. To generate rough and smooth CoNNear models, we trained two CoNNear models using 2310 70-dB-SPL speech fragments from the TIMIT dataset. This dataset was either given as an input to the reference TL-model that included 5% irregularities (rough) or not (smooth) to generate input-output data pairs for training rough and smooth CoNNear models. The original CoNNear model in [8] was trained using BM displacement simulations generated with a rough cochlear TL-model. It is important to note that the irregularities were generated using a uniform random number generation. The TL model that was used to generate the training set for the rough CoNNear model and the reference TL model had the same random-number-generator seed and ran on the same computer to allow for a fair comparison between both model implementations.

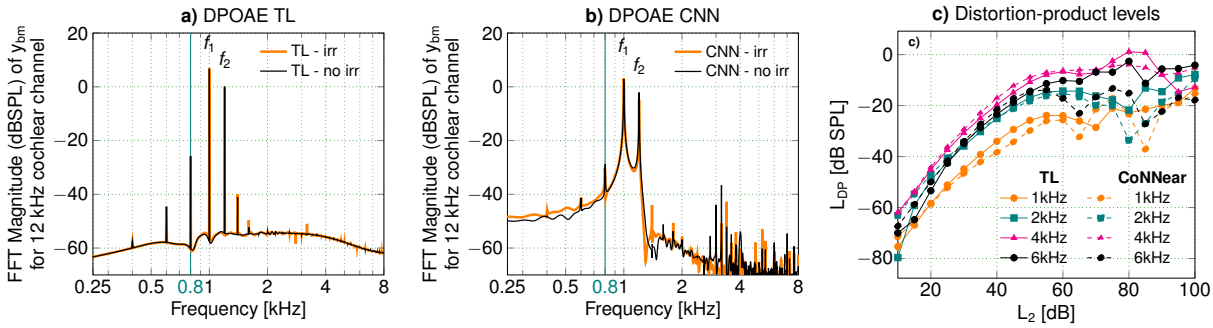


FIGURE 2. Simulated DPOAEs across model architecture. The frequency response of the 12-kHz CF channel (i.e., fast Fourier transform of the BM displacement waveform) was considered as a proxy for OAEs recorded in the ear-canal. Panels (a) and (b) show simulations in response to two pure tones of f_1 of 1.0 and f_2 of 1.2 kHz with L_2 = 50 dB SPL and L_1=0.4 L_2+39 [16]. The $2f_1-f_2$ DP at 0.8 kHz was indicated with a teal line. (c) Simulated $2f_1-f_2$ levels (L_{DP}) for TL and the CoNNear model implementations. f_1 was varied between 1 and 6 kHz, and DP stimuli followed a f_2/f_1 ratio of 1.2.

RESULTS

Distortion-source OAEs

Because distortion-source OAE generators are predominantly a local cochlear phenomenon, this is the first feature we investigate in Fig. 2. Panels a & b compare simulated DPOAEs from the 12-kHz BM response that was either simulated using a smooth (no irr) or rough (irr) cochlear TL/CoNNear model. Note that the sampling frequency in the TL model was 100 kHz, against 20 kHz in the CoNNear model, but that TL simulations were downsampled to 20 kHz before we compared their frequency responses. While both models were able to simulate a $2f_1-f_2$ DP, there was more off-DP, high-frequency noise introduced in CoNNear, and there were differences in both the amplitude of the DPOAE and the OAE_{RS} component (i.e., difference between rough and smooth model simulations) across implementations. The spectral energy in the $2f_1-f_2$ DP was up to 5 dB smaller in CoNNear than in the TL model, and peaks departed from higher energy plateaus. The *spectral-peak-to-noisefloor* was hence reduced in the CNN approximation compared to the original TL model.

Level-dependent properties of simulated DPOAEs in the rough model implementations are shown in the DPOAE growth functions for f_2 primaries between 1 and 6 kHz (panel c). For stimulus levels below 70 dB SPL, CoNNear captures the nonlinear growth of TL-simulated DPOAEs well and for higher levels we see a divergence that can be attributed to the adopted training procedure. CoNNear was trained using a speech corpus presented at 70 dB SPL and hence performs better at lower stimulus levels, whereas the TL model accuracy is level independent. We conclude that CoNNear is qualitatively able to capture the growth characteristics of TL-simulated and recorded DPOAEs for L_2 stimulation levels up to 70 dB, and that the nature of the CoNNear OAE_{RS} should be investigated further.

Reflection-source OAEs

The OAE_{RS} component is better observed in click-evoked otoacoustic emissions (CEOAEs) that are separated from the evoking click in the time domain. Characteristic to CEOAEs is that high-frequency energy occurs earlier in time than low-frequency energy. In TL models of CEOAE generation, this trait originates through reverse traveling waves that stem from local BM reflection sources (i.e., BM_{RS}) traveling in reverse through the transmission line. Figure 3 shows normalized ear-canal pressure (a) and extracted reflection-source OAEs (OAE_{RS}, b) simulated in the reference TL model. TL-model simulated OAE_{RS}s and corresponding frequency spectra (Fig. 4) together capture the typical experimentally observed CEOAE-latency-decrease for increasing OAE component frequency [17]. Because CoNNear was not designed to simulate reverse middle-ear-transmission [18], a fair comparison between OAE_{RS} simulations can only be performed using simulated BM responses. We took the output of the 12-kHz basal BM CF response (maximally represented CF in CoNNear) as a proxy for the global cochlear input to the middle-ear transmission and ear-canal in both implementations (panel c & d). The derived $BM_{12k,RS}$ displacement response in the TL model (c)

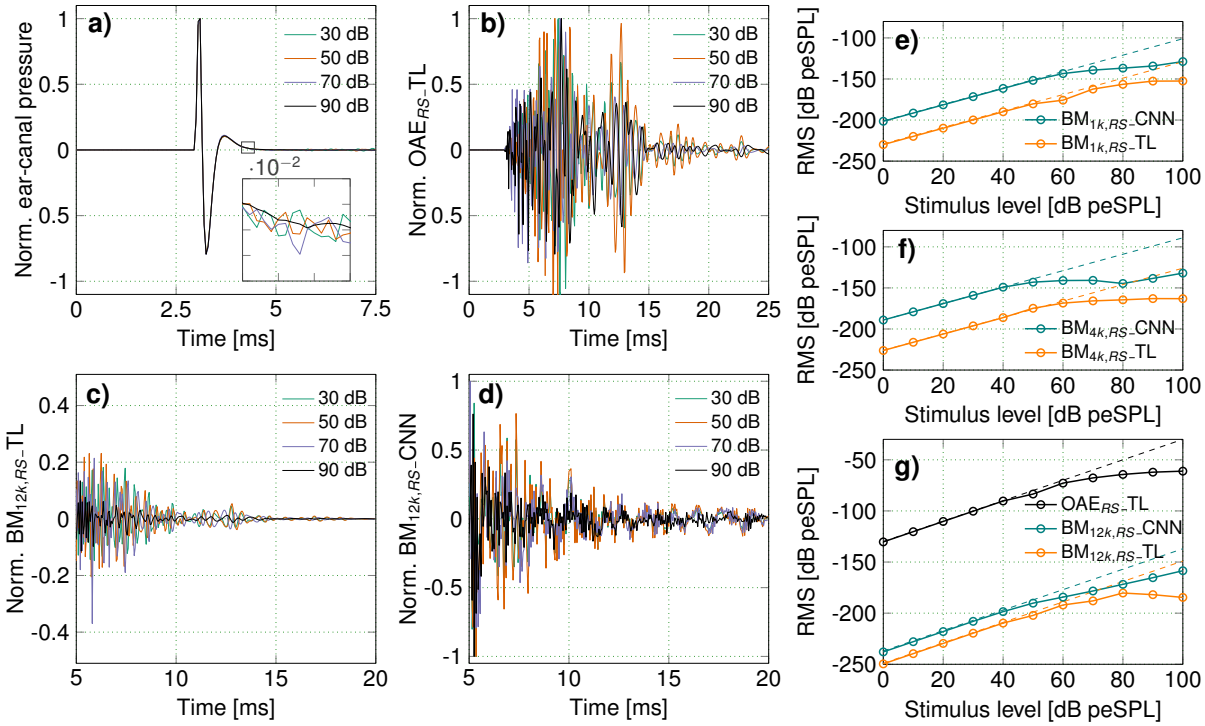


FIGURE 3. Simulated reflection-sourece OAEs (a) Simulated ear-canal pressure in the reference TL model with cochlear irregularities. The inset shows the OAE which is present after the linear click response has passed. (b) Normalized OAE_{RS}_TL for clicks of 4 different levels (c) Normalized $BM_{12k,RS}$_TL (d) Normalized $BM_{12k,RS}$_CNN. Both (c) and (d) plots were normalized for each input sound level with respect to the max value of BM displacement after 5 ms (i.e. 2 ms after the click onset) for the CNN model for the designated level. (e)-(g) displays the $BM_{Xk,RS}$_TL/CNN for 3 different frequencies. The level curves are calculated as the level of the click in dB peSPL versus the rms level of the $BM_{Xk,RS}$_TL/CNN (i.e., waveform starting 2 ms after click onset). (g) Additionally shows the level of the click in dB peSPL versus the rms level of the OAE_{RS}_TL (black curve).

shows similar characteristics to the ear-canal-pressure derived OAE_{RS} (b) and is therefore used to study the OAE_{RS} generator mechanisms of CoNNear. Panel c shows that CoNNear-generated $BM_{12k,RS}$ responses are generally more noisy, but also contain longer latency, low-frequency components for stimulus levels up to 70 dB (i.e. in the 10-15 ms region). Even though CoNNear does not explicitly allow for reverse traveling waves in its architecture, the higher CF-channels appear to have learnt to represent energy associated with global reverse traveling waves of the original TL-model that was used for training. However, there are quantitative differences. Panels (b) and (c) of Fig.4 compare BM_{RS} responses at 12 and 8 kHz resp., and show that the frequency content of CoNNear is typically higher in the lower frequencies and more distributed. Secondly, when comparing the strength of locally generated BM_{RS} responses for different stimulus levels (Fig.3 panels e,f,g), we see that CoNNear (green) typically generates stronger BM reflection sources than the TL model (orange). It is possible that BM_{RS} responses in CoNNear stem from baseline CNN training noise that differs between the trained smooth and rough cochlear models, with no further relation to reflection-source emissions. The level-dependence of simulated CoNNear BM_{RS} responses appears to counter this notion. Fig.3e,f,g shows that simulated BM_{RS} responses grow compressively with increasing stimulus level. While machine noise is not expected to depend on stimulus level, the size of the reflection source emissions is. Given a fixed-size roughness pattern and compressively growing forward-traveling waves, the resulting reflection-source OAE will show compression [3], as is typically observed in the growth pattern of recorded CEOAEs [19]. Despite the apparent differences between simulated CoNNear and TL BM_{RS} responses, both the visual time-frequency distribution and compressive growth of simulated CEOAEs show promise in that the CoNNear model may have learnt some aspects of the reflection-source OAE generators associated with TL-models of OAE generation.

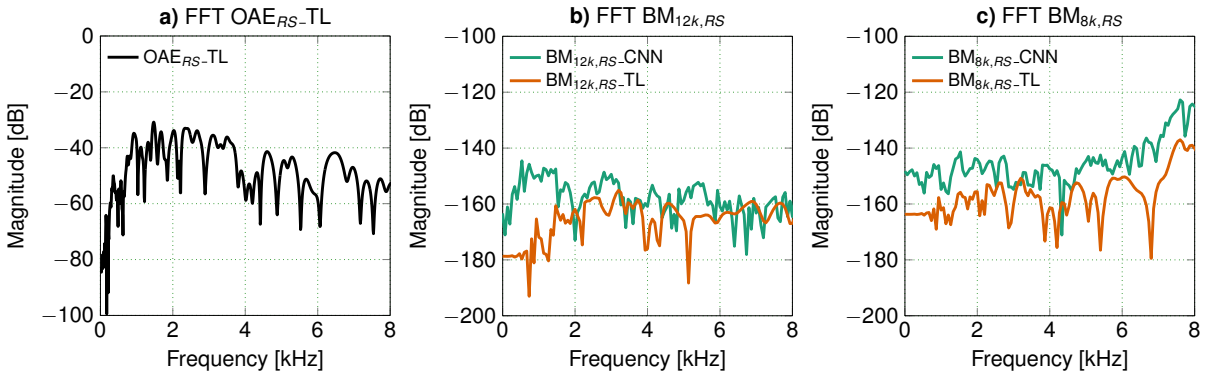


FIGURE 4. Spectra of simulated reflection-source emissions (a) FFT of OAE_{RS}_TL (b) FFT of $BM_{8k,RS}$_TL/CNN (c) FFT of $BM_{12k,RS}$_TL/CNN. Magnitude spectra were computed using $20\log_{10}(\text{abs}(\text{FFT}(x)))$ where x corresponds to the 5-20-ms window of the time-domain response.

DISCUSSION

Even though CoNNear was trained to minimize the L1-loss between 201 TL and CoNNear simulated BM responses [8], some aspects associated with OAE generation were also learnt throughout the process. For one, Fig.2b,c show that the *tanh*-nonlinearity between the auto-encoder layers was able to simulate distortion products that bear resemblance to those generated in a TL-model of cochlear processing. Particularly, Fig.2c illustrates nonlinear growth of the $2f_1-f_2$ DP that aligns well with the TL DP growth for several primary frequencies and stimulus levels within the training range. Obtaining improved DP simulation accuracy for stimulus levels above 70 dB, might require retraining with the corpus presented at higher stimulation levels. We conclude that distributed fixed-shape nonlinearlities within auto-encoder structures can approximate that of a local BM nonlinearlity with feedback [2, 12]. We earlier demonstrated that the shape of the activation function does matter in this regard: CNN networks that use a ReLU nonlinearity (linear growth for positive input) cannot be trained to capture BM compressive growth nor DP generation [8].

Assessing how well CoNNear succeeded at simulating reflection-source emissions as resulting from local reflection-source generators and reverse traveling waves was more challenging. First, CoNNear simulations were generally more noisy than TL-model simulations, and this is best observed in the DPOAE simulations in Fig.2a,b where CoNNear spectral peaks departed from a higher plateau. The baseline CNN noise of course has consequences for simulated BM_{RS} responses which form the difference signal between a trained smooth (no irr) and rough (irr) CoNNear model. CNN training noise can cause the difference signal (BM_{RS}) to be larger than initially intended from the introduced 5% impedance irregularity in the rough TL cochlear model. Indeed, the 10+ dB larger BM_{RS} responses at different CFs (Fig.3e,f,g), along with their broadband spectra (Fig.4b,c) support the notion that there was broadband CNN training noise in the simulated CoNNear BM_{RS} simulations that is interfering with the intended roughness pattern associated with local BM_{RS} generator sources. However, simulated local BM_{RS} responses did show compressive growth (Fig.3e,f,g) and the time-domain $BM_{12k,RS}$ proxy for reflection-source OAEs (Fig.3d) showed a visual similarity for lower frequencies occurring later in time than higher frequencies. It thus appears that CoNNear learnt some aspects associated with reflection-source OAEs. This is promising but not entirely convincing because it remains to be tested under which conditions the CNN training noise can be reduced and whether OAE-specific training would improve the rough CoNNear model quality. Multi-layer CNN architectures are fundamentally different from mechanistic TL models, and while the former might generate accurate outputs, it computes them in a fundamentally different way. The only way in which this -in theory- could work is when the CNN accurately learnt to solve the set of coupled, nonlinear time-domain equations that describe the TL-model implementation.

Before using CoNNear in a closed-loop circuit to design experimental DPOAE protocols or hearing diagnostic tools (Fig.1b), we suggest the following improvements. Modify the CoNNear architecture such to introduce an additional time-domain output that relates couples the basal BM displacement via a mechanical bandpass filter [18] to the ear-canal-pressure. This can be achieved using a blackbox-like approach that adds a specialized CNN layer that can learn to represent the middle-ear-transduction based on the ear-canal-pressure simulations that are generated along with BM displacement simulations in the original TL model. Secondly, we suggest using an adjusted or combined loss function in the training procedure that takes into account the difference signal between CoNNear-simulated and

TL-model simulated reflection-source emissions such as CEOAEs. As local distortion-sources were already realistic during the original training procedure with BM displacement patterns to a speech corpus, it might not be necessary to use DPOAEs during the training phase. This brings the advantage that the realism with which DPOAE reflection and distortion sources are captured by the trained $\text{CoNNear}_{\text{OAE}}$ model can remain an independent evaluation metric. Lastly, just as the TL-reference model is able to simulate different roughness patterns by changing the seed in the random number generator to simulate the myriad of experimentally observed person-specific CEOAEs, $\text{CoNNear}_{\text{OAE}}$ should be able to be tailored to the individual. To this end, we suggest to use the transfer-learning approach we earlier adopted to go from normal-hearing to hearing-impaired implementations of CoNNear using a limited training set (50 in stead of 2310 sentences) [20]. The limited training set could be generated using a TL-model that is individualized on the basis of recorded DPOAEs [6, 21] or CEOAEs [22].

ACKNOWLEDGMENTS

Work supported by European Research Council ERC-StG-678120 (SV,DB,FD) and FWO-SBO 1SB4421N (AVDB).

REFERENCES

1. E. de Boer, "Auditory physics. physical principles in hearing theory. i," Physics reports **62**, 87–174 (1980).
2. G. Zweig, "Finding the impedance of the organ of corti," The Journal of the Acoustical Society of America **89**, 1229–1254 (1991).
3. C. A. Shera and J. J. Guinan Jr, "Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian oaes," The Journal of the Acoustical Society of America **105**, 782–798 (1999).
4. A. Altoè and C. A. Shera, "The cochlear ear horn: geometric origin of tonotopic variations in auditory signal processing," Scientific reports **10**, 1–10 (2020).
5. S. Keshishzadeh, M. Garrett, and S. Verhulst, "Towards personalized auditory models: Predicting individual sensorineural hearing-loss profiles from recorded human auditory physiology," Trends in Hearing **25**, 2331216520988406 (2021).
6. S. Keshishzadeh and S. Verhulst, "Individualized cochlear models based on distortion product otoacoustic emissions," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (IEEE, 2021) pp. 403–407.
7. F. Drakopoulos, V. Vasilkov, A. O. Vecchi, T. Wartenberg, and S. Verhulst, "Model-based hearing-restoration strategies for cochlear synaptopathy pathologies," bioRxiv (2022).
8. D. Baby, A. Van Den Broucke, and S. Verhulst, "A convolutional neural-network model of human cochlear mechanics and filter tuning for real-time applications," Nature machine intelligence **3**, 134–143 (2021).
9. F. Drakopoulos, D. Baby, and S. Verhulst, "A convolutional neural-network framework for modelling auditory sensory cells and synapses," Communications Biology **4**, 1–17 (2021).
10. S. Verhulst, T. Dau, and C. A. Shera, "Nonlinear time-domain cochlear model for transient stimulation and human otoacoustic emission," The Journal of the Acoustical Society of America **132**, 3842–3848 (2012).
11. S. Verhulst, A. Altoe, and V. Vasilkov, "Computational modeling of the human auditory periphery: Auditory-nerve responses, evoked potentials and hearing loss," Hearing research **360**, 55–75 (2018).
12. C. A. Shera, "Intensity-invariance of fine time structure in basilar-membrane click responses: Implications for cochlear mechanics," The Journal of the Acoustical Society of America **110**, 332–348 (2001).
13. F. Chollet *et al.*, "Keras," https://keras.io (2015).
14. M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," https://www.tensorflow.org/ (2015), software available from tensorflow.org.
15. S. Verhulst and C. A. Shera, "Relating the variability of tone-burst otoacoustic emission and auditory brainstem response latencies to the underlying cochlear mechanics," in *AIP conference proceedings*, Vol. 1703 (AIP Publishing LLC, 2015) p. 090003.
16. P. Kummer, T. Janssen, P. Hulin, and W. Arnold, "Optimal l1- l2 primary tone level separation remains independent of test frequency in humans," Hearing research **146**, 47–56 (2000).
17. R. Sisto and A. Moleti, "Transient evoked otoacoustic emission latency and cochlear tuning at different stimulus levels," The Journal of the Acoustical Society of America **122**, 2183–2190 (2007).
18. S. Puria, "Measurements of human middle ear forward and reverse acoustics: implications for otoacoustic emissions," The Journal of the Acoustical Society of America **113**, 2773–2789 (2003).
19. S. Verhulst, J. M. Harte, and T. Dau, "Temporal suppression of the click-evoked otoacoustic emission level-curve," The Journal of the Acoustical Society of America **129**, 1452–1463 (2011).
20. A. Van Den Broucke, D. Baby, and S. Verhulst, "Hearing-impaired bio-inspired cochlear models for real-time auditory applications," in *21st Annual Conference of the International Speech Communication Association (INTERSPEECH 2020)* (International Speech Communication Association (ISCA), 2020) pp. 2842–2846.
21. B. N. Buran, G. P. McMillan, S. Keshishzadeh, S. Verhulst, and N. F. Bramhall, "Predicting synapse counts in living humans by combining computational models with auditory physiology," The Journal of the Acoustical Society of America **151**, 561–576 (2022).

22. S. Keshishzadeh and S. Verhulst, "From derived-band envelope-following responses to individualized models of near-and supra-threshold hearing deficits," in *International Symposium on Auditory and Audiological Research (ISAAR 2019)*, Vol. 7 (2019) pp. 13–20.
23. G. Zweig and C. A. Shera, "The origin of periodicity in the spectrum of evoked otoacoustic emissions," The Journal of the Acoustical Society of America **98**, 2018–2047 (1995).
24. V. Vasilkov, M. Garrett, M. Mauermann, and S. Verhulst, "Enhancing the sensitivity of the envelope-following response for cochlear synaptopathy screening in humans: The role of stimulus envelope," Hearing Research **400**, 108132 (2021).

Diagnosis of mechanical ear pathologies using a classification model

Kristine Elisabeth Eberhard[1,2], Gabrielle R. Merchant[3], Hideko Heidi Nakajima[1], Stephen T. Neely[3,b]

[1]*Harvard Medical School and Massachusetts Eye and Ear, 243 Charles Street, Boston, MA, USA*
[2]*Copenhagen Hearing and Balance Centre, Department of Otolaryngology, Head and Neck Surgery & Audiology, Rigshospitalet, Inge Lehmanns Vej 8, 2100 Copenhagen, Denmark*
[3]*Boys Town National Research Hospital, 555 North 30th Street Omaha, NE, USA*

[a)]Corresponding Author: stephen.neely@boystown.org

Abstract. Wideband tympanometry (WBT) has potential as a non-invasive diagnostic tool for pathologies that cause changes to mechanics of the middle and inner ear. However, because WBT data are complex and thus difficult to interpret, WBT has not been widely implemented in the clinic. Adult patients with conductive hearing loss and normal otoscopic exam have diagnostic uncertainty which can be helped by early diagnosis. In addition to middle-ear pathology, inner-ear pathology such as superior canal dehiscence (SCD) can be better managed with early detection. In this study we measured WBT in adult ears with SCD, stapes fixation (SF) and normal ears. A total of 204 ears were included. WBT was measured using the Titan WBT hardware and the Research Platform software (Interacoustics, Middelfart, Denmark) to obtain the raw WBT recordings which were then filtered by our mitigation software to reduce low-frequency noise. Air-bone gap (ABG) was calculated from audiometric measurements. We developed an automated classification model that used multivariate logistic regression, and based on the highest probability, classified ears as SCD, SF or normal. The model features included principal components of WBT absorbance and frequency-averaged ABG. A cross-validation approach with 1000 iterations was used to train and validate the model to improve generalizability of results. In the validation set, correct classification of SCD, SF and normal ears were achieved for 94.8% of ears. Sensitivity for SCD was 80.9% and for SF 97.1%, whereas specificity for SCD was 97.5% and for SF 99.2%. Performance of an absorbance-only and an ABG-only model was also tested. The technique we developed performed surprisingly well for the ABG-only model for the dataset recorded. Adding WBT absorbance to the ABG-only model further improved performance, especially for SF sensitivity. We find that automation of SCD and SF differential diagnosis based on our regression models on absorbance and ABG is feasible and will provide clinicians with a powerful user-friendly diagnostic tool.

INTRODUCTION

Several studies have shown that wideband tympanometry (**WBT**) and the derived metric *absorbance* have great potential to diagnose mechanical pathologies of the ear (Feeney, Grant and Marryott, 2003; Hunter, Bagger-Sjöbäck and Lundberg, 2008; Shahnaz, Longridge and Bell, 2009; Nakajima et al., 2012, 2013; Merchant et al., 2015, 2021; Feeney et al., 2020). However, WBT has not been widely adopted by clinicians. A major reason clinicians cannot use WBT data to diagnose pathology is because WBT data are complex, and no diagnostic algorithm exists to interpret the complex data for various diagnoses. Additionally, much of the research in this area has been on group level (e.g. averaged) effects and patterns associated with certain conditions. Translation of group-level effects to what a clinician will observe in an individual patient is not straightforward. For example, many of the fine structure details in individual WBT and absorbance measurements are smoothed by group-level summation metrics such as averaging. As such, there is a need for automated methods that can perform accurate diagnosis from an individual ear's measurements.

With an algorithm that focuses on interpreting individual WBT data and automation, this non-invasive, easy and objective measurement will be feasible in a clinical setting.

In adult patients with conductive hearing loss and a normal otoscopy exam, determining the pathology causing the hearing loss non-invasively can be difficult (Masud et al., 2019; Dewyer, Quesnel and Santos, 2020). Imaging (i.e., computed tomography (**CT**) scan) is not performed as part of the initial assessment. A CT scan has radiation exposure, is costly, and cannot always provide a conclusive diagnosis (Fuse et al., 1992; Masud et al., 2019). While surgical middle-ear exploration can often confirm middle-ear pathology such as stapes fixation (**SF**), the surgery is invasive and it may also be difficult to visualize or interpret lesions with palpation of the middle-ear chain (Schimanski, Schimanski and Berthold, 2011). In addition, middle-ear exploration cannot diagnose inner-ear mechanical lesions that cause conductive hearing loss such as superior canal dehiscence (**SCD**), a condition for which diagnosis is often delayed due to having similar symptoms to other otologic and neurologic conditions, such as low-frequency conductive hearing loss and dizziness (Minor, 2005; Zhou, Gopen and Poe, 2007; Niesten et al., 2012; Dewyer, Quesnel and Santos, 2020; Eberhard, Chari, et al., 2021).

In this study, we measured WBT in adult ears with mechanical middle- and inner-ear pathologies (SF and SCD) as well as in normal control ears. We developed an automated classification model (a similar approach to Merchant et al., 2021) based on WBT absorbance and air-bone gap (**ABG**) that differentiates between ears with SCD, SF and normal ears and provides the likely diagnosis (Merchant et al., 2021).

METHODS

Subjects & Ears

A total of 204 ears were included in the study comprising 32 SCD ears (30 subjects), 38 SF ears (31 subjects) and 134 normal ears (81 subjects). The number of male ears range from 41-45% among ear category, and mean age was 52, 45 and 31 years old for SCD, SF and normal ears respectively.

Only ears with a single confirmed pathology were included. SF was confirmed by surgery and SCD by high-resolution CT.

Normal control ears had air conduction thresholds ≤ 15 dB HL at 0.25-4 kHz and ≤ 20 dB HL at 6-8 kHz and normal tympanograms (pressure: -100 to +150 daPa; admittance: 0.3 – 1.8 mL; volume: 0.3 – 2.5 mL). They had no history of significant ear disease and had a normal otoscopic exam.

WBT Measurement & Absorbance

WBT was measured using the commercially available Titan WBT hardware and the Research Platform software (Interacoustics, Middelfart, Denmark). This software allows access to the raw microphone pressure response that results from a train of click stimuli presented while changing the ear canal static pressure. This results in a microphone pressure response at a wide range of static pressures. Customized data collection and processing software were developed with MATLAB for extraction and processing of raw data measured by the Research Platform to absorbance, as described in Eberhard et al. 2021 (Eberhard, Ravicz, et al., 2021).

Absorbance at tympanometric peak pressure (**TPP**) was found beneficial compared to absorbance at 0 daPa. As described by previous studies, using absorbance at TPP compensates for differences between middle-ear pressure and ear-canal pressure that could reduce pathology-specific effects (likely from stiffening the tympanic membrane) (Liu et al., 2008; Sanford et al., 2013; Sun, 2016; Feeney et al., 2017). Also, absorbance was computed using the surge method described in Merchant et al. 2019 instead of a constant ear canal area of 44.18 mm^2 because the classification model performed better using the surge method (Merchant et al., 2019). All absorbance measurements in this paper are therefore at TPP and were computed using the surge method.

Audiometric Testing

Participants underwent pure tone air and bone audiometric testing by a clinical audiologist as part of their clinical diagnostic workup or research study participation either on the same day (82% of ears) or a separate day within half a year of their research study participation (on average within 12 days). ABG was computed for this study by taking

differences at each frequency between air-conduction and bone-conduction thresholds in individual ears and averaging across frequencies from 0.5 to 4 kHz. Other variations of ABG were explored, but this average was determined to result in the best overall performance of the classification model (see more below). For ears without measured air-bone gaps due to normal air-conduction thresholds, the ABG was set to zero.

Automated Classification Model

Multivariate logistic regression models were built to use absorbance and ABG features to relate them to the three categorical output variables: SCD, SF and normal. Three models were built: one including only absorbance features, one with only ABG, and one with a combination of absorbance and ABG features. As described above, ABG was averaged across frequencies from 0.5-4 kHz and included as a single feature in both the ABG-only model and the combined absorbance and ABG model.

Training of the model was performed on 70% of the data. To test model performance on unseen clinical data, the accuracy of the model predictions were validated on the remaining 30% of the data. The model was trained and validated iteratively 1000 times using cross-validation to improve generalizability of the results.

Absorbance was measured at 512 frequencies, and because absorbance at the different frequencies will be correlated, a problem with collinearity occurs. To reduce dimensionality of the absorbance data and to solve issues with collinearity, principal component analysis was used to transform absorbance data to fewer features. Since the model performed better using multiple principal components (≥10), regularization was introduced to minimize overfitting. Regularization adds an additional parameter, lambda, when training the model that reduces the risk of overfitting.

The regression model computed the predicted probability of each of the three conditions (SCD, SF and normal). The condition with the highest predicted probability was determined as the likely diagnosis. This probability-modelling approach is similar to Merchant et al., 2021 (Merchant et al., 2021).

RESULTS

Absorbance

Mean absorbance showed pathology specific features (Fig. 1). The mean absorbance below 1 kHz was higher for ears with SCD compared with SF and normal ears and demonstrated a peak around 900 Hz. The mean absorbance of SF ears showed slightly lower absorbance than normal ears both below 1 kHz but also at frequencies from 2-5 kHz.

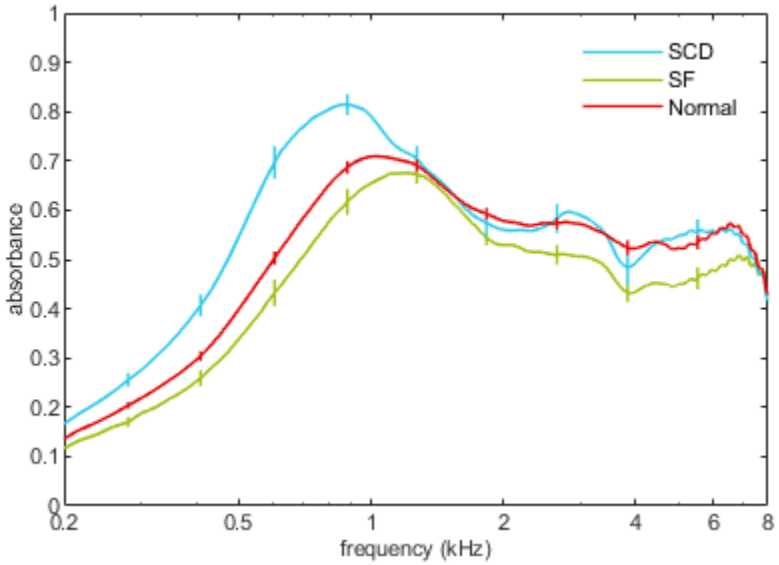


Fig. 1. Mean absorbance and standard error (bars) for ears with SCD (blue), SF (green) and normal ears (red).

Air-Bone Gap

Per definition, normal ears had ABG ≤15dB at frequencies 0.25-4 kHz. If bone conduction thresholds were not measured due to a patient having normal air conduction thresholds, the ABG was set to 0. Figure 2 shows a box plot of the frequency-averaged ABG (0.5-4 kHz) according to ear category, SCD, SF and normal ears. Median ABGs for ears with SCD and SF were 9 and 28 dB, respectively.

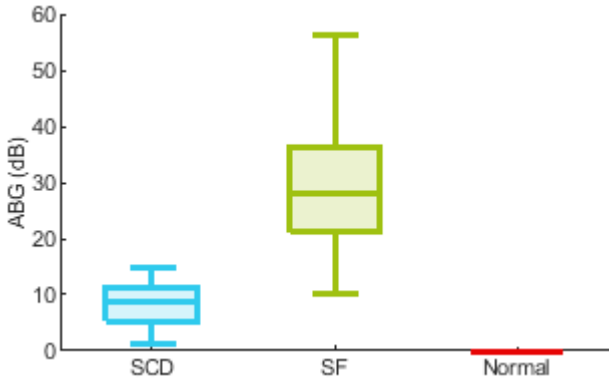


Fig. 2. Boxplot of the frequency-averaged ABG according to ear category, SCD (blue), SF (green) and normal ears (red). The boxplot illustrates median (line within box), interquartile range (box edges) and min/max range (whiskers).

Classification Model Performance

Model performance was assessed based on performance of the validation dataset. In the validation dataset, absorbance-only and ABG-only total error, sensitivities and specificities for each pathology are shown in Table 1. Overall, the ABG-only model performed better than the absorbance-only model.

The best performance was obtained when both absorbance and ABG were included in the model with error rate of 5.2% under the following conditions: 10 principal components of absorbance and the regularization parameter λ of 2e-3. Under these conditions, the combined absorbance and ABG model correctly classified 94.8% of SCD, SF and normal ears in the validation dataset. The sensitivity for SCD was 80.9% and for SF 97.1%. The specificity for SCD was 97.5% and for SF 99.2%. Table 2 presents a confusion matrix summarizing the results of the 1000 validation datasets of the combined absorbance and ABG model.

TABLE 1.
Total error, sensitivity and specificity for SCD and SF of the absorbance-only model, the ABG-only model and the combined absorbance and ABG model.

	Total error (%)	**SCD sensitivity (%)**	**SF sensitivity (%)**	**SCD specificity (%)**	**SF specificity (%)**
Absorbance only (10 principal components; λ=2e-3)	36.6	31.3	12.5	93.7	91.6
ABG only (λ=0)	5.7	79.6	92.5	97.1	98.8
Absorbance & ABG (10 principal components; λ=2e-3)	**5.2**	**80.9**	**97.1**	**97.5**	**99.2**

TABLE 2. Validation confusion matrix of the combined absorbance and ABG model (n=61000).

Error* 5.2%		**Predicted class**		
		SCD	**SF**	**Normal**
True class	**SCD**	8088	380	1532
	SF	322	10678	0
	Normal	950	0	39050

*Confusion matrix error was defined as one minus the proportion of correctly classified ears, i.e. the ratio of (1) the sum of the off-diagonal elements to (2) the sum of all elements.

DISCUSSION

We have developed a classification model that can automatically provide suggested diagnoses for individual ears by combining information from WBT and standard audiogram. It performs very well, with low error (5.2%), moderate to high sensitivity (SCD 80.9%, SF 97.1%), and high specificity (SCD 97.5%, SF 99.2%). While the ABG-only model performs surprisingly well, absorbance provides additional objective information on ear mechanics, further improving diagnosis. With our combined (absorbance & ABG) model, we can automate the analysis and interpretation of WBT and audiometric measurements at initial consultation by an audiologist. This would provide important diagnostic information that is not yet available for clinical decision making to improve patient care.

The strong performance of ABG-only model unexpectedly shows separation between SCD and SF (Fig. 2), providing support that ABG carries useful diagnostic information for such diagnoses. However, clinicians find that ABG alone is not sufficient for differentiating between SCD and SF, or sometimes between SCD and normal ears (Minor, 2005; Zhou, Gopen and Poe, 2007; Dewyer, Quesnel and Santos, 2020; Eberhard, Chari, et al., 2021).

The most common cause of SF is *otosclerosis*, a slowly progressing remodeling of bone causing fixation of the stapes footplate, thereby reducing sound energy to the cochlea (Hüttenbrink, 2003; Shahnaz, Longridge and Bell, 2009). Because of the slow process, ears with early-stage otosclerosis will have small ABGs and therefore resemble ears with SCD in Fig. 2 (Niesten et al., 2014; Quesnel, Ishai and McKenna, 2018; Eberhard, Chari, et al., 2021). For early stages of otosclerosis, absorbance may be helpful in differentiating SCD and SF because SCD and SF have opposite effects on absorbance, as illustrated in Fig. 1. SF increases mechano-acoustic input impedance observed from the ear canal, whereas SCD reduces input impedance (Hüttenbrink, 2003; Rosowski et al., 2004; Shahnaz, Longridge and Bell, 2009). The importance of absorbance in differentiating between SCD and early stages of otosclerosis may not be observable in our data. Our inclusion criteria required surgical confirmation of SF pathology, which may have precluded early stages of otosclerosis with small ABGs. For ears with ABGs that are similar between SF and SCD, absorbance would likely be critical in diagnosis.

The absorbance-only model classified 63% of ears correctly, and while sensitivities for SCD and SF were low, specificity was high (93.7% for SCD and 91.6% for SF). Because the absorbance-only model's specificity for SCD and SF is high, there is high certainty for having SCD or SF if this model predicts it. That means that in the clinic where ABG already provides information that the ear is abnormal, the absorbance-only model that identifies SCD or SF has high likelihood of being correct.

In future work, further improvements in prediction performance might be achieved by applying computational modeling techniques as in Merchant and Neely, 2021 and Masud, 2020 (Masud, 2020; Merchant and Neely, 2021). Additionally, separating low-frequency and high-frequency ABG effects could potentially improve performance as well as extend to differentiating other mechanical pathologies (Masud, 2020).

Limitations

In this study, our model only included two types of mechanical pathologies, SCD and SF. Adding additional pathologies (e.g. malleus fixation and ossicular discontinuity) to the model would be important for future studies.

There may be selection bias in the inclusion criteria because only SF that had surgery were included. Thus our cohort may have had larger ABGs than those with earlier stages of otosclerosis. Earlier stages of otosclerosis would have similar low-frequency ABG to SCD where absorbance would increase accuracy with our diagnostic method.

Conclusion

This study demonstrates that automation of diagnosis using both absorbance and ABG based on our regression method is feasible. Here we demonstrate the diagnostic ability to detect SCD and SF from normal ears. Because diagnosis of mechanical pathologies such as SF and SCD can have uncertainties, our diagnostic method is expected to improve patient care.

ACKNOWLEDGMENTS

We are grateful for the assistance of Michael Ravicz and Denis Fitzpatrick in preparation of the WBT data for analysis. We thank the otology surgeons at Massachusetts Eye and Ear (Drs. Alicia Quesnel, Bradley Welling, Daniel Lee, David Jung, Dunia Abdul-Aziz, Felipe Santos, and Ronald de Venecia) for their help identifying and recruiting subjects with SCD and SF. Also, we thank Dr. Stéphane Maison, Anita Mepani and Kelsie Grant for their assistance collecting data in normal hearing ears.

Research was supported by the National Institutes of Health through award numbers P20GM109023, R01DC008318, L30DC017300, and P50DC015857. Kristine Eberhard received a PhD fellowship from the William Demant Foundation.

REFERENCES

Dewyer, N.A., Quesnel, A.M. and Santos, F. (2020) ‘‘A Case Series of Patients With Concurrent Otosclerosis and Superior Semicircular Canal Dehiscence’’, Otology & Neurotology, 41(2), e172–e181. doi:10.1097/MAO.0000000000002487.

Eberhard, K.E., Chari, D.A., Nakajima, H.H., Klokker, M., Cayé-Thomasen, P. and Lee, D.J. (2021) ‘‘Current Trends, Controversies, and Future Directions in the Evaluation and Management of Superior Canal Dehiscence Syndrome’’, Frontiers in Neurology, 12(April). doi:10.3389/fneur.2021.638574.

Eberhard, K.E., Ravicz, M.E., Merchant, G.R., Masud, S.F., Maison, S.F., Neely, S.T. and Nakajima, H.H. (2021) ‘‘Preserving Wideband Tympanometry Information With Artifact Mitigation’’, Ear & Hearing, 43(2), 563–576. doi:10.1097/aud.0000000000001117.

Feeney, M.P., Keefe, D.H., Hunter, L.L., Fitzpatrick, D.F., Garinis, A.C., Putterman, D.B. and McMillan, G.P. (2017) ‘‘Normative Wideband Reflectance, Equivalent Admittance at the Tympanic Membrane, and Acoustic Stapedius Reflex Threshold in Adults’’, Ear and hearing, 38, e142–e160. doi:10.1097/AUD.0000000000000399.

Feeney, M.P., Keefe, D.H., Hunter, L.L., Fitzpatrick, D.F., Putterman, D.B. and Garinis, A.C. (2020) ‘‘Effects of Otosclerosis on Middle Ear Function Assessed With Wideband Absorbance and Absorbed Power’’, Ear & Hearing, 42(3), 547–557. doi:10.1097/aud.0000000000000968.

Feeney, M.P., Grant, I.L. and Marryott, L.P. (2003) ‘‘Wideband Energy Reflectance Measurements in Adults With Middle-Ear Disorders’’, Journal of Speech, Language, and Hearing Research, 46(4), 901–911. doi:10.1044/1092-4388(2003/070).

Fuse, T., Aoyagi, M., Koike, Y. and Sugai, Y. (1992) ‘‘Diagnosis of Ossicular Chain in the Middle Ear By High-Resolution Ct’’, ORL J Otorhinolaryngol Relat Spec., 54(5), 251–254. doi:10.3950/jibiinkoka.95.247.

Hunter, L.L., Bagger-Sjöbäck, D. and Lundberg, M. (2008) ‘‘Wideband reflectance associated with otitis media in infants and children with cleft palate’’, International Journal of Audiology, 47(SUPPL. 1), S57-61. doi:10.1080/14992020802294057.

Hüttenbrink, K.B. (2003) ‘‘Biomechanics of Stapesplasty: A Review’’, Otology and Neurotology, 24(4), 548–557. doi:10.1097/00129492-200307000-00004.

Liu, Y.-W., Sanford, C.A., Ellison, J.C., Fitzpatrick, D.F., Gorga, M.P. and Keefe, D.H. (2008) ‘‘Wideband absorbance tympanometry using pressure sweeps: System development and results on adults with normal hearing’’, The Journal of the Acoustical Society of America, 124(6), 3708–3719. doi:10.1121/1.3001712.

Masud, S.F., Knudson, I.M., Stankovic, K.M. and Nakajima, H.H. (2019) ‘‘Fracture of the Incus Caused by Digital Manipulation of the Ear Canal and its Diagnosis Using Wideband Acoustic Immittance’’, Otology & Neurotology, 40(2), e115–e118. doi:10.1097/MAO.0000000000002103.

Masud, S.F. (2020) ‘*Diagnosis of Mechanical Ear Pathologies Using Structure-Based Modeling and Machine Learning Techniques [dissertation]*’. Harvard University. Available at: https://dash.harvard.edu/handle/1/37365115.

Merchant, G.R., Roösli, C., Niesten, M.E.F., Hamade, M.A., Lee, D.J., McKinnon, M.L., Ulku, C.H., Rosowski, J.J., Merchant, S.N. and Nakajima, H.H. (2015) ‘‘Power reflectance as a screening tool for the diagnosis of superior semicircular canal dehiscence’’, Otology and Neurotology, 36(1), 172–177. doi:10.1097/mao.0000000000000294.

Merchant, G.R., Siegel, J.H., Neely, S.T., Rosowski, J.J. and Nakajima, H.H. (2019) ‘‘Effect of Middle-Ear Pathology on High-Frequency Ear Canal Reflectance Measurements in the Frequency and Time Domains’’, Journal of the Association for Research in Otolaryngology, 20(6), 529–552. doi:10.1007/s10162-019-00735-1.

Merchant, G.R., Al-Salim, S., Tempero, R.M., Fitzpatrick, D. and Neely, S.T. (2021) ‘‘Improving the Differential Diagnosis of Otitis Media With Effusion Using Wideband Acoustic Immittance’’, Ear & Hearing, 42(5), 1183–1194. doi:10.1097/aud.0000000000001037.

Merchant, G.R. and Neely, S.T. (2021) ‘‘The influence of otitis media with effusion on middle-ear impedance estimated from wideband acoustic immittance measurements’’, The Journal of the Acoustical Society of America, 150(2), 969–978. doi:10.1121/10.0005822.

Minor, L.B. (2005) ‘‘Clinical manifestations of superior semicircular canal dehiscence’’, Laryngoscope, 115, 1717–1727. doi:10.1097/01.mlg.0000178324.55729.b7.

Nakajima, H.H., Pisano, D. V., Röösli, C., Hamade, M.A., Marchant, G.R., Mafoud, L., Halpin, C.F., Rosowski, J.J. and Merchant, S.N. (2012) ‘‘Comparison of Ear-Canal Reflectance and Umbo Velocity in Patients With Conductive Hearing Loss: A Preliminary Study’’, Ear & Hearing, 33(1), 35–43. doi:10.1063/1.3658147.

Nakajima, H.H., Rosowski, J.J., Shahnaz, N. and Voss, S.E. (2013) ‘‘Assessment of ear disorders using power reflectance’’, Ear and hearing, 34(Suppl. 1), 48S-53S.

Niesten, M.E.F., Mckenna, M.J., Grolman, W. and Lee, D.J. (2012) ‘‘Clinical factors associated with prolonged recovery after superior canal dehiscence surgery’’, Otology and Neurotology, 33(5), 824–831. doi:10.1097/MAO.0b013e3182544c9e.

Niesten, M.E.F., Hamberg, L.M., Silverman, J.B., Lou, K. V., McCall, A.A., Windsor, A., Curtin, H.D., Herrmann, B.S., Grolman, W., Nakajima, H.H. and Lee, D.J. (2014) ‘‘Superior Canal Dehiscence Length and Location Influences Clinical Presentation and Audiometric and Cervical Vestibular-Evoked Myogenic Potential Testing’’, Audiology and Neurotology, 19(2), 97–105. doi:10.1159/000353920.

Quesnel, A.M., Ishai, R. and McKenna, M.J. (2018) ‘‘Otosclerosis: Temporal Bone Pathology’’, Otolaryngologic Clinics of North America, 51(2), 291–303. doi:10.1016/j.otc.2017.11.001.

Rosowski, J.J., Songer, J.E., Nakajima, H.H., Brinsko, K.M. and Merchant, S.N. (2004) ‘‘Clinical, experimental, and theoretical investigations of the effect of superior semicircular canal dehiscence on hearing mechanisms’’, Otology and Neurotology, 25(3), 323–332. doi:10.1097/00129492-200405000-00021.

Sanford, C.A., Hunter, L.L., Patrick Feeney, M. and Nakajima, H.H. (2013) ‘‘Wideband acoustic immittance: Tympanometric measures’’, Ear and Hearing, 34(Suppl. 1), 65S-71S. doi:10.1097/AUD.0b013e31829c7250.

Schimanski, G., Schimanski, E. and Berthold, M.R. (2011) ‘‘Diagnostic findings in stapes revision surgery - A retrospective of 26 years’’, Otology and Neurotology, 32(3), 373–383. doi:10.1097/MAO.0b013e3182096da1.

Shahnaz, N., Longridge, N. and Bell, D. (2009) ‘‘Wideband energy reflectance patterns in preoperative and post-operative otosclerotic ears’’, International Journal of Audiology, 48(5), 240–247. doi:10.1080/14992020802635317.

Sun, X.-M. (2016) ‘‘Wideband Acoustic Immittance: Normative Study and Test–Retest Reliability of Tympanometric Measurements in Adults’’, Journal of Speech, Language, and Hearing Research, 59(4), 819–34. doi:10.1044/2016.

Zhou, G., Gopen, Q. and Poe, D.S. (2007) ‘‘Clinical and diagnostic characterization of canal dehiscence syndrome: A great otologic mimicker’’, Otology and Neurotology, 28(7), 920–926. doi:10.1097/MAO.0b013e31814b25f2.

Modelling the Recovery of Residual Acoustic Hearing after Cochlear Implantation by Using Feasible Intracochlear Acoustic Devices

David Slater,[a] Stephen Elliott,[b] and Carl Verschuur[c]

(Institute of Sound and Vibration Research, University of Southampton, University Road, Southampton SO17 1BJ, England.)

[a] *Corresponding author: dhs1e17@soton.ac.uk*
[b] *Electronic mail: sje@soton.ac.uk*
[c] *Electronic mail: c.a.verschuur@soton.ac.uk*

Abstract. Cochlear implants are now used to enhance the high frequency hearing of patients that still have significant residual acoustic hearing at low frequencies. The hybrid combination of electrical and acoustic hearing can give significant improvements in speech perception, compared with just one or the other. The process of cochlear implantation, however, typically degrades the low frequency residual acoustic hearing. To study the mechanical effects of cochlear implantation on residual acoustic hearing, a two-chamber model of the cochlea has been developed that enables the upper and lower chambers to be modelled with different physical characteristics, so that the effects of an implant in the lower chamber can be predicted. In particular, the model has been used to predict the reduction in residual acoustic hearing caused by the stiffening of the round window, due to its penetration by the electrode array of an implant. In the frequency range 100 – 1000 Hz, the predicted hearing loss is of the order of 20 dB, which is consistent with the average post-operative hearing losses measured previously [1] The new model has been used to show that the loss can be reduced substantially by incorporating a gas bubble of a few cubic millimetres near the round window; the bubble then fulfils the pressure release function of the normal round window. The effect on the residual acoustic hearing of changing the size and position of the bubble will be discussed. The model has also been used to simulate the intracochlear stimulation that could be achieved by incorporating an acoustic actuator within a cochlear implant, as an alternative to the high-powered external hearing aids that are currently used. To study the feasibility of designing an actuator with sufficient output to overcome the patient's low frequency hearing loss, yet that is small enough to fit within an electrode array, both its acoustic output and its internal acoustic impedance need to be considered. At least two types of actuator appear to be feasible in this application. Combined intracochlear electrical acoustic excitation opens up the possibilities of both ipsilateral and bilateral hybrid hearing.

INTRODUCTION AND MOTIVATION

Introduction

A cochlear implant, CI, is now an established procedure that is approved by health regulators in most countries, including NICE in the UK, for patients with profound or severe hearing loss.

Generally, cochlear implantation enables hearing of speech and its recognition for those who had not been able to do so. However, it can also impair low frequency hearing. One cause of such impairment is the stiffening of the round window (RW), thereby reducing its ability to allow the movement of cochlear fluid at frequencies below 1 kHz. This can cause a hearing loss up to 30 dB, which diminishes the ability to hear traffic noise, music and non-verbal sounds that are helpful to speech recognition and understanding.

The aim of the work has been to improve the experience of implanted patients that have suffered low frequency hearing loss as a result of RW stiffening.

Approach and Method

To fulfill the aim of the work, it was necessary first to confirm the magnitude of the hearing loss due to the stiffening, and then to assess the extent of mitigation of the loss by various remedies, with a view to finding an achievable remedy that would significantly reduce the loss or eliminated it. (Hearing loss occurring after implantation can result from a variety of causes [2].)

The approach taken has been that of deriving a model of the cochlea from well-established first principles. This gives assurance that the model can be used to cover different circumstances, without need to produce a new derivation to suit the circumstances of each use.

The model has two chambers, bounded at the basal end with windows and aqueducts whose impedances can be set independently of each other. model was implemented in MatLab and then verified and validated as follows:

- The model was loaded with data for two identical chambers and windows, and the results were again compared with those of an established single chamber model [3].
- Comparisons with established lumped component models have been used to verify the model [4] and [5].
- It was also loaded with the input data used in an early two-chamber model [6], and the results compared.
- Some validation was obtained by comparing the model's results with measured values. The only measurements found that could be compared with modelled results were the position along the cochlea of maximum Basilar membrane (BM) vibration, the cochlea input impedance and the impedance of the RW.

The validated model has been used to explore the effects of various potential remedies, including:

- incorporating a small compliant bubble within the implant, as a substitute for the compliance of the RW [4], and
- calculating the output velocity and pressure that could be expected from various actuators if they could be made small enough to fit within an implant.

THE MODEL

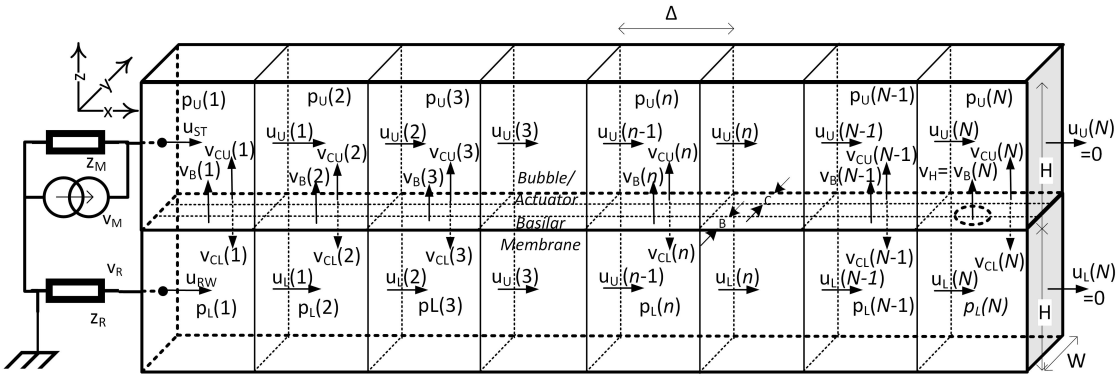


FIGURE 1. The Elemental Variables

The developed model is one of an uncoiled cochlea, with a uniform cross sectional chamber area and a uniform BM width. The length of the cochlea has been discretized into 350 similar elements, each one of which can contain a compressible element (such as a bubble) and a velocity source (such as the middle ear at the base or an actuator along the cochlea). The second elements from the base in the upper and lower chambers are used to model the vestibular and cochlear aqueducts respectively. The compressible elements and excitation sources in one chamber can be different from those in the corresponding element of the other chamber. Each element of the BM can have a source of excitation, which could be used for future modelling of otoacoustic emissions and for comparison with certain previous single chamber models. Each element of the passive BM is represented by a series spring, mass and damper impedance that is a function of the distance of the element from the base and excitation frequency. The cross sections of the two chambers are the same. The cochlear fluid is considered to be incompressible throughout this paper; compressibility can be modelled readily but makes virtually no difference below 7 kHz.

Mathematically, the model uses the law of conservation of mass to relate the BM acoustic velocity into the elemental volume to the net flow out of it (to the compressible element, if any and to the adjacent elements). For the upper chamber, the conservation of mass equation expresses, for each elemental volume, n, the BM acoustic velocity (in

m³/s), $q_B(n)$ as a function of the acoustic velocities along the cochlea, $q_U(n-1), q_U(n)$ and the chamber pressure, $p_U(n)$ in the elemental volume:

$$q_B(n) = q_U(n) - q_U(n-1) - q_{EU}(n) + Y_{CU}(n)p_U(n), \tag{1}$$

where $q_{EU}(n)$ is the unloaded acoustic velocity of the elemental excitation source, if any, in the compressible element, and $Y_{CU}(n)$ is the associated acoustic admittance.

The law of conservation of momentum equates the force applied to the fluid in the elemental volume to the product of its mass and its acceleration along the cochlea. The force is equal to the product of the pressure difference between the entry and exit of the elemental volume, which is $p_U(n-1) - p_U(n)$ and the effective cross-sectional area of the chamber, A_{CH}. The force is also equal to the product of the acceleration of the fluid and the fluid mass, which is $\rho A_{CH}\Delta$; where Δ is the length of each element, measured along the cochlea and ρ is the density of the fluid, which is assumed to be constant when compressibility is not modelled. Because all of the time varying quantities are assumed to be sinusoidal, the mechanical acceleration is $i\omega q_U(n)/A_{CH}$. This results in the following equations:

$$(p_U(n) - p_U(n+1))A_{CH} = \frac{i\omega q_U(n)}{A_{CH}}\rho A_{CH}\Delta = i\omega u_U(n)\rho\Delta \tag{2}$$

$$\frac{q_U(n) - q_U(n-1)}{A_{CH}} = \frac{(-p_U(n-1) + 2p_U(n) - p_U(n+1))A_{CH}}{i\omega\rho\Delta} \tag{3}$$

Noting that $A_{CH} = Bh$, where B is the Basilar membrane width, and h is the effective height of the chamber. This equation can now be substituted into equation (1), and acoustic quantities converted to mechanical quantities by dividing by the BM elemental area, $B\Delta$, to give:

$$\frac{q_B(n)}{B\Delta} = \frac{(-p_U(n-1) + 2p_U(n) - p_U(n+1))Bh}{i\omega\rho\Delta B\Delta} - \frac{q_{EU}(n) - Y_{CU}(n)p_U(n)}{B\Delta} = v_B(n). \tag{4}$$

Where $v_B(n)$ is the mechanical BM velocity, positive when upwards.

$$\text{Or } \frac{(-p_U(n-1) + 2p_U(n) - p_U(n+1))h}{i\omega\rho\Delta^2} + y_{CU}(n)p_U(n) = v_B(n) + v_{EU}(n) \tag{5}$$

Where $y_{CU}(n)$ is the mechanical admittance of the compressible element, and $v_{EU}(n)$ is the velocity of the excitation source within the compressible element. For simplicity, the width of the compressible element is the same as that of the BM. In all of the above, the index, n is the number of the finite element in the discretized model of the chamber, when $n > 0$; when $n = 0$, the indexed variable is a property of the middle ear for the upper chamber and the round window for the lower chamber. In the matrix equations that follow each of the variables $\mathbf{p}_U, \mathbf{v}_B$, and $\mathbf{v}_{EU}$, together with the corresponding lower chamber variable, is a column vector, for $n > 0$. stapes excitation, an additional column vector, $\mathbf{v}_M$ of the same length is required; this will have the $n = 0$ value, v_M as its first element and zeros for the remainder. This boundary vector can then be added to the other velocity vectors in the following matrix equations.

$$(\mathbf{Y}_{FC} + \mathbf{Y}_{CU} + \mathbf{Y}_M)\mathbf{p}_U = \mathbf{v}_B + \mathbf{v}_{EU} + \mathbf{v}_M, \text{ or} \tag{6}$$

$$\mathbf{p}_U = (\mathbf{Y}_{FC} + \mathbf{Y}_{CU} + \mathbf{Y}_M)^{-1}(\mathbf{v}_M + \mathbf{v}_{EU} + \mathbf{v}_B), \text{ where:} \tag{7}$$

$$\mathbf{Y}_{FC} = \mathbf{u}_U/\mathbf{p}_U$$

$$\mathbf{Y}_{FC} = \frac{h}{i\omega\rho\Delta^2}\begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix} \tag{8}$$

$\mathbf{Y}_{CU}$ is an n by n matrix, with y_{CU} on the principal diagonal and zeros elsewhere,

$\mathbf{y}_M$ is the reverse middle ear admittance, and

$\mathbf{Y}_M$ is an n by n matrix, with y_M in the first column of the first row and zeros elsewhere. (9)

The corresponding equation for the lower chamber is:

$$\mathbf{p}_L = (\mathbf{Y}_{FC} + \mathbf{Y}_{CL} + \mathbf{Y}_R)^{-1}(\mathbf{v}_{EL} - \mathbf{v}_B), \text{ where:} \quad (10)$$

$\mathbf{Y}_{CL}$ is an n by n matrix, with $y_{CL}(n)$ on the principal diagonal and zeros elsewhere,

$\mathbf{y}_R$ is the round window admittance, and

$\mathbf{Y}_R$ is an n by n matrix, with y_R in the first column of the first row and zeros elsewhere. (11)

The equations that give $\mathbf{p}_U$ and $\mathbf{p}_L$ can be subtracted to give the pressure difference, $\mathbf{p}_D$ between the two chambers, which is simply related to $\mathbf{v}_B$.

$$\mathbf{p}_D = (\mathbf{Y}_{FC} + \mathbf{Y}_{CU} + \mathbf{Y}_M)^{-1}(\mathbf{v}_M + \mathbf{v}_{EU} + \mathbf{v}_B) - (\mathbf{Y}_{FC} + \mathbf{Y}_{CL} + \mathbf{Y}_R)^{-1}(\mathbf{v}_{EL} - \mathbf{v}_B) = \mathbf{Z}_B(\mathbf{v}_{EB} - \mathbf{v}_B) \quad (12)$$

Where $\mathbf{Z}_B$ is an n by n matrix, with $z_B(n)$ on the principal diagonal and zeros elsewhere,

$$\mathbf{Z}_B(\mathbf{v}_{EB} - \mathbf{v}_B) = (\mathbf{Y}_{FC} + \mathbf{Y}_{CU} + \mathbf{Y}_M)^{-1}(\mathbf{v}_M + \mathbf{v}_{EU} + \mathbf{v}_B) - (\mathbf{Y}_{FC} + \mathbf{Y}_{CL} + \mathbf{Y}_R))^{-1}(\mathbf{v}_{EL} - \mathbf{v}_B) \quad (13)$$

$$\text{If } \mathbf{Z}_U = (\mathbf{Y}_{FC} + \mathbf{Y}_{CU} + \mathbf{Y}_M)^{-1}, \text{ and} \quad (14)$$

$$\mathbf{Z}_L = (\mathbf{Y}_{FC} + \mathbf{Y}_{CL} + \mathbf{Y}_R)^{-1}, \text{ then} \quad (15)$$

$$\mathbf{v}_B = (\mathbf{Z}_U + \mathbf{Z}_B + \mathbf{Z}_L)^{-1}(\mathbf{Z}_L\mathbf{v}_{EL} - \mathbf{Z}_U(\mathbf{v}_{EU} + \mathbf{v}_M) + \mathbf{Z}_B\mathbf{v}_{EB}) \quad (16)$$

Usually, there will be no excitation sources in any compressible elements in the upper chamber, and if otoacoustic emissions are not considered, then $\mathbf{v}_{EB} = \mathbf{v}_{EU} = 0$. This simplifies equation (16) to become:

$$\mathbf{v}_B = (\mathbf{Z}_U + \mathbf{Z}_B + \mathbf{Z}_L)^{-1}(\mathbf{Z}_L\mathbf{v}_{EL} - \mathbf{Z}_U\mathbf{v}_M) \quad (17)$$

AVOIDING THE HEARING LOSS

Two alternative approaches have been investigated for mitigating acoustic hearing loss caused by stiffening of the RW:

- incorporating a small bubble into the cochlear implant, and
- incorporating a miniaturized actuator into the cochlear implant, to provide intracochlear excitation.

A range of different bubble volumes and positions along the lower chamber have been investigated, to find the optimum distance from the RW and the relationship between bubble volume and hearing loss. Eight actuator technologies have also been studied, to investigate the feasibility of achieving the required acoustic output from a device sufficiently miniaturized to fit within an implant.

Bubble Performance

Figure 2 shows the predicted BM velocity, referenced to unit ear canal pressure, both as a function of position along the cochlea when exited at a single frequency of 500Hz and as a function of frequency when calculated at a single position along the cochlea, at 25mm. Results are shown with a normal RW stiffness, and when the RW is stiffened by a factor of 170, with no bubble or with a bubble of different volumes adjacent to the RW in the lower fluid chamber. At this frequency the peak BM response is reduced by about 15dB when the RW is stiffened but restored to almost its original level when even the smallest bubble, of 1.5mm³, is introduced. By running a series of simulations at different frequencies, the change in the peak BM response after the RW has been stiffened, corresponding to the hearing loss, has been calculated and is plotted in Figure 3 as a function of frequency for various bubble volumes. These results are also used to calculate the hearing loss as a function of the bubble volume at several frequencies in Figure 4.

It is clear from Figure 3 that a bubble with a volume of 15 mm³ adjacent to the RW will eliminate virtually all hearing loss due to RW stiffening, but such a bubble would be too large to fit into an implant. Locating the bubble away from the RW makes little difference to the residual hearing loss in the frequency range of interest (100 - 1000 Hz). Even a 1.5 mm³ bubble, however, would leave no more than 3 dB of hearing loss at frequencies above 100 Hz. (Frequencies below this do not contribute much to speech recognition, the hearing of traffic, or the appreciation of music.)

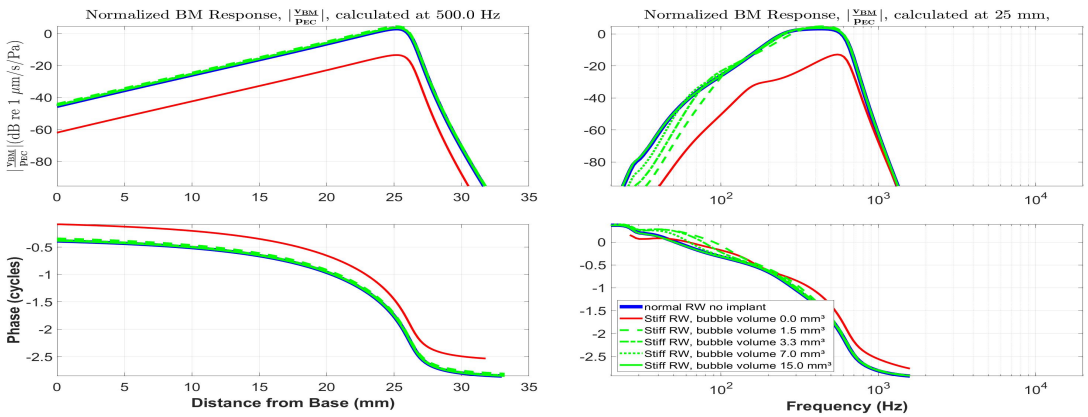


FIGURE 2. BM velocity for a normal RW (blue line), a stiffened RW (red line) and a stiffened RW with bubbles of four volumes (green lines) in the spatial domain at 500 Hz and the frequency domain at 25 mm.

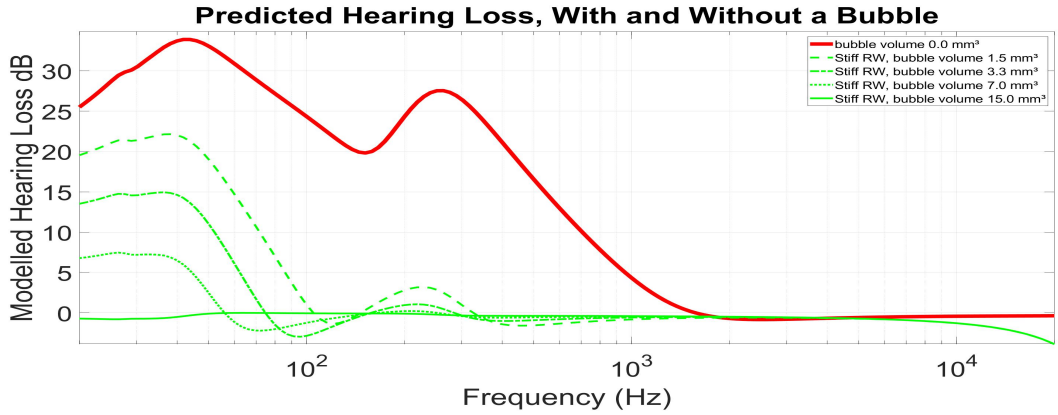


FIGURE 3. Hearing loss as a function of frequency, with bubble volume as a parameter. Hearing loss is calculated for each frequency as the difference in dB between the BM response at the position where it is a maximum (ie best place) for a normal ear and for an ear with a stiffened RW.

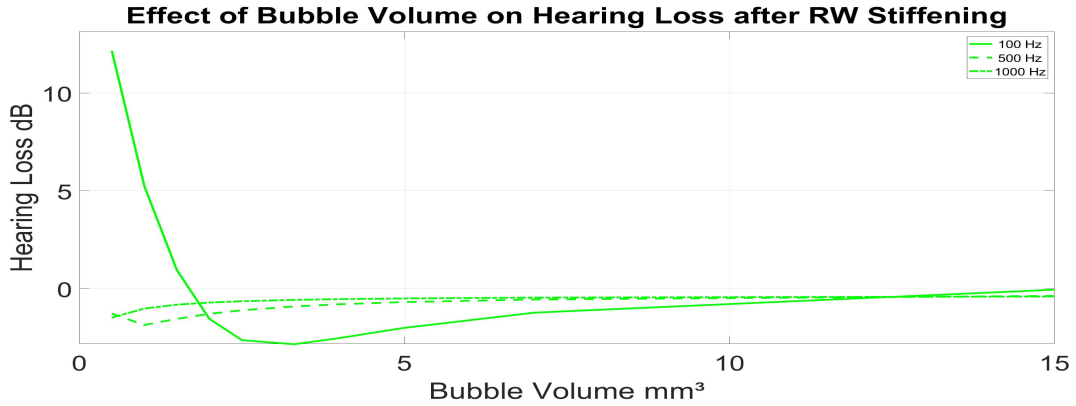


FIGURE 4. Hearing loss due to a stiffened RW, as a function of bubble volume, with frequency as a parameter.

Actuator Performance

An actuator needs to be capable of producing sufficient acoustic velocity and pressure to achieve the greatest level of sensation that is reached when listening with a normal ear to speech, music or traffic noise, without excessive distortion. The model above has been used to calculate the requirements for such an actuator, which are shown in Table I.

The feasibilities of a number of actuator technologies have been evaluated in terms of their ability to meet these requirements. The result of the calculations are summarized in Table II, which shows that several technologies appear feasible, if suitable sized actuators could be fabricated.

TABLE I. Velocity, displacement and pressure requirements of an intracochlear actuator, together with the cochlear impedance at the actuator. The maximum value of each variable is shown in **Bold**. The conversion from linear (mechanical) to volume (acoustic) velocity uses an diaphragm area of 15×0.32=4.8 mm².

f	**Linear velocity**	**Volume velocity**	**Linear Displacement**	**Required pressure**	**Cochlear Impedance**
Hz	mm/s	mm^3/s	μm	kPa	GΩ
100	**2.5**	**12**	**4**	0.95	77
152	1.9	9.1	2	**1.3**	148
159	1.8	8.8	1.8	1.3	**151**
999	1.8	8.9	0.3	0.66	75

TABLE II. Feasibilities of the considered technologies.

Actuator Technology	**Feasible**	**Note**
Balanced armature with permanent magnets	Yes	1
Ceramic and polymeric piezoelectric	No	2
Polymeric piezoelectric constrained dome	Yes	3
MEMS xylophone (cantilever)	No, but	4
Attracted armature without permanent magnet	No	5
Moving coil	No	6
Magnetostrictive	No	7
Externally excited permanent magnet	Yes, but	8
For an explanation of 'but' see associated Note.		

Notes on Table II

1. Balanced armature actuators are established technology, which would be feasible, provided both that they can be fabricated at the required scale and that the resultant electrode array can be satisfactorily inserted through the RW, when containing a rigid device that will be placed between 1 and 16 mm from the RW.

2. Linear piezoelectric actuators are not feasible, because the electric field required exceeds the electric breakdown strength of suitable piezoelectric material.

3. Piezoelectric actuators appear to be feasible if a thin film of the material can be formed into a longitudinally constrained dome, whose longitudinal elongation is limited so that a longitudinal electric field causes amplified transverse movement. This technology is feasible provided both that the dome can be adequately constrained, and that the electrode array can be inserted without causing damage to either the cochlea or the dome.

4. MEMS cantilever actuators designed as microphones are probably not feasible as actuators, because of the rather excessive voltage requirements (c. 300 V). However, the proven fabrication technique [7] of this type of microphone could be beneficially investigated for use in combination with the constrained piezoelectric dome technology.

5. Balanced armature actuators without permanent magnets are not feasible for this purpose, because of the excessive current requirement that results in a coil that would not fit into an electrode array.

6. Moving coil actuators are not feasible, because the axial length of the coil would need to be too large to fit transversely into an electrode array.

7. Magnetostrictive actuators are not feasible, because they require excessive currents.

8. An actuator composed of a permanent magnet within an electrode array excited by a coil attached to the scalp may be also feasible, provided that an electrode array containing a rigid magnet from 1 to 16 mm from the RW could be inserted without causing damage to the cochlea.

Figure 5 shows the BM velocity predicted by the two-chamber model, when driven at the stapes, with or without a stiffened RW as above, or when driven by an actuator having the same volume velocity as the stapes driving the lower fluid chamber. The actuator is either assumed to act at a single point, at different locations along the CI, or in a distributed manner from 1mm to 16mm along the length of the CI. It can be seen that such an actuator is predicted to drive the BM at the same or a greater extent than with stapes excitation.

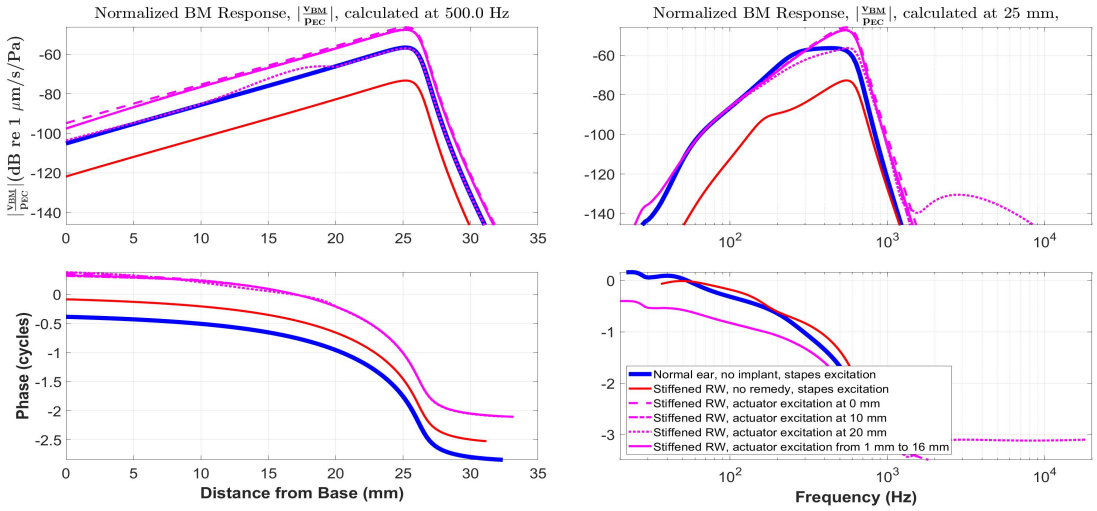


FIGURE 5. Normalized Basilar membrane linear velocity response predicted at 500 Hz in the spatial domain and at 25 mm in the frequency domain for a point actuator just one finite element long at each of three positions and for a 15 mm long actuator. For comparison, graphs are also plotted for a normal ear and for an implanted ear, each with excitation from the stapes. The excitation volume velocity from the actuator is the same as that from the stapes when the ear canal sound pressure is 115 dB above Normal Hearing Level (BS EN ISO 389-7 :2019).

CONCLUSIONS

The model has predicted that RW stiffening will cause hearing loss that is commensurate with measured results. It has also shown that there are feasible remedies that will mitigate or eliminate that loss. In particular, it has shown that a bubble that is small enough to fit within an implant is predicted to reduce the maximum loss to about 3 dB within the frequency range of interest (100 to 1000 Hz), Also, if they could be suitably miniaturised it is predicted that at least two actuator techniques could drive the BM at sensation levels suitable for listening to speech, music and traffic noise.

REFERENCES

1. C. Verschuur, W. Hellier, and C. Teo, "An evaluation of hearing preservation outcomes in routine cochlear implant care: Implications for candidacy," Cochlear implants international **17**, 62–65 (2016).
2. A. Causon, C. Verschuur, and T. A. Newman, "A retrospective analysis of the contribution of reported factors in cochlear implantation on hearing preservation outcomes," Otology & Neurotology **36**, 1137–1145 (2015).
3. S. J. Elliott and G. Ni, "An elemental approach to modelling the mechanics of the cochlea," Hearing research **360**, 14–24 (2018).
4. S. J. Elliott, G. Ni, and C. A. Verschuur, "Modelling the effect of round window stiffness on residual hearing after cochlear implantation," Hearing Research **341**, 155–167 (2016).
5. L. Xue, H. Liu, W. Wang, J. Yang, Y. Zhao, and X. Huang, "The role of third windows on human sound transmission of forward and reverse stimulations: A lumped-parameter approach," The Journal of the Acoustical Society of America **147**, 1478–1490 (2020).
6. L. Peterson and B. P. Bogert, "A dynamical theory of the cochlea," The Journal of the Acoustical Society of America **22**, 369–381 (1950).
7. C. Zhao, *Ultraminiature Piezoelectric Implantable Acoustic Transducers for Biomedical Applications*, Ph.D. thesis, Mechanical Engineering, The University of Michigan (2018).

A graph signal processing model of the cochlea

Melia E. Bonomo,[1, a)] Santiago Segarra,[2, b)] and Robert M. Raphael[1, c)]

[1)]*Department of Bioengineering, Rice University, 6100 Main St, Houston, Texas 77005, USA.*
[2)]*Department of Electrical and Computer Engineering, Rice University, 6100 Main St, Houston, Texas 77005, USA.*

[a)]*Corresponding author: mbonomo@rice.edu*
[b)]*Electronic mail: segarra@rice.edu*
[c)]*Electronic mail: rraphael@edu.edu.*

Abstract.
We take a novel graph approach to study the complex cochlear mechanics during music encoding for patients with normal hearing and hearing loss. Audiograms from AudGenDB, a pediatric clinical database, are used in tandem with the UR Ear 2020b simulation of inner hair cell voltage responses to different sound stimuli. The music stimuli that we test consisted of single pitches, single timbres, pitch combinations (chords), timbre combinations, and song excerpts. Graph signal processing is used to model the inner hair cell responses of each patient as a new data structure that we term a "cochlea graph." By calculating various graph-theoretic properties, we show that these cochlea graphs carry distinguishing information based on a patient's hearing loss diagnosis. Our results are poised for application to a data science approach for improving cochlear implant signal processing of music, in which these music-encoding cochlea graphs can be used to develop individualized signal processing solutions based on a user's underlying hearing health.

INTRODUCTION

The cochlea is a spiral-shaped organ responsible for converting acoustic information into neural signals in the peripheral auditory system [1, 2]. Briefly, acoustic vibrations coming into the cochlea cause a traveling wave to move through the Organ of Corti, which contains approximately 12,000 outer hair cells and 4,000 inner hair cells (IHC). Often compared to a piano, the cochlea exhibits linear tonotopy: there is a one-dimensional spatial arrangement of positions where different frequencies of the incoming sound are processed along the cochlear spiral [3]. In other words, high-frequency components cause peak excitation in sensory cells near the base of the cochlea, low-frequencies cause peak excitation in cells near the apex, and mid-range frequencies cause peak excitation along a logarithmic spacing of cells in between. The excited IHCs release neurotransmitters to about 30,000 auditory nerve fibers, which then relay their signals to the central auditory system via the vestibulocochlear nerve.

For those who are deaf or severely hard-of-hearing due to hair cell damage, the linear tonotopy along the cochlear spiral has been the mathematical basis for multi-channel cochlear implant design and surgical placement [4]. The device filters incoming sound into frequency channels that excite 12–22 electrodes in a configuration that loosely mimics the tonotopy of the undamaged cochlea. The electrodes then each stimulate a group of underlying auditory nerve fibers. There is inevitably a large reduction in the frequency resolution due to device hardware limitations, which is one of the major hinderances for processing complex auditory signals such as music [5]. Additionally, individual differences in hearing performance for cochlear implant users are not well understood but are thought to be related to differences in underlying neural health. There is a need for improved music processing in cochlear implants. A data-based approach that incorporates individual users' hearing health with a representation of cochlear mechanics that is more nuanced than linear tonotopy has high potential.

In working towards that goal, here we report on a graph signal processing model for encapsulating a patient's hearing health information and the auditory biophysics of music processing into a new data structure that we term a patient's "cochlea graph." A graph in biology is a network comprised of units (nodes) that are linked to one another by means of anatomical connections or based on relationships in the functional activity of different units. The set of links between nodes define the graph's topology. Signals on a graph are a set of values associated with the nodes, and thus graph signal processing is the application of concepts and techniques from classical signal processing to these graphs [6]. Furthermore, graph-learning methods can be used to determine the topology of a graph based on the observation of signals on the nodes [7]. In our cochlea graphs, we designate IHCs as nodes, the sets of IHC voltage responses to various musical stimuli are signals, and the links between nodes are learned based on the voltage signals.

Though we obtain the IHC voltages using an existing model of the human auditory periphery that incorporates linear tonotopy [8], our model encodes a more complex representation of cochlear mechanics (Figure 1). Indeed, the one-dimensional, piano-likeness of the cochlea appears in the IHC responses, the second layer in our model. The first layer contains information about how healthy or damaged each IHC node in the model is, thereby influencing how the

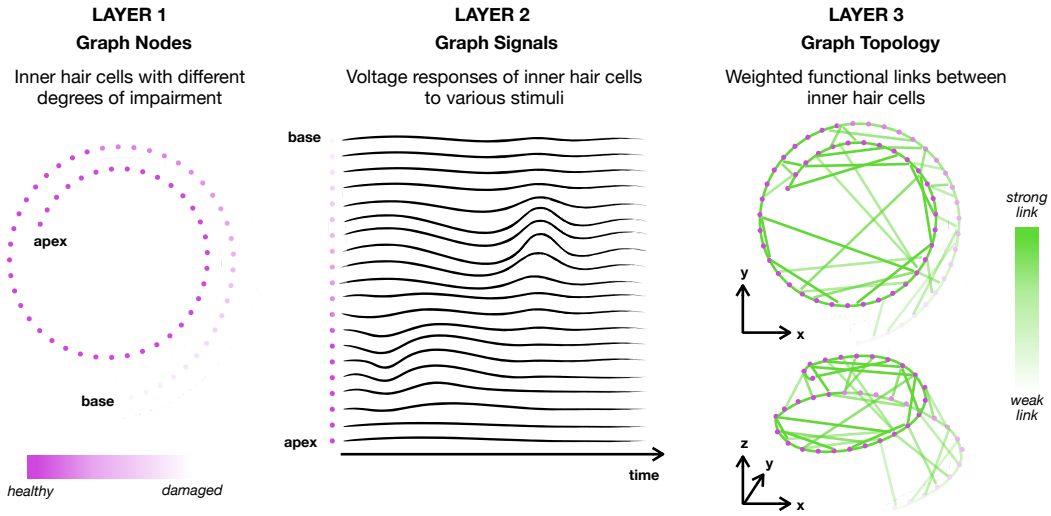


FIGURE 1. Cartoon representation of the different layers of cochlear mechanics in our graph signal processing model. Layer 1: The foundational layer is comprised of the inner hair cells, which each have a certain amount of impairment based on how healthy or how damaged they are. These cells are considered nodes of our "cochlea graph." Layer 2: Based on linear tonotopy and on the degree of impairment of the hair cells as defined in Layer 1, the second layer contains the voltage responses of these cells to a collection of stimuli. The average voltages are considered the graph signals. Layer 3: Based on the signals in Layer 2, the last and most complex layer of the model is the learned connections between nodes, which are links with varying weights. This produces the topology of our cochlea graph and illustrates the nonlinear functional relationships between hair cells.

IHCs will each respond to stimuli in the second layer. The third and final layer of cochlear mechanics in our model is a network of links between IHC nodes that is nonlinear, i.e., three-dimensional across the cochlear spiral. The links are learned based on how a collection of stimuli are processed in the second layer. Therefore, the complete graph signal processing model embodies IHC health, linear tonotopy, and new information about multi-dimensional sound encoding.

We constructed cochlea graphs for patients with normal hearing, conductive hearing loss, sensorineural hearing loss, or both types of hearing loss from AudGenDB, a clinical audiology database [13]. Audiograms from these patients were input to the UR Ear 2020b simulation of human sound encoding [8, 9, 10] with various musical stimuli that tested different pitch and timbre mixtures. The output IHC voltage responses were analyzed in a graph signal processing toolbox [11] to learn the links between IHCs. We found quantitative differences in the features of the resulting cochlea graphs based on hearing diagnoses, including that the graphs of patients with more severe hearing loss had more links, but these links had lower weights than the graphs of patients with normal hearing or less severe hearing loss. Importantly, we have shown that these cochlea graphs, constructed from the cochlear response to musical stimuli, carry information about a patient's hearing loss. This data structure has the potential to inform applications for improving music processing in cochlear implants.

METHODS

Audiograms from de-identified patients with normal hearing and diagnosed hearing loss were extracted from the AudGenDB database [13]. This database is HIPAA compliant and maintained by the Children's Hospital of Philadelphia. It contains data from electronic health records, audiological instruments, radiological imagery, clinical genetics results, and genomics research records. We selected all audiograms from patients that were complete (i.e., having responses for at least 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz) and performed in air (i.e., as opposed to other audiometry test conditions), and then we randomly sampled 400 audiograms across four diagnosis groups (Table I): normal hearing (NH), only conductive hearing loss (HL), only sensorineural HL, or both conductive and sensorineural HL. For conductive HL, sound conduction is impeded through the external ear, the middle ear, or both. For sensorineural HL, there is a problem in the cochlea, e.g., damaged cells, or the neural pathway to the auditory cortex.

Decibel loss across the cochlea was calculated by loading each audiogram into a fitting function in the UR Ear

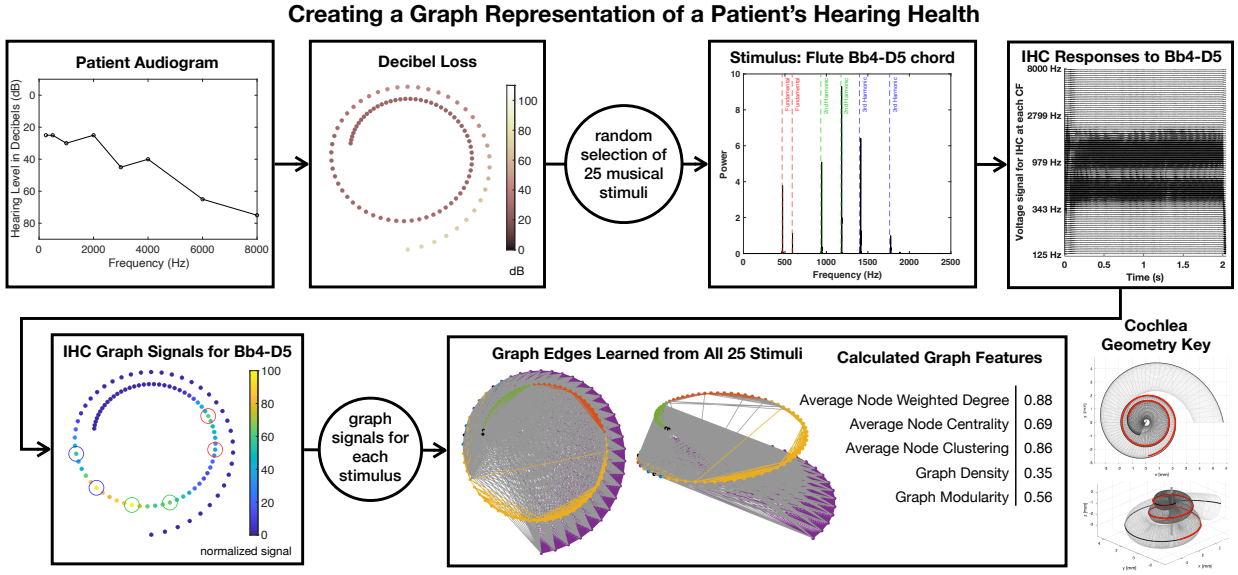


FIGURE 2. Analysis pipeline with the processing of an example patient with both types of HL shown. Patient audiograms from AudGenDB were put into the UR Ear 2020b toolbox [8, 9, 10] to determine the decibel loss for inner and outer hair cells along 100 positions along the cochlea. Each position is associated with a characteristic frequency (CF) and corresponds to an inner hair cell (IHC) node in our graph. For each of 25 randomly selected stimuli from a pool containing single instrument notes, single instrument chords, multi-instrument chords, and song excerpts, IHC voltage responses were determined using the UR Ear 2020b. In this example, a two-note major flute chord is played with fourth octave B-flat and fifth octave D. The stimulus power spectrum shows that there are multiple harmonics for both notes. The voltage traces for each IHC are offset on the y-axis according to the IHC's CF. Graph signals were determined by averaging the voltage time series. In this example with the flute chord, we see heightened signals for the IHCs that are responding to the harmonics, which are circled and color-coded as they are labelled in the stimulus power spectrum. The graph signals from all 25 stimuli were then put into the GSPBox toolbox [11] to learn weighted links between the 100 nodes, and graph features were calculated for each graph and compared across patients. The geometry of the cochlea in the Key was based on measurements from [12].

TABLE I. Information about the patients randomly sampled from AudGenDB in each hearing diagnosis group. HL: hearing loss. Native: Native American, Native Alaskan, or Native Hawaiian. NaN: race/ethnicity information was not available in the database.

	Normal Hearing	Conductive HL	Sensorineural HL	Both Types of HL	Total
Number	100	100	100	100	400
Age at Encounter (yrs)	9.1 ± 0.7	9.0 ± 0.4	10.9 ± 0.6	10.4 ± 0.4	9.1 ± 0.3
Sex					
Female	45	47	52	39	183
Male	55	53	48	61	217
Race					
White	61	75	54	72	262
Black	16	9	25	11	61
Asian	1	2	4	5	12
Native	0	0	1	1	2
Other/Multiple	11	10	10	7	38
NaN	11	4	6	4	25
Ethnicity					
Hispanic/Latinx	7	8	9	8	32
Not Hispanic/Latinx	66	77	71	70	284
NaN	27	15	20	22	84

2020b MATLAB toolbox [8, 9, 10]. This function computes outer and inner hair cell impairment factors based on the hearing level recorded at each frequency tested in the audiogram. We studied graphs with 100 nodes, which were comprised of 100 IHCs whose characteristic frequencies were evenly log spaced between 125 Hz and 8,000 Hz.

The stimuli used in the model were musical notes, chords, and song excerpts. We acquired 255 notes across four instruments (flute, trumpet, violin, piano) from the Musical Instrument Samples dataset [14], and we remixed these to create 461 dyad (two-note) chords played by one type of instrument and 10,212 triad (three-note) chords played by any combination of the available instruments. The chords were either major or minor chords with frequency tuning ratios of 4:5:6 or 10:12:15, respectively. Additionally, 472 two-second excerpts were clipped from five orchestral songs: The Orchestra Improvisation, Emotional Soundtrack, Classical Kanon, and Beautiful Flute Melody and Strings by the M33 Project and Vivaldi's Concerto in C Major for Oboe and Orchestra andante by the Advent Chamber Orchestra. This led to 11,400 generated musical stimuli. We randomly sampled 25 stimuli to test on each patient such that 5% of the tested stimuli were single pitch, single timbre (individual note); 10% were harmonic pitch mixtures, single timbre (two-note chords played by one instrument); 55% were harmonic pitch mixture, timbre mixture (three-note chords played by multiple instruments); and 30% were complex pitch and timbre mixtures (song excerpts).

The IHC model of the UR Ear 2020b toolbox, which simulates the biophysics of sound processing in the inner ear, was used to compute hair cell voltage responses to each musical stimulus. The average of each voltage time series was assigned as the graph signal of each node, and all signals were normalized between 0 and 100 to improve the signal to noise ratio. We assumed that the signals changed smoothly across adjacent nodes [15] and used the Graph Signal Processing Toolbox (GSPBox) [11] to learn the weighted graph links between IHC nodes from the smoothness in their signals. To visualize our resulting cochlea graphs, we used measurements from a three-dimensional anatomical model of the cochlear geometry determined by means of corrosion casting [12].

The NetworkX [16] Python package was used to calculate various network features of the resulting cochlea graph for each patient. These included: degree centrality, the fraction of nodes each node is connected to, averaged over all nodes; average clustering coefficient, the fraction of possible triangles through a node that exist, averaged over all nodes; average weighted degree, the sum of weights of links connected to a node, averaged over all nodes; and graph density, the fraction of all possible links that exist. Community detection, or the partition of nodes into densely-connected communities known as modules, was performed via modularity maximization [17, 18]. Modularity is the fraction of links that are within modules, and the average number of modules is the number of modules in the network configuration that maximizes modularity. All of these graph features were compared for patients across different diagnosis groups.

The full analysis pipeline that we developed for performing graph signal processing of the audiograms for individual patients is illustrated in Figure 2.

RESULTS

There were qualitative differences in the resulting cochlea graphs among the different diagnoses groups. Figure 3 shows an example from each group. All of the graphs had a linear component, where IHCs were connected with their anatomical neighbors along the spiral, as expected. However, we also saw a fair amount of nonlinear connections across the three-dimensional structure of the cochlea. In the NH and conductive HL graphs, these cross-cochlea links appeared more organized into modules around different sections of the spiral. In the sensorineural HL and both types of HL graphs, the cross-cochlea connections appeared to saturate parts of the spiral. In all diagnosis groups, there were many connections between the base and apex. Some of these may be due to a boundary effect from lack of sufficient stimuli that excite the most basal or most apical cells, which needs to be explored further.

When calculating various features to quantify the structures of these graphs in each diagnosis group, the largest differences were seen between the graphs of patients with NH and both types of HL. The average node centrality for patients with NH was 0.48 ± 0.01 ($\pm$ standard error), while it was 0.69 ± 0.02 for patients with both types of HL. The graph density for patients with NH was 0.24 ± 0.01, while it was 0.34 ± 0.01 for patients with both types of HL. The average weighted degree was 1.14 ± 0.02 for NH patients and 0.89 ± 0.02 for patients with both types of HL. The higher average centrality, higher graph density, and lower average weighted degree in graphs of patients with HL suggests that these graphs generally have more links, but that these links have lower weights. In other words, the links are less significant, and these graphs were exhibiting noisier behavior than the NH graphs.

There was not a large difference in modularity for patients with NH (0.64 ± 0.00), conductive HL (0.63 ± 0.00), sensorineural HL (0.63 ± 0.00), or both types of HL (0.62 ± 0.00). Patients with both types of HL had slightly more modules (6.14 ± 0.14) compared to NH (5.72 ± 0.09), however this may be in part due to these graphs having a

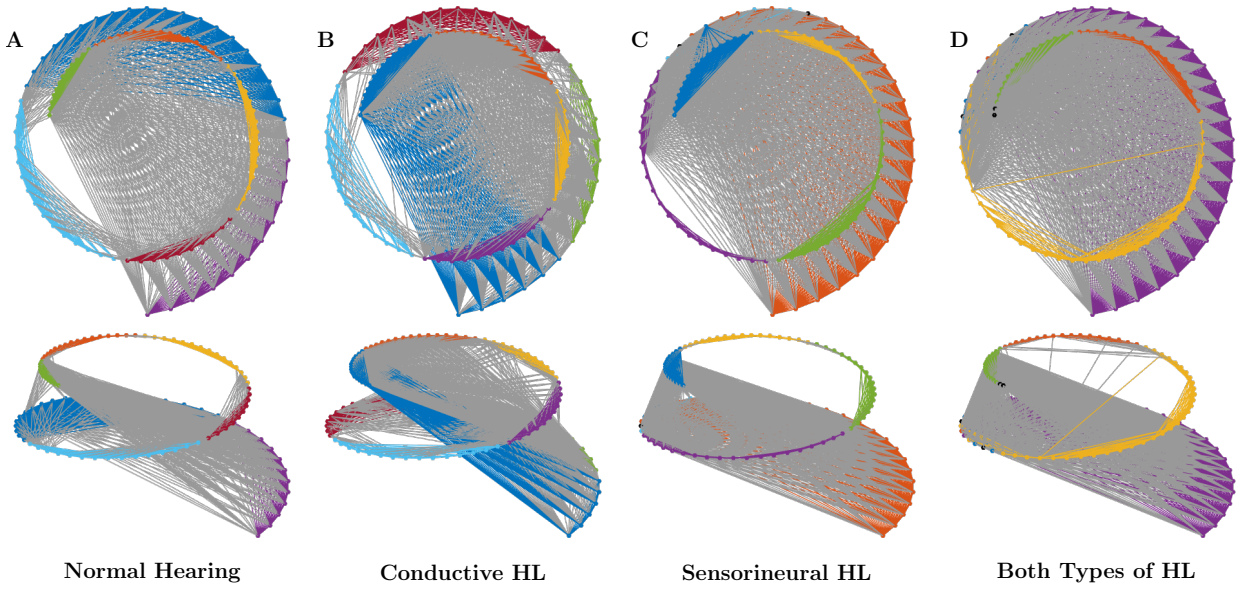


FIGURE 3. Cochlea graph examples for individual patients in each diagnosis group. Patients were chosen for having graphs with most features falling within one standard error of the averages for their respective diagnosis groups (Figure 4). Nodes and within-module links are colored according to their module allegiance after performing community detection. Between-module links are gray. Module colors across patients are arbitrary. The three-dimensional views are the same as in Figure 2.

greater number of very small modules that the algorithm could not sufficiently place into a larger module due to lack of reliable links. There also was a small difference in average clustering, for graphs from patients with NH (0.76 ± 0.00) and those with both types of HL (0.80 ± 0.01).

These results are summarized in Figure 4.

DISCUSSION

Linear tonotopy is essential in models of the cochlea's geometry, e.g., [19], and of the full human auditory periphery, which typically incorporate the phenomenon via a transmission-line architecture, e.g., [20], or parallel filter banks, e.g., [8]. Here, we have reported the first time that the outputs of these existing models of cochlea response, which inherently comprise IHCs exhibiting peak responses to their characteristic frequencies, have been taken a step further to construct a nonlinear graph representation of cochlear mechanics (Figure 1). The graph approach allowed us to examine how the response of individual cells is coordinated at the level of the whole, three-dimensional cochlea for passing complex information to the auditory nerve fibers. This method was a means for developing a new graph data structure, and we were able to successfully distinguish patients with different diagnoses in this graph space to validate that the cochlea graphs have useful information. Future work will explore the impact of graph size and number of graph signals on the learned graph topologies. Additionally, especially with the application to cochlear implants in mind, we will look at the impact of age on the graph features. Age of implantation has been associated with differences in difficulties with pitch and timbre discrimination [21].

Even during NH, the mechanical response of the cochlea to music is complex [22]. Though rhythm perception in cochlear implant users is comparable to NH listeners, melody and timbre perception have been shown to be hindered by limitations of standard cochlear implant music encoding [23]. Pitch information is believed to be processed from spectral and high frequency temporal cues in NH individuals. Pitch discrimination difficulties arise in cochlear implant users from, for example, surgical limitations of electrode placement [24], imprecise spatial stimulation of the cochlea [25], and disruption of the frequency ratios that convey sound consonance [26]. These difficulties further disrupt perception of musical emotion [27] and melodic contour [28]. Timbre is perceived from the temporal envelope of a sound and the presence of harmonics (integer multiples of the fundamental frequency). Poor timbre discrimination in cochlear implant users arises from alterations in the intervals between harmonics of the sound that are encoded

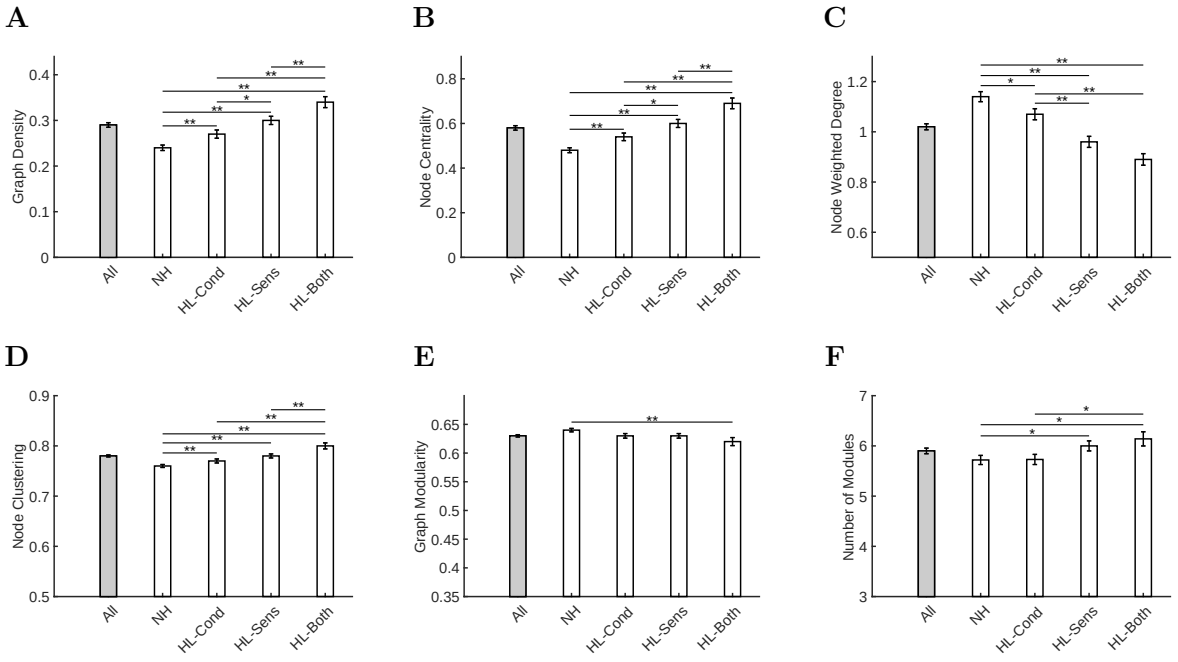


FIGURE 4. Graph features calculated and averaged over all patients ('All') and within in each diagnosis group ('NH' is patients with normal hearing, 'HL-Cond' is patients diagnosed with conductive hearing loss only, 'HL-Sens' is patients diagnosed with sensorineural hearing loss only, 'HL-Both' is patients with both conductive and sensorineural hearing loss). Error bars are standard error of the mean. Student t-tests were performed for all pairs of diagnosis groups. Single asterisk indicates p-value < 0.05, and double asterisk indicates p-value < 0.01.

by the cochlear implant due to the electrodes stimulating a broad frequency range of auditory nerve fibers [23]. For both pitch and timbre, the preservation of frequency ratios and intervals is necessary but difficult given the hardware limitations on the number of and distance between electrodes. By constructing a patient's nonlinear cochlea graph from their personal hearing health data and instances of encoding various harmonies and harmonics, graph deep learning methods [29, 30] may provide new solutions for how music can be processed with "higher fidelity" [4] through these devices.

CONCLUSION

In summary, we explored a new model of cochlear mechanics by means of graph signal processing. This model constructed a data structure that combined clinical audiograms with the biophysics of music processing. The UR Ear 2020b simulation generated IHC voltage responses for various musical stimuli, including a variety of single pitches, single timbres, pitch combinations (chords), timbre combinations, and song excerpts. These voltage responses were converted into graph signals that were then used to learn links between the IHC nodes using GSPBox. Calculating various graph-theoretic features of the resulting graphs allowed us to find quantitative differences for patients with NH, conductive HL, sensorineural HL, and both types of HL. Given that the mechanics of hearing is altered during hearing loss, our patient cochlea graphs have potential for informing efforts to optimize music processing in cochlear implants customized to individual users.

ACKNOWLEDGMENTS

M.E.B. is supported by a training fellowship from the Gulf Coast Consortia, on the NIH National Library of Medicine Training Program in Biomedical Informatics and Data Science (T15LM007093). Musical pieces by the M33 Project

are licensed under CC BY-NC 4.0, and the musical piece by the Advent Chamber Orchestra is licensed under CC BY-SA 3.0 US.

REFERENCES

1. M. A. Rutherford, H. von Gersdorff, and J. D. Goutman, "Encoding sound in the cochlea: from receptor potential to afferent discharge," The Journal of Physiology **599**, 2527–2557 (2021). https://doi.org/10.1113/JP279189
2. G. D. Langner, *The neural code of pitch and harmony* (Cambridge University Press, 2015).
3. D. D. Greenwood, "A cochlear frequency-position function for several species–29 years later," The Journal of the Acoustical Society of America **87**, 2592–2605 (1990). https://doi.org/10.1121/1.399052
4. G. M. Clark, "The multi-channel cochlear implant: Multi-disciplinary development of electrical stimulation of the cochlea and the resulting clinical benefit," Hearing Research **322**, 4–13 (2015). https://doi.org/10.1016/j.heares.2014.08.002
5. C. J. Limb and A. T. Roy, "Technological, biological, and acoustical constraints to music perception in cochlear implant users," Hearing Research **308**, 13–26 (2014). https://doi.org/10.1016/j.heares.2013.04.009
6. A. Ortega, P. Frossard, J. Kovačević, J. M. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," Proceedings of the IEEE **106**, 808–828 (2018). https://doi.org/10.1109/JPROC.2018.2820126
7. G. Mateos, S. Segarra, A. G. Marques, and A. Ribeiro, "Connecting the dots: Identifying network structure via graph signal processing," IEEE Signal Processing Magazine **36**, 16–43 (2019). https://doi.org/10.1109/MSP.2018.2890143
8. M. S. Zilany, I. C. Bruce, and L. H. Carney, "Updated parameters and expanded simulation options for a model of the auditory periphery," The Journal of the Acoustical Society of America **135**, 283–286 (2014). https://doi.org/10.1121/1.4837815
9. L. H. Carney, T. Li, and J. M. McDonough, "Speech coding in the brain: representation of vowel formants by midbrain neurons tuned to sound fluctuations," Eneuro **2** (2015). https://doi.org/10.1523/ENEURO.0004-15.2015
10. I. C. Bruce, Y. Erfani, and M. S. Zilany, "A phenomenological model of the synapse between the inner hair cell and auditory nerve: Implications of limited neurotransmitter release sites," Hearing Research **360**, 40–54 (2018). https://doi.org/10.1121/1.3238250
11. N. Perraudin, J. Paratte, D. I. Shuman, V. Kalofolias, P. Vandergheynst, and D. K. Hammond, "GSPBOX: A toolbox for signal processing on graphs," ArXiv e-prints (2014), arXiv:1408.5781 [cs.IT]. https://doi.org/10.48550/arXiv.1408.5781
12. M. Pietsch, L. Aguirre Dávila, P. Erfurt, E. Avci, T. Lenarz, and A. Kral, "Spiral form of the human cochlea results from spatial constraints," Scientific reports **7**, 1–11 (2017). https://doi.org/10.1038/s41598-017-07795-4
13. J. A. Germiller, E. B. Crenshaw, I. Krantz, J. Peterson, M. Reinders, P. White, and M. Italia, "AudGenDB: A Public, Internet-Based, Audiologic/Otologic/Genetic Database for Pediatric Hearing Research," Otolaryngology–Head and Neck Surgery **145**, P235–P236 (2011). https://doi.org/10.1177/0194599811415823a329
14. L. Fritts, "Musical instrument samples," University of Iowa (1997). https://theremin.music.uiowa.edu/MIS.html
15. V. Kalofolias, "How to learn a graph from smooth signals," in *Artificial Intelligence and Statistics* (PMLR, 2016) pp. 920–929. https://proceedings.mlr.press/v51/kalofolias16.html
16. A. Hagberg, D. Schult, and P. Swart, "NetworkX," GitHub repository (2022). https://github.com/networkx
17. M. E. Newman, "Modularity and community structure in networks," Proceedings of the National Academy of Sciences USA **103**, 8577–8582 (2006). https://doi.org/10.1073/pnas.0601602103
18. Z. Zuo, "Python modularity maximization," GitHub repository (2018). https://zhiyzuo.github.io/python-modularity-maximization/
19. A. Altoè and C. A. Shera, "The cochlear ear horn: geometric origin of tonotopic variations in auditory signal processing," Scientific reports **10**, 1–10 (2020). https://doi.org/10.1038/s41598-020-77042-w
20. S. Verhulst, A. Altoe, and V. Vasilkov, "Computational modeling of the human auditory periphery: Auditory-nerve responses, evoked potentials and hearing loss," Hearing research **360**, 55–75 (2018). https://doi.org/10.1016/j.heares.2017.12.018
21. Y. Zaltz, R. L. Goldsworthy, L. Kishon-Rabin, and L. S. Eisenberg, "Voice discrimination by adults with cochlear implants: The benefits of early implantation for vocal-tract length perception," Journal of the Association for Research in Otolaryngology **19**, 193–209 (2018). https://doi.org/10.1007/s10162-017-0653-5
22. M. R. Jones, R. R. Fay, and A. N. Popper, eds., *Music Perception* (Springer, New York, NY, 2010).
23. F. Sorrentino, F. Gheller, N. Favaretto, L. Franz, E. Stocco, D. Brotto, and R. Bovo, "Music perception in adult patients with cochlear implant," Hearing, Balance and Communication **18**, 3–7 (2020). https://doi.org/10.1080/21695717.2020.1719787
24. F.-G. Zeng, S. Rebscher, W. Harrison, X. Sun, and H. Feng, "Cochlear implants: system design, integration, and evaluation," IEEE Reviews in Biomedical Engineering **1**, 115–142 (2008). https://doi.org/10.1109/RBME.2008.2008250
25. K. L. Plant, H. J. McDermott, R. J. Van Hoesel, P. W. Dawson, and R. S. Cowan, "Factors influencing electrical place pitch perception in bimodal listeners," The Journal of the Acoustical Society of America **136**, 1199–1211 (2014). https://doi.org/10.1121/1.4892790
26. N. T. Jiam, M. T. Caldwell, and C. J. Limb, "What does music sound like for a cochlear implant user?" Otology & Neurotology **38**, e240–e247 (2017). https://doi.org/10.1097/MAO.0000000000001448
27. M. Caldwell, S. K. Rankin, P. Jiradejvong, C. Carver, and C. J. Limb, "Cochlear implant users rely on tempo rather than on pitch information during perception of musical emotion," Cochlear Implants International **16**, S114–S120 (2015). https://doi.org/10.1179/1467010015Z.000000000265
28. J. J. Galvin III, Q.-J. Fu, and R. V. Shannon, "Melodic contour identification and music perception by cochlear implant users," Annals of the New York Academy of Sciences **1169**, 518 (2009). https://doi.org/10.1111/j.1749-6632.2009.04551.x
29. M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst, "Geometric deep learning: going beyond euclidean data," IEEE Signal Processing Magazine **34**, 18–42 (2017). https://doi.org/10.1109/MSP.2017.2693418
30. T. M. Roddenberry, N. Glaze, and S. Segarra, "Principled simplicial neural networks for trajectory prediction," in *International Conference on Machine Learning* (PMLR, 2021) pp. 9020–9029. http://proceedings.mlr.press/v139/roddenberry21a.html

Simulation of Conductive Hearing Loss and Its Impact on Distortion-product Otoacoustic Emissions Using a Hydrodynamic Cochlea Model

Ernst Dalhoff,[1, a)] Anthony W. Gummer,[1, b)] and Dennis Zelle[1, 2, c)]

[1)] *Section of Physiological Acoustics and Communication, Dept. of Otolaryngology, Eberhard-Karls-University Tübingen, Elfriede-Aulhorn-Str. 5, 72076 Tübingen, Germany*

[2)] *Earlab GmbH, Etterschlager Str. 4, 82237 Wörthsee, Germany*

[a)] *Corresponding author: ernst.dalhoff@uni-tuebingen.de*

[b)] *Electronic mail: anthony.gummer@uni-tuebingen.de*

[c)] *Electronic mail: d.zelle@gmail.com*

Abstract.
Distortion-product otoacoustic emissions (DPOAEs) are sound signals in the ear canal arising from nonlinear amplification of hydrodynamic oscillations in the cochlea. The analysis of the growth behavior of DPOAE pressure as function of stimulus level offers a quantitative assessment of the mechanical state of the cochlea. However, middle-ear dysfunction directly influences DPOAEs and their diagnostic accuracy. The present work simulates the impact of middle-ear dysfunction on DPOAEs using a hydrodynamic model of the human cochlea coupled to a middle-ear model for five frequencies f_2 = 1 to 4 kHz. DPOAEs in the ear canal are simulated by simultaneously solving the equations of motion representing the dynamics of the middle-ear and cochlea model in the time domain. Increasing the damping and stiffness of the annular ligament in the middle-ear model introduces conductive hearing loss (CHL) of various degrees. The changes in estimated distortion-product thresholds (EDPTs) and slopes of input-output (I/O) functions derived from the simulated DPOAE growth behavior are compared to changes in the middle-ear transfer functions (METFs). Pooled over all frequencies and degrees of CHL, the relative change of the METF at f_2 exhibits a linear dependency on the relative change of EDPT level while the relative change of the METF at the distortion-product frequency f_{DP} is linearly related to the relative change of the slope. The statistically significant linear relationships of the change in EDPT level and slope of the I/O functions with changes in the METF indicate that conductive hearing loss can be objectively quantified from the DPOAE growth behavior.

INTRODUCTION

Distortion-product otoacoustic emissions (DPOAEs) are sound signals in the ear canal that emerge from nonlinear amplification of hydrodynamic oscillations in the cochlea in response to two tones of frequencies f_1 and f_2 ($f_2/f_1 = 1.2$). DPOAEs at the cubic distortion-product frequency $f_{\mathrm{DP}} = 2f_1 - f_2$ offer a quantitative assessment of the mechanical state of the cochlea [1], and, according to a widely accepted model [2], mainly consist of two components, a nonlinear-distortion and a coherent-reflection component. Semi-logarithmic DPOAE input-output (I/O) functions that plot the DPOAE amplitude as function of stimulus level L_2 enable the computation of estimated distortion-product thresholds (EDPTs), which have been shown to predict behavioral thresholds with high accuracy [3] when accounting for wave interference between the nonlinear-distortion and coherent-reflection components [4].

However, middle-ear dysfunction directly influences DPOAEs and their diagnostic accuracy since conductive hearing loss (CHL) reduces both the forward propagation of the acoustic stimuli and the backward transmission of the DPOAEs. For example in neonates, middle-ear dysfunction due to amniotic fluid hampers the identification of cochlear disorders using otoacoustic emissions [5]. CHL has been shown to yield an increase of EDPT accompanied with a reduced slope in the I/O function [5]. Additionally, changes in middle-ear transfer function (METF) result in an altered optimal relationship between the stimulus levels L_1 and L_2 that facilitate DPOAE generation in the cochlea [6], further decreasing the amplitude of the DPOAE response.

Recently, the computation of I/O functions was refined using a least-squares fit approach to reproduce the DPOAE growth behavior from so-called DPOAE level maps, which sample the DPOAE growth behavior at multiple stimulus-level combinations in the L_1, L_2 space [7]. The derivation of ridge-based I/O functions from these model level maps provide EDPTs with high precision but without the need of predefined optimal stimulus-level pairs, therefore, presenting a promising method to acquire DPOAEs in the presence of CHL. The present work simulates the growth behavior of the nonlinear-distortion component, i.e. without an interfering reflection component, in the time domain using a hydrodynamic model of the human cochlea coupled to a middle-ear model [8] for various degrees of CHL.

METHODS

Hydrodynamic Cochlea Model

To simulate DPOAE time responses, the cochlea model reproduces the transversal motion of the basilar membrane (BM) and the radial motion of the stereocilia of the outer hair cells (OHCs) using mechanical, longitudinally coupled oscillators [9, 10]. For the normalized longitudinal position x, the transversal BM displacement q_y is given by

$$F(x,t) = m(x)\ddot{q}_y(x,t) + [d(x) + \partial_x s(x)\partial_x]\,\dot{q}_y(x,t) + k(x)q_y(x,t) \tag{1}$$

with the mass density $m(x)$ of the Corti organ, the viscous damping $d(x)$ of the cochlear fluids, and the BM stiffness $k(x)$. The hydrodynamic force density $F_{\mathrm{H}}(x,t)$ and the viscous shearing resistance $s(x)$ provide longitudinal coupling between the oscillators [9]. The effective force density on the BM is

$$F(x,t) = F_{\mathrm{H}}(x,t) + F_{\mathrm{CA}}(x,t) \tag{2}$$

consisting of $F_{\mathrm{H}}(x,t)$ and $F_{\mathrm{CA}}(x,t)$, which reproduces the force density due to active cochlear amplification. The nonlinear cochlear-amplifier force is a function of the stereocilia displacement

$$-g_{\mathrm{TM}}(x)\ddot{q}_y(x,t) = \ddot{q}_z(x,t) + \left[\gamma_{\mathrm{TM}}(x) + \frac{\partial_x s_{\mathrm{TM}}(x)\partial_x}{m_{\mathrm{TM}}(x)}\right]\dot{q}_z(x,t) + \omega_{\mathrm{TM}}^2(x)q_z(x,t) \tag{3}$$

where $\gamma_{\mathrm{TM}}(x)$, $\omega_{\mathrm{TM}}(x)$, and $s_{\mathrm{TM}}(x)$ represent the damping, the resonance frequency, and the shearing resistance of the tectorial membrane (TM), respectively [10]. Similarly to the model described in [9, 10], the TM inertia elicits stereocilia motion due to a rotation of the Corti organ. Hence, BM acceleration excites Eq. (3) via the geometrically derived coupling coefficient [11]

$$g_{\mathrm{TM}}(x) = \frac{m(x)}{g(x)m_{\mathrm{TM}}(x)} \tag{4}$$

with $b(x)$ and $h(x)$ as the effective BM width and the height of the Corti organ, respectively. In order to achieve locally confined cochlear amplification, the resonance frequency $\omega_{\mathrm{TM}}(x)$ is set to be a half of an octave below that of the BM at the tonotopic place x [12]. The viscous damping due to the cochlear fluid is compensated by cochlear amplification if $F_{\mathrm{CA}}(x,t) \geq d(x)\dot{q}_y(x,t)$, yielding the cochlear force density

$$F_{\mathrm{CA}}(x,t) = -\lambda(x)\frac{\gamma_{\mathrm{TM}}(x)d(x)}{\beta a(x)g_{\mathrm{TM}}(x)}S[a(x)q_z(x,t)] \tag{5}$$

derived from the linearized small-signal behavior for low stimulus levels [10]. In Eq. (5), $S[\cdot]$, β, and $a(x)$ replicate the nonlinearity of the mechanoelectrical transduction of the OHC, and $\lambda(x)$ represents an empirically chosen scaling factor to counteract shearing resistance yielding a maximum cochlear amplification of 40 dB at 4 kHz. Cochlear amplification is determined relative to the maximum of the traveling wave in the passive model, i.e. in a model without cochlear amplification.

For incompressible and laminar cochlear fluids, the hydrodynamic force density is defined as

$$F_{\mathrm{H}}(x,t) = G_{\mathrm{S}}(x)\ddot{q}_u(t) - \int_{\tilde{x}=0}^{\tilde{x}=1} G(x,\tilde{x})\ddot{q}_y(\tilde{x},t)d\tilde{x} \tag{6}$$

with the hydrodynamic coupling terms $G_{\mathrm{S}}(x)$ and $G(x,\tilde{x})$ relating the pressure changes due to accelerations of the stapes $\ddot{q}_u$ at $x = 0$ and of the BM at $\tilde{x}$ to the effective force at x [9].

Middle-Ear Model

The equations of motion of the three-mass oscillator depicted in Fig. 1 replicates the middle-ear transfer characteristics. A least-squares fit in the frequency domain with the METF

$$H_{\mathrm{ME}}(j\omega) = \frac{q_u(j\omega)}{P_{\mathrm{Ty}}(j\omega)} \tag{7}$$

of a realistic finite-element model of the middle ear [13] yields the model parameters d_i, k_i, and m_i. In the least-squares fit, A_{Ty}, A_{s}, m_3, k_4, d_4, and d_{C} are predefined, referring to the the tympanic-membrane area, the stapes-footplate area, the stapes mass, the stiffness and damping of the annular ligament, and the cochlear damping, respectively. The displacement u_1 of the tympanic membrane is converted into ear-canal sound pressure according to [14] and provides the simulation of DPOAEs that are transferred via the middle ear from the hydrodynamic force $F_{\mathrm{H,0}}$ at $x = 0$.

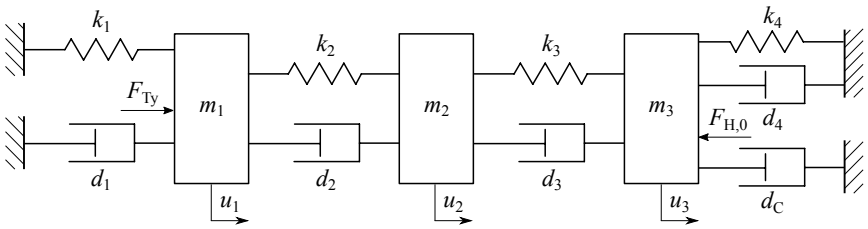


FIGURE 1. Simplified middle-ear model. F_{Ty}: force at the tympanic membrane; $u_3 = q_u$: stapes displacement; $F_{\mathrm{H,0}}$: hydrodynamic force at $x = 0$.

Numerical Simulations

Short-pulse DPOAEs [3, 7] in the ear canal are simulated by simultaneously solving the equations of motion in the time domain using the Dormand-Prince method in MATLAB (Ver. 9.9, MathWorks, Natick, MA). To enable cancellation of the stimulus tones and analysis of the DPOAE time signal, pulsed stimuli for four different primary-tone phase combinations are simulated [15]. The solution is computed for equidistant discretization of 300 elements along the BM and frequency-dependent sampling frequencies ranging from 26 kHz at $f_2 = 1$ kHz and 104 kHz at $f_2 = 4$ kHz. Increasing the damping and stiffness of the annular ligament in the middle-ear model by setting $d_4' = \alpha d_4$ and $k_4' = \alpha k_4$ with $\alpha = 1$ to 10 introduces conductive hearing loss (CHL) of various degrees. Numerical simulations of the DPOAE growth behavior in the L_1, L_2 space sampled at 240 stimulus-levels pairs for five frequencies $f_2 = 1$, 1.5, 2, and 4 kHz yield DPOAE level maps [7]. For each degree of CHL, model level maps were computed from the simulated data to derive ridge-based I/O functions. The changes in EDPTs and slopes of these I/O functions are compared to METF changes.

RESULTS

Model Properties

The model mimics experimentally visible traveling waves and shows a good agreement with the tonotopic frequency map of the human cochlea for frequencies between 0.125 and 10 kHz and realistic amplification between 0.5 and 6 kHz. Figure 2A shows a traveling wave for the active model elicited with a pure tone of $f_2 = 1.5$ kHz and $L = 35$ dB SPL (black solid line). With decreasing cochlear amplification, e.g. due to cochlear disorders or increasing stimulus level, the traveling wave broadens and shifts towards the cochlear base (Fig. 2A, gray dashed line). Simultaneous stimulation with two tones of frequencies f_1 and f_2 elicits distortion products in the overlapping region of the two corresponding traveling waves that are visible in the BM velocity (Fig. 2B) close to and at the f_2-tonotopic place, and as DPOAE in the ear canal (Fig. 2C).

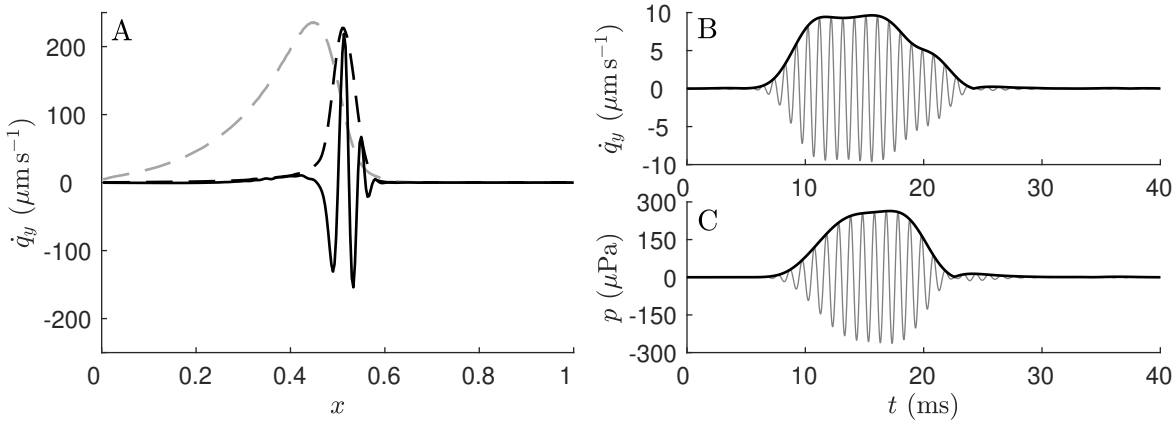


FIGURE 2. A: Simulated traveling wave (black solid line) with envelope (black dashed line) for the active model at $L = 35$ dB SPL and $f = 1.5$ kHz. Gray dashed line: traveling-wave envelope for the passive model ($F_{CA} = 0$) at $L = 63$ dB SPL and $f = 1.5$ kHz. **B**: BM velocity and its envelope at tonotopic place of f_2 at $x = 0.52$ for a short-pulse distortion product after primary-tone phase variation [15] with $f_2 = 1.5$ kHz, $L_1 = 45$ dB SPL, $L_2 = 35$ dB SPL. The f_2 pulse is presented between $t = 5$ and 17 ms, while the f_1 pulse starts at $t = 0$ ms with a duration of 40 ms. **C**: Simulated DPOAE response in the ear canal for the parameters in B.

DPOAE Level Maps

Figure 3A depicts the DPOAE level map for $f_2 = 2$ kHz representing the growth behavior of the nonlinear-distortion component for normal middle-ear transfer characteristics. In general, for each L_2 level there exists a L_1 level producing maximum DPOAE amplitude (red circles), representing optimally evoked DPOAE amplitudes. While the DPOAE amplitude at optimal stimulus-level pairs increases linearly with increasing L_2 for low to moderate levels, the DPOAE amplitude shows considerable saturation at high stimulus levels with $L_2 \geq 60$ dB SPL. Figure 3B shows the model level map derived by fitting a mathematical model [7] to the simulated data shown in A.

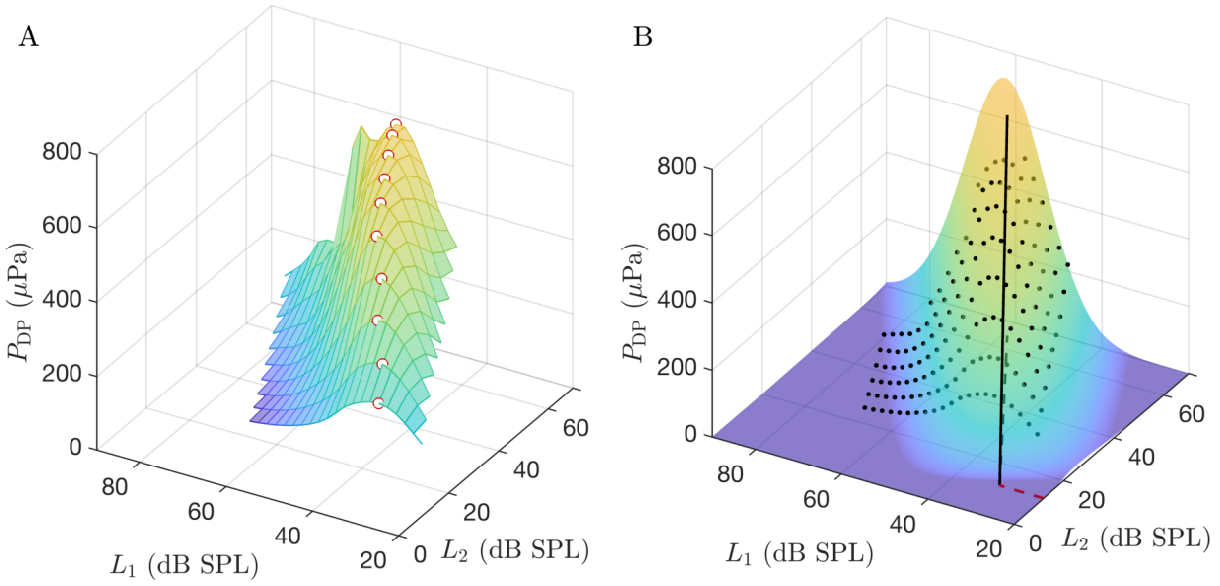


FIGURE 3. A: Simulated DPOAE level map of the nonlinear-distortion component at $f_2 = 2$ kHz with regular METF. **B**: Model level map derived from the simulated data (black dots) in A using a least-squares fit [7]. Parameters of the ridge-based I/O function (black solid line): EDPT (red dashed line) $L_{EDPT} = 11.88 \pm 0.48$ dB SPL, slope $s = 10.88 \pm 0.15$ μPa/dB SPL.

The mathematical model describes the DPOAE growth behavior in a rotated coordinate system as a sequence of parabolas located perpendicular and symmetrical to the optimal path. The fit yields the optimal stimulus-path parameters a and b, the scaling c of the parabolas, the EDPT level L_{EDPT} (red dashed line), and the slope s of the ridge-based I/O function (black solid line).

Dependence of DPOAE Growth Behavior on Conductive Hearing Loss

The increased damping and stiffness of the annular ligament yields a considerable reduction of sound propagation for frequencies below the first resonance frequency of the three-mass oscillator (Fig. 4A). With increasing coefficient α the resonances in the METF become more pronounced and shift slightly towards higher frequencies. The reduced sound conduction offset the simulated DPOAE level maps towards higher stimulus levels and considerably reduce the DPOAE growth, which becomes apparent in a decrease of the slope of the DPOAE I/O functions (Fig. 4B).

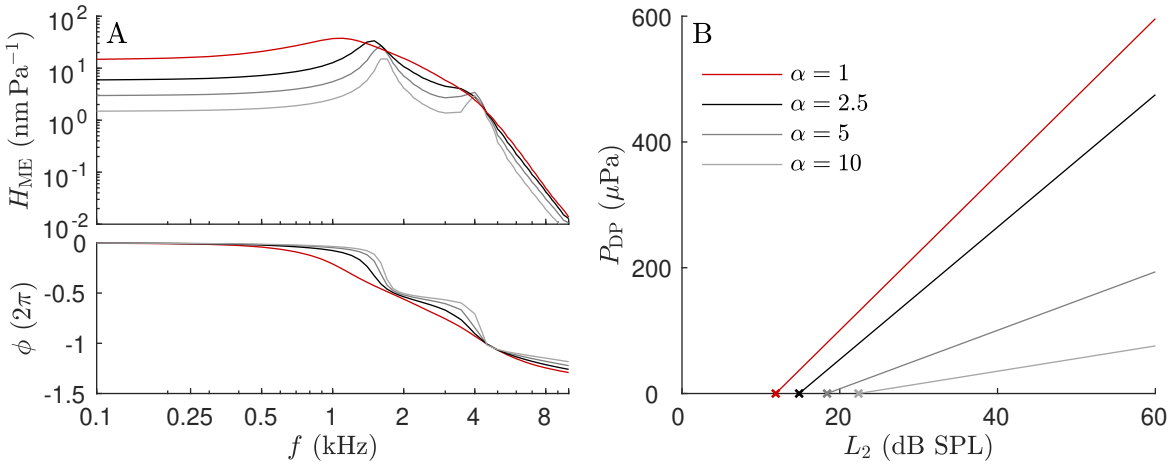


FIGURE 4. METFs and ridge-based DPOAE I/O functions for increasing CHL. **A**: Exemplary METFs H_{ME} for increasing scaling factor α of the damping d_4 and stiffness k_4 of the annular ligament. $\alpha = 1$ (red line) corresponds to regular METF. **B**: Ridge-based I/O functions extracted from model level maps for $f_2 = 2$ kHz for the METFs depicted in A.

Pooled over all frequencies and degrees of CHL, the relative change of the METF at the stimulus-tone frequency f_2 exhibits a significant linear dependency on the relative change of EDPT level, ΔL_{EDPT}, with $r^2 = 0.972$ ($p < 0.001$, $N = 70$) and an estimation error computed as the standard deviation of the residuals $\sigma = 1.14$ dB. The linear relationship can be quantified with the regression line $\Delta H_{\mathrm{ME}}(f_2) = [0.94 \pm 0.02]\Delta L_{\mathrm{EDPT}} - [0.22 \pm 0.17]$ dB (Fig. 5A).

Additionally, the relative change of the METF at f_{DP} can be linearly related to the relative change of the slope of the ridge-based I/O function, Δs, with $r^2 = 0.989$ ($p < 0.001$, $N = 70$) and $\sigma = 0.70$ dB. Again, a regression line describes the relationship sufficiently with $\Delta H_{\mathrm{ME}}(f_{\mathrm{DP}}) = [1.07 \pm 0.01]\Delta s + [0.44 \pm 0.14]$ dB (Fig. 5B).

DISCUSSION

Simulation of DPOAE responses in an active, nonlinear time-domain model of the cochlea as shown here demonstrate that measurement of DPOAE level-map characteristics offers the possibility to separate conductive middle-ear loss from sensorineural hearing loss, i.e. loss of the cochlear amplification, with high accuracy. The underlying principle has been already predicted and also tested experimentally.

In case of CHL, the slope of DPOAE I/O functions decreases with reverse-transmission loss of the middle ear, while the EDPT, i.e. the extrapolated DPOAE threshold, shifts towards higher levels because the forward middle-ear loss must be compensated with higher stimulus levels in order to produce an intra-cochlear stimulus that is comparable to the situation without CHL [5]. Concurrently, forward-transmission loss that is approximately equal for both stimuli is expected to yield a shift of the optimal stimulus-level path to higher L_1 values and a decrease in the slope a of the

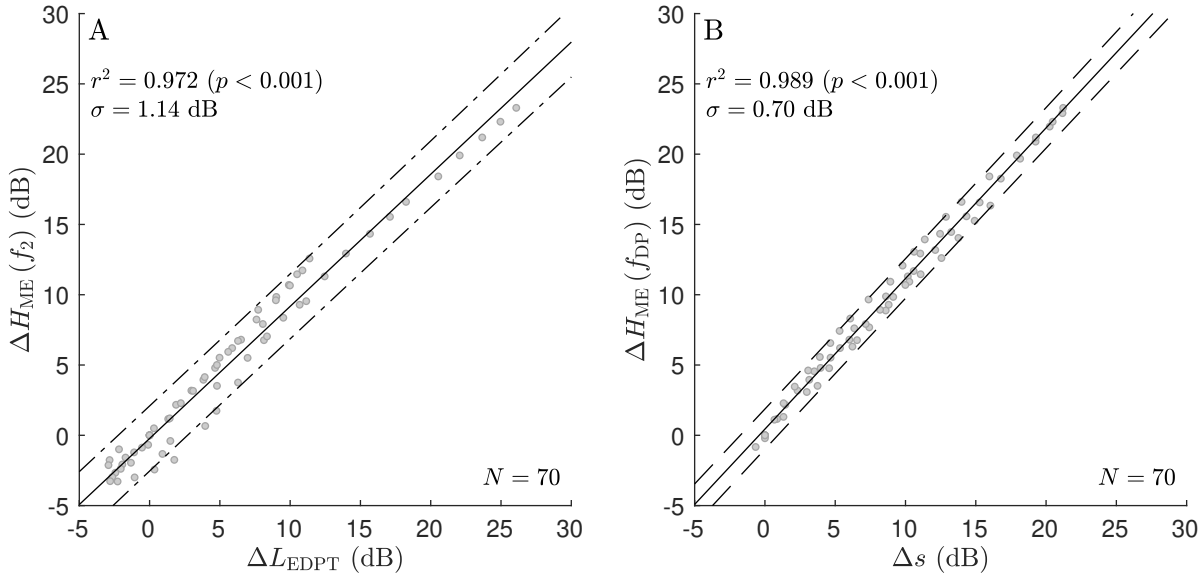


FIGURE 5. Prediction of middle-ear transmission change using objective diagnostic parameters derived from ridge-based I/O functions. **A**: Relative change in METF at f_2, ΔH_{ME}, plotted against the relative change in EDPT level, ΔL_{EDPT}. **B**: Relative change in METF at f_{DP}, ΔH_{ME}, plotted against the relative change in slope of the I/O function, Δs.

optimal stimulus-level path in the L_1, L_2-plane. In newborn screening and follow-up visits, these changes appeared, most probably because of vanishing amniotic fluid between both visits [6].

The middle-ear model used in this study mimics the overall transfer characteristics of a realistic finite-element model [13] and, therefore, its constituent parts cannot readily be related to anatomical elements of a real middle ear. However, the simplified three-mass oscillator considerably reduces computational effort and enables the simulation of DPOAEs directly in the time domain for various degrees of CHL.

The correlations shown in Fig. 5A and B and the small standard deviation of the residue, encompassing testing frequencies of 1, 1.5, 2, and 4 kHz with conductive losses of up to 20 dB, show that the principle may be expected to have high methodological accuracy. This is remarkable as the simulated conductive loss in the middle-ear model shows clear resonances, which are expected to affect both stimulus tones at some of the testing frequencies differently. Clearly, to exploit this principle optimally, several aspects should be considered. First, the EDPT should be derived from a DPOAE level map [7] and rely on a technique separating the nonlinear-distortion component from the reflection component, as done in this simulation. Second, a shift of the optimum stimulus-level path towards higher L_1-levels and a decrease of the slope s as compared to normative data should be evaluated together in a conductive-loss separation algorithm.

By subtraction of the estimated CHL from the estimated hearing loss derived from the EDPT, the gain of the cochlear amplifier can be measured independently from middle-ear disorders.

CONCLUSION

The statistically significant linear relationship of the change in EDPT level and slope of the I/O functions with changes in the METF with very low standard deviation of the residue indicates that conductive hearing loss can be objectively quantified using I/O functions derived from DPOAE level maps. This technique is especially promising for newborn screening.

ACKNOWLEDGMENTS

The authors would like to thank S. Oßmann and M. Bornitz for providing the METF of their finite-element middle-ear model. This work was supported by the German Research Council, DFG DA 487/4-1.

REFERENCES

1. P. Avan, B. Büki, and C. Petit, "Auditory distortions: Origins and functions," Physiological Reviews **93**, 1563–1619 (2013).
2. C. A. Shera and J. J. Guinan Jr., "Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs," The Journal of the Acoustical Society of America **105**, 782–798 (1999).
3. D. Zelle, L. Lorenz, J. P. Thiericke, A. W. Gummer, and E. Dalhoff, "Input-output functions of the nonlinear-distortion component of distortion-product otoacoustic emissions in normal and hearing-impaired human ears," The Journal of the Acoustical Society of America **141**, 3203–3219 (2017).
4. E. Dalhoff, D. Turcanu, A. Vetešník, and A. W. Gummer, "Two-source interference as the major reason for auditory-threshold estimation error based on DPOAE input-output functions in normal-hearing subjects," Hearing Research **296**, 67–82 (2013).
5. T. Janssen, D. D. Gehr, A. Klein, and J. Müller, "Distortion product otoacoustic emissions for hearing threshold estimation and differentiation between middle-ear and cochlear disorders in neonates," The Journal of the Acoustical Society of America **117**, 2969 (2005).
6. S. C. Marcrum, P. Kummer, and T. Steffens, "Estimation of Minor Conductive Hearing Loss in Humans Using Distortion Product Otoacoustic Emissions," Ear and Hearing **38**, 391–398 (2017).
7. D. Zelle, K. Bader, L. Dierkes, A. W. Gummer, and E. Dalhoff, "Derivation of input-output functions from distortion-product otoacoustic emission level maps," The Journal of the Acoustical Society of America **147**, 3169–3187 (2020).
8. D. Zelle, *Innenohrdiagnostik mittels Zeitbereichsanalyse distorsionsprodukt-otoakustischer Emissionen* (TUDpress, Dresden, 2020) p. 272.
9. F. Mammano and R. Nobili, "Biophysics of the cochlea: Linear approximation," The Journal of the Acoustical Society of America **93**, 3320–3332 (1993).
10. R. Nobili, A. Vetešník, L. Turicchia, and F. Mammano, "Otoacoustic emissions from residual oscillations of the cochlear basilar membrane in a human ear model," Journal of the Association for Research in Otolaryngology **4**, 478–494 (2003).
11. J. B. Allen, "Cochlear micromechanics – A physical model of transduction," The Journal of the Acoustical Society of America **68**, 1660–1670 (1980).
12. A. W. Gummer, W. Hemmert, and H.-P. Zenner, "Resonant tectorial membrane motion in the inner ear: its crucial role in frequency tuning," Proceedings of the National Academy of Sciences **93**, 8727–8732 (1996).
13. M. Bornitz, H.-J. Hardtke, and T. Zahnert, "Evaluation of implantable actuators by means of a middle ear simulation model," Hearing Research **263**, 145–151 (2010).
14. C. L. Talmadge, A. Tubis, G. R. Long, and P. Piskorski, "Modeling otoacoustic emission and hearing threshold fine structures," The Journal of the Acoustical Society of America **104**, 1517–1543 (1998).
15. M. L. Whitehead, B. B. Stagner, G. K. Martin, and B. L. Lonsbury-Martin, "Visualization of the onset of distortion-product otoacoustic emissions, and measurement of their latency," The Journal of the Acoustical Society of America **100**, 1663–1679 (1996).

Bone Conduction Stimulation of the Temporal Bone with the Inner Ear: a Finite Element Study

Paweł Borkowski[1, a)], Kazimierz Niemczyk[2] and Magdalena Lachowska[2]

[1]*Institute of Aeronautics and Applied Mechanics, Warsaw University of Technology, 00-665 Warsaw, Poland*
[2] *Medical University of Warsaw, Department of Otolaryngology, 02-097 Warsaw, Poland*

[a)] Corresponding author: pbork@meil.pw.edu.pl

Abstract. A numerical model of the human temporal bone with a precisely reconstructed cochlea was used to investigate bone conduction stimulation applied on the otic capsule. A passive cochlear model was used. Validation of the cochlea was based on the experimental data. The round window membrane and cortical bone were assumed as viscoelastic, with the storage and loss moduli defined as functions of frequency. Other solid tissues were modelled as linear isotropic, and the perilymph fluid was assumed as viscous and compressible. The model was subjected to harmonic analysis in the frequency range from 0.2 to 10 kHz. The bone conduction stimulation was applied as a harmonic force to the mass attached to the otic capsule surface. The boundary of the temporal bone was fixed on the sutures between the adjacent skull bones. The aim of the study was to examine: 1) the type of boundary condition applied on the surface of the petrous part of the temporal bone adjacent to the petro-occipital synchondrosis, 2) the value of the surface density used in the model to include the mass of soft tissues in the vicinity of the temporal bone, assumed as uniformly distributed on its outer surface (inside and outside the skull). Two conditions related to the mobility of the stapes were considered. To assess effects of the above-mentioned factors, amplitude and phase of vibrations obtained from numerical simulation were used, including the volume displacement on the round window membrane, displacements on the basilar membrane and at the spiral lamina along the cochlea, and vibration of the promontory bone at the edge of the round window. The values of forces needed for bone conduction stimulation corresponded to those used in previous experimental studies on cadavers. The results from numerical analysis were compared with the experimental data.

INTRODUCTION

Bone conduction implants placed in the skull surface have been known for many years and used in clinical practice, but new sites of stimulation, such as the otic capsule, are also being investigated [1]. The closer to the cochlea, the less energy is needed for stimulation [2]. Research on bone conduction (BC) can be conducted using experimental tests [3, 4] and the models using electro-mechanical analogy [5] or finite element (FE) method [6, 7].

In the case of BC stimulation of the FE model of the temporal bone, some parameters seem to be important, such as the associated mass of soft tissues and cerebrospinal fluid vibrating with the bone, the effect of inertia of the middle ear ossicles [5] and the type of support applied at the bone boundary. These parameters can affect vibration amplitudes on the flexible parts of the cochlea, such as the basilar membrane (BM) and the round window membrane (RW) but can also affect vibrations of the cochlear promontory bone in the proximity of the RW. In this study, the type of support at petro-occipital synchondrosis (POS), mobility of the stapes, and soft tissues mass were investigated to help validate the FE model based on the experimental data.

METHODS

The subject of the analysis was a three-dimensional model of the temporal bone, previously published in [7]. The FE model (Fig. 1) had well-reconstructed structures of cortical and trabecular parts and included the cochlea, semicircular canals, annular ligament, RW membrane, BM, spiral lamina (SL), spiral ligament, perilymph fluid, helicotrema, and a titanium cylinder representing the bone conduction implant, which was attached to the otic

capsule bony surface above the lateral semicircular canal. Other parts of the inner ear such as the balance receptors, Reissner's membrane, tectorial membrane, organ of Corti, and the cochlear duct were not included. Hence, the passive linear behavior of BM was assumed. The middle and outer ear structures, as the tympanic membrane, malleus, incus, and the middle-ear ligaments and muscles were not included, except the stapes. Consequently, only two of the five possible BC sound pathways [5] were considered: the inertia of the perilymph fluid and compression of the cochlear walls. Sound radiation in the ear canal and in middle-ear cavity and sound transmission in the cerebrospinal fluid were not considered. The effect of inertia of the middle-ear ossicles was partially considered in the model with a movable stapes. In contrast, the otosclerotic condition on the annular ligament was assumed for the case, in which the ligament was replaced with a cortical bone (Fig. 2b).

The RW membrane and cortical bone were assumed as viscoelastic, with the storage and loss moduli defined as functions of frequency. Other solid tissues were modelled as linear isotropic. Some properties of the BM and SL were defined as functions of the spatial coordinates and frequency. The cochlear fluid was assumed as viscous and compressible [7]. A dynamic viscosity of the fluid was used to the vibration damping. The cochlear part of the model was validated based on experimental data.

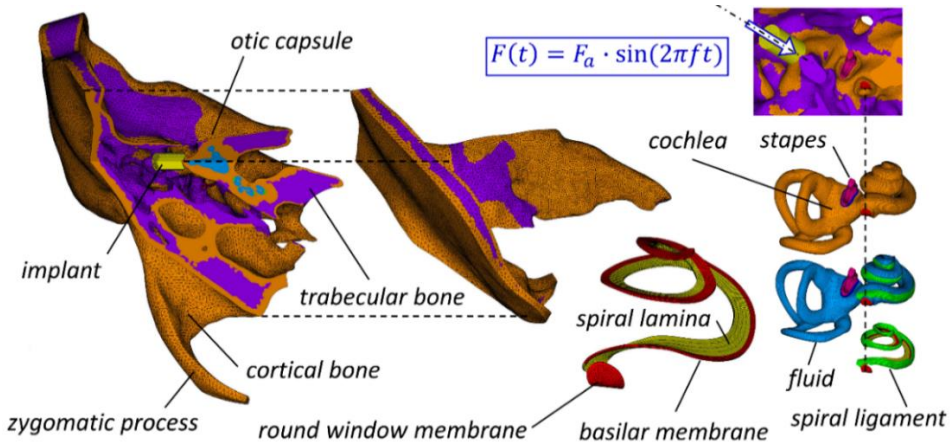


FIGURE 1. FE model of the temporal bone with the bone conduction (BC) implant located at the otic capsule. An arrow indicated the direction of bone conduction stimulation and was coincident with the axis of the cylinder.

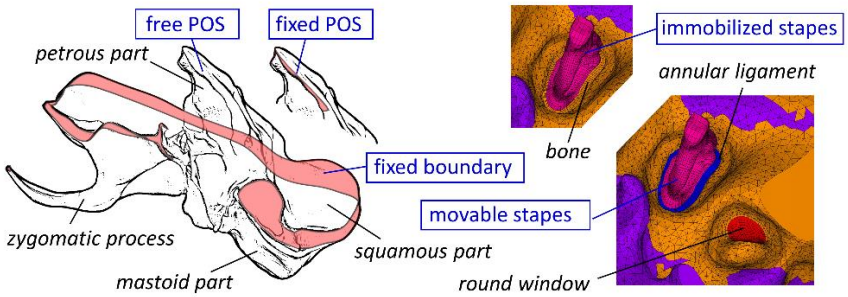


FIGURE 2. The four cases considered in harmonic analysis: (a) fixation of the bone surface (common) and the two types of boundary conditions at the petro-occipital synchondrosis (POS): 'free'- with no ossification, 'fixed'- with full ossification, (b) the two conditions of the stapes mobility: a movable and an immobilized.

Boundary conditions in harmonic analysis

Harmonic analyses for 17 frequencies, based on the one-third octave band (0.2, 0.3, 0.4, 0.5, 0.63, 0.8, 1.0, 1.25, 1.6, 2.0, 2.5, 3.15, 4.0, 5.0, 6.3, 8 and 10 kHz) were performed using the ANSYS program (v17.2, ANSYS Inc., Canonsburg, Pennsylvania, US). The temporal bone was supported by fixing all translational degrees of freedom on

its boundary (Fig. 2a). Additionally, two conditions related to the petro-occipital synchondrosis (POS) state were considered. The fixed POS imitated full ossification, and the free POS - no ossification. The temporal bone surface (inside area of 4960 mm^2, outside area of 5080 mm^2) was covered by uniformly distributed mass, imitating soft tissues vibrating with the bone. Three values of the surface density were considered: 0, $2.5 \cdot 10^{-3}$, and $5 \cdot 10^{-3}$ g/mm^2, so the mass of associated soft tissues was respectively 0, 25 or 50 grams.

The BC implant was assumed as a titanium cylinder with a diameter of 3 mm, and a height of 5 mm, which was modeled as a concentrated mass (0.16 g) coupled to the nodes lying on the otic capsule surface and close to the base area of the cylinder (Fig. 1). A harmonic force $F(t)$ with constant amplitude $F_a = 0.1$ N was applied to the mass finite element representing the BC implant. The force direction was denoted as the primary direction in [7], and the amplitude was the same as used in the experiment [3].

RESULTS

Four cases were subjected to harmonic analysis, considering two cases related to mobility of the stapes and two types of POS supports. For each case, three values of the soft tissues mass were considered. The nodes located on the RW edge were used to calculate the averaged movement of the promontory bone, including translation in the direction perpendicular to the RW surface at its center, and rotations about the two axes tangent to that surface. The volume displacement (VD) on the RW membrane was based on the relative displacements and calculated by subtracting the averaged movement of the promontory bone at a given location from the displacements of the nodes located on the RW surface, similarly as in [9]. The relative amplitude on the basilar membrane (BM) in each cross-section of the cochlea was evaluated as the difference between absolute vibrations of two points: one in the center of the BM and the other located at the base of the spiral lamina (SL) [7]. The considered direction of vibration was parallel to the modiolus axis. All quantities were computed and processed as complex numbers.

Volume displacement on the round window membrane

The VD amplitudes at low frequencies up to 630 Hz did not depend on both the mobility of the stapes and POS support (Fig. 3). If the stapes was immobilized, the amplitudes increased a few times for frequencies above 0.8 kHz for the free POS or above 1.25 kHz for the fixed POS, relatively to the results obtained for the movable stapes. For frequencies greater than 2.5 kHz, the VD amplitudes for the immobilized stapes were nearly independent on the POS state (they were slightly greater if the POS was fixed).

The VD amplitudes at high frequencies of 8 and 10 kHz depended significantly on the stapes mobility. For the immobilized stapes, VD amplitudes were of the same level as obtained for low frequencies. For the movable stapes, VD amplitudes were of one order less than those obtained for low frequencies. If the stapes was movable, the amplitudes were slightly greater for the fixed POS than for free POS.

The increase in the mass of soft tissues shifted the maximum of VD amplitude towards lower frequencies, which can be clearly seen for the fixed POS and immobilized stapes: 2 kHz for 50 g, 2.5 kHz for 25 g, and 3.15 kHz - if the mass was not included.

The widest phase angle range for VD was obtained with the case of movable stapes and free POS conditions (Fig. 3). In the model with movable stapes and fixed POS, the phase angle fluctuations occurred at higher frequencies, relatively to the free POS. This effect was caused by the stiffness increase of the structure due to an additional support. In the model with immobilized stapes and fixed POS, the phase drop started at higher frequencies, than for the free POS.

Vibration of the promontory bone

The displacements of the cochlear promontory bone, as expected, did not depend on the mobility of the stapes and were shown only in two cases related to the POS state (Fig. 4). The promontory vibration amplitude was almost constant at low frequencies (below 0.63 Hz for the free POS and at 1.25 kHz for fixed POS), but the values obtained for fixed POS were about 5 times smaller than for free POS. Displacements for frequencies for frequencies above 5 kHz were quite similar for both conditions applied on the POS.

The resonant frequencies of the temporal bone for the case with the fixed POF were around 5 kHz, depending on the value of the associated mass of soft tissues. For the free POS, the resonance occurred at lover frequencies between 1 and 2 kHz. The course of the vibration phase for the free POS indicated several resonances in the

considered frequency range. For both types of boundary conditions applied on the POS, the phase decline started at a lover frequency for a greater value of the soft tissues mass.

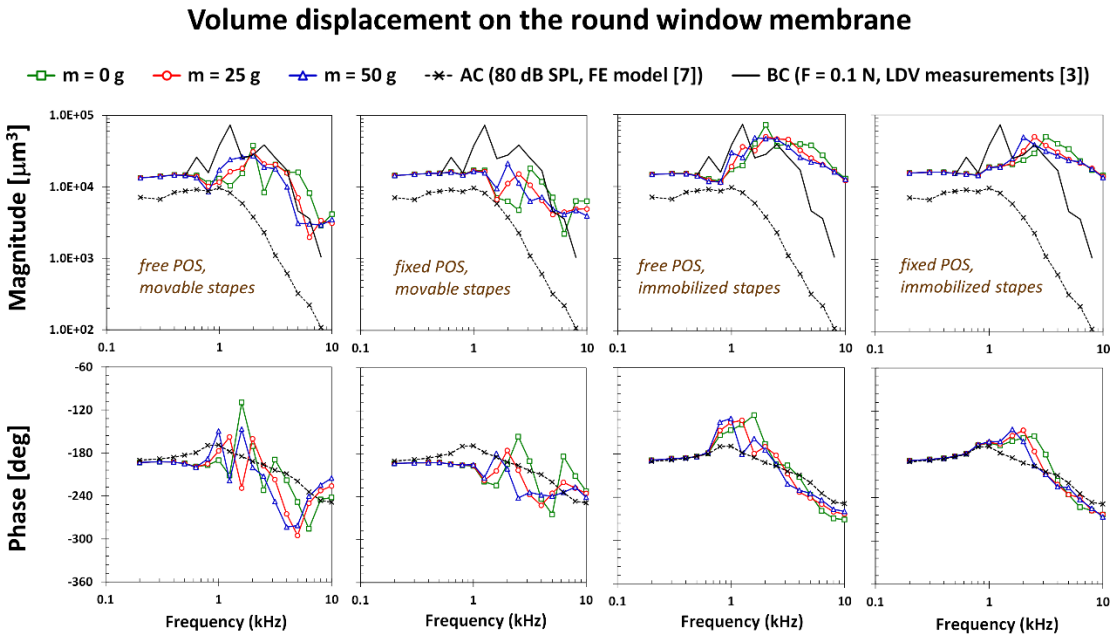


FIGURE 3. The volume displacement (VD) at the round window membrane (RW) for two types of supports related to the petro-occipital synchondrosis (free POS and fixed POS), and two conditions related to the mobility of the stapes (movable and immobilized), for three masses of soft tissues: AC - results of air-conduction stimulation applied on the stapes footplate in the presented FE model [7], BC - results of bone conduction stimulation conducted on a human cadaver under the same force amplitude[3].

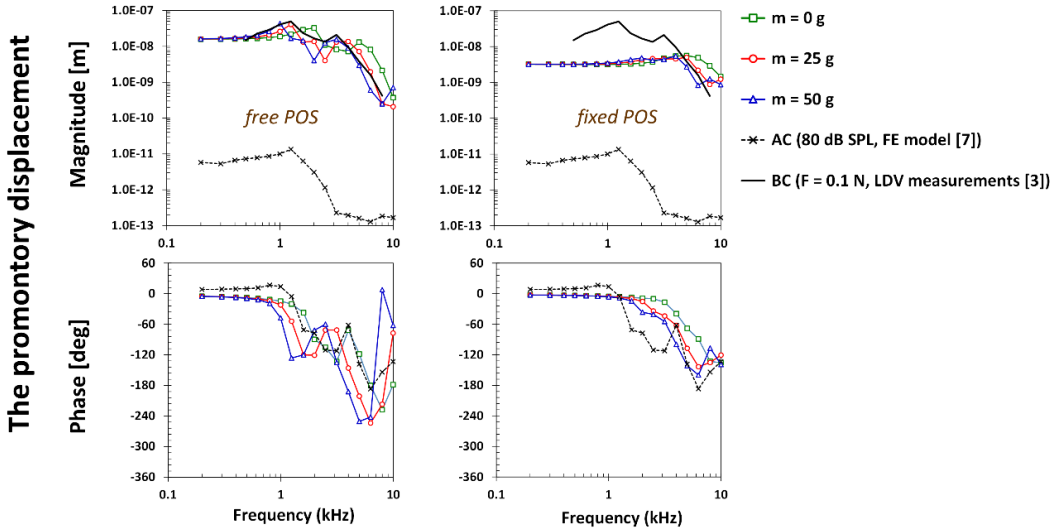


FIGURE 4. The promontory displacement in the direction perpendicular to the round window membrane (RW) for two types of supports related to the petro-occipital synchondrosis (free POS and fixed POS). The results did not depend on the mobility of the stapes. AC - results of air-conduction stimulation applied on the stapes footplate in the presented FE model [3], BC - results of bone conduction stimulation conducted on a human cadaver under the same force amplitude [3].

Vibrations on the basilar membrane and the spiral lamina along the cochlea

The displacements on BM along the cochlea were presented as absolute amplitudes, marked in Fig. 5 with a single solid line, and as relative amplitudes (double solid line). To compute the relative amplitude on the BM, the vibration at the SL base (dotted line) was subtracted from the absolute amplitude of the BM center in each cross-section of the cochlea.

For frequencies below 0.8 kHz, the SL vibrated more if the POS was free, but relative amplitudes at the SL were not dependent on the stapes condition, nor on the POS fixation and the mass of associated soft tissues. The position of the maximum absolute vibration at the BM was shifted towards the base of the cochlea, while the position of the relative amplitude agreed with the Greenwood function [10]. At 1 kHz and for the free POS, the SL amplitudes were 2-3 times greater, relatively to the fixed POS, but the relative amplitude at the BM was greater only by about 40%.

Vibration on the SL at 2.5 kHz significantly depended on the mass of soft tissues. If the mass was equal to 0, the amplitude at the SL base for the case with the fixed POS was one fourth the value obtained for the free POS. Conversely, if the mass was 25 g, the SL amplitude for the fixed POS was two times greater (Fig. 5). For the mass of 50 g, SL amplitudes were small and not dependent on POS conditions. The relative amplitude on the BM at 2.5 kHz was also dependent on the mass of soft tissues. For movable stapes and mass equal to zero, the amplitude was 60% of the value obtained for the immobilized stapes, and the results were similar for both free and fixed POS. For the mass values of 25 g and 50 g, the relative amplitude on the BM was several times smaller if the stapes was movable, and for 25 g depended also on the POS state, resulting in a greater amplitude for the fixed POS (Fig. 5).

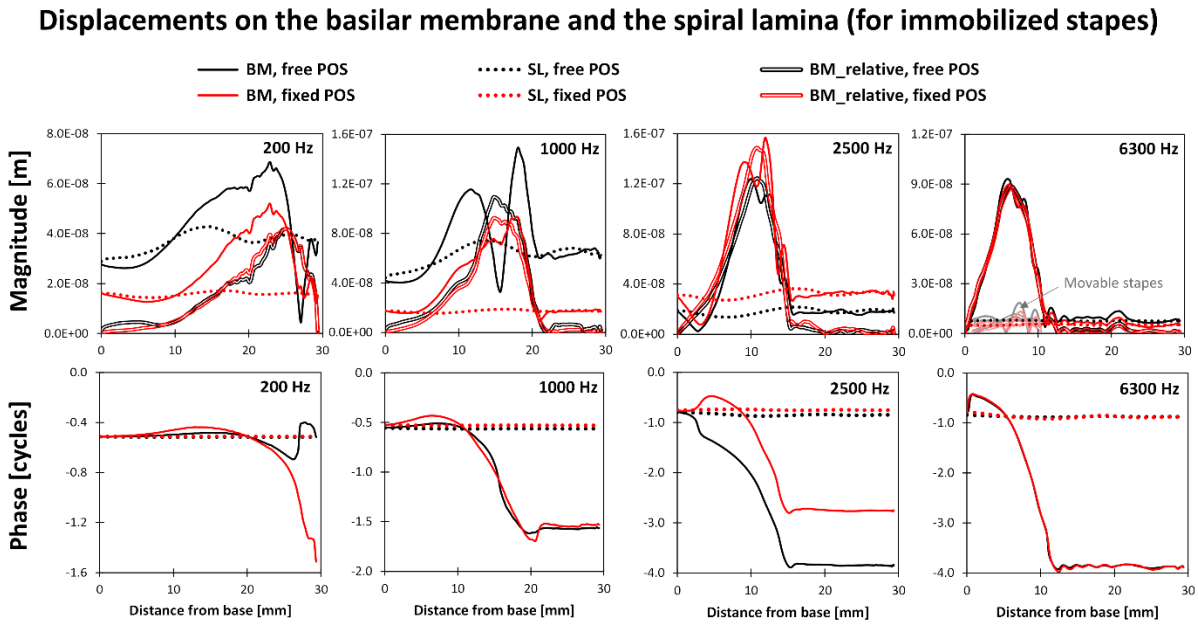


FIGURE 5. The basilar membrane (BM) displacement in the direction parallel to the modiolus axis, for two conditions of the petro-occipital synchondrosis (free and fixed POS). The stapes was immobilized. The results were shown for selected frequencies (0.2, 1, 2.5 and 6.3 kHz) as functions of the distance measured along the cochlea. The relative amplitudes were obtained by subtracting the two complex numbers representing vibrations at the BM center and at the base of the spiral lamina (SL). The mass of soft tissues was 25 g.

Vibration amplitudes on the SL at 6.3 kHz were small compared to those obtained for lover frequencies (Fig. 5). They were independent on the POS state, except for the case without the soft tissues mass, when the amplitude at the apex part of the cochlea was twice as high for the fixed POS. The relative amplitudes on the BM at 6.3 kHz in the case of the immobilized stapes were nearly the same for both free and fixed POS. On the other hand, for a movable stapes and no mass of soft tissues, the relative vibrations on the BM were only 15% (free POS) or 30% (fixed POS) of the value obtained for the immobilized stapes.

For the mass value of 25 g, SL vibrations at 6.3 kHz were small for both conditions applied on the POS (30% lover for the fixed POS, Fig. 5). Distributions of BM vibrations along the cochlea for the case with the immobilized stapes at 6.3 kHz were close to the result of air conduction stimulation [7]. Conversely, BM amplitudes for the movable stapes were significantly suppressed in the basal region of the cochlea as was shown in Fig. 5, both for fixed and free POS.

The SL and BM vibrations at 6.3 kHz and with mass 50 g were slightly smaller than those obtained for 25 g, except the case of movable stapes and free POS, when the travelling wave in the BM did not occur.

The phases of absolute vibrations on the BM and SL were similar for both conditions applied on the POS at the selected frequencies except 2.5 kHz, when the difference between free and fixed POS was about 1 cycle (Fig. 5).

DISCUSSION

The boundary conditions considered in this work at the petro-occipital synchondrosis can be treated as extreme. As shown in [8], the POS state may change depending on age: no ossification was observed in a twenty-year-old, and advanced ossification - in an eighty-year-old. In the presented work, the fixed POS caused reduction of the vibration amplitude at the cochlear promontory bone. It should be mentioned here that even if the POS is ossified, under real conditions it can vibrate along with the occipital bone. Since the occipital bone has a certain compliance, the resonant frequencies obtained for the fixed POS seem to be overestimated.

Resonant frequencies of the petrous part of the temporal bone depended on boundary conditions applied at the POS and were different than the resonant frequencies of the middle ear. It was seen, that in the case of free POS and movable stapes, there was an increase in the volume displacement on the RW membrane at medium frequencies, while for the fixed POS and immobilized stapes – the increase occurred at higher frequencies (Fig. 3). The results obtained for the free POS were closer to the experimental results [3].

In this work, the conditions related to the stapes mobility were related either to the exclusion of one of the bone conduction pathways (effect of inertia of the middle ear - for an immobilized stapes), or to the assumption of the non-physiological and reduced impedance at the oval window (for a movable stapes). When the stapes was immobilized, it was easier to interpret the results, as one of the bone conduction pathways was eliminated. On the other hand, this case reflected the otosclerotic state and was of practical importance. The condition of the movable stapes modified the system response mainly at medium frequencies, where the middle ear plays an important role [5]. The resonant frequencies calculated for the movable stapes were certainly not the same as those obtained considering the entire middle ear - therefore should be treated with caution. For the movable stapes, the maximum of the VD amplitude was shifted towards lower frequencies, particularly if the POS was fixed (Fig. 3).

Results of bone conduction stimulation for low frequencies showed that the mobility of the stapes had no effect on vibrations on round window and basilar membranes. Since VD amplitudes at high frequencies were greater when the stapes was immobilized, also the relative BM vibrations were higher for that case. The results obtained for the model with the movable stapes were characterized by a moderate decrease in the BM vibrations at medium frequencies and a radical decrease at high frequencies, relatively to the case with the immobilized stapes.

In the presented FE model, the cortical bone volume was $8.9 \cdot 10^3$ mm^3, which at a density of 1.8 g/cm^3 gave a mass of 16 g. The volume of the trabecular bone was $9.78 \cdot 10^3$ mm^3, which gave a mass of 12 g with a density of 1.2 g/cm^3, i.e., the mass of the temporal bone was 28 g (not counting the fluid and compliant structures in the cochlea). Thus, the soft tissues mass of 25 g was nearly the same as that of the temporal bone, while for 50 g it was approximately twice as large. The effect of the mass on the VD vibration was most noticeable for the case with the movable stapes at medium frequencies, as it influenced the difference between the resonant frequencies of the temporal bone and the resonant frequencies of the cochlea. The influence of the soft tissue mass on the BM vibration was most significant for the model with the moving stapes at high frequencies. As the mass increased, the BM vibration decay progressed, while the POS status was negligible. If the stapes was immobilized, the amplitudes on the BM were several times higher, and both the type of POS support and the mass of soft tissues did not matter.

The results obtained from the numerical simulation were fitted to the results of the experiment [3] and the best match was obtained for the free support of the petro-occipital synchondrosis and the soft tissues mass of 25 g. Unfortunately, the experiment [3] was carried out only on one cadaver and the direction of stimulation could also slightly differ, hence these results should be treated with caution. Contrary to the experiment, the presented FE model did not consider the middle and outer ear structures, which had a significant impact on the differences between the simulation results and the experimental data.

CONCLUSIONS

The fixed support applied at the petrous-occipital synchondrosis caused reduction of the cochlear promontory vibrations, especially at low and medium frequencies, relatively to the free POS. The volume displacement on the round window membrane did not depend on the POS support at low and high frequencies. In the cochlea, the state of synchondrosis had a greater effect on the bone vibration amplitudes at the spiral lamina, than on the relative amplitudes on the basilar membrane. The SL vibrations were similar as the bone vibrations in the vicinity of the cochlea and did not depend on the mobility of the stapes. Overall, the BM vibrations were affected mainly by the stapes mobility.

The mass of soft tissues turned out to be significant mainly for medium and high frequencies. The mass increase resulted in a decrease in the resonant frequencies of the temporal bone, which particularly affected VD and BM vibrations in the case of a movable stapes, therefore the middle and outer ear structures should be considered in further numerical simulations.

The selected conditions (soft tissues mass of 25 g, free POS and immobilized stapes) were then used in the temporal bone model to compare the two directions of bone conduction stimulation [7].

ACKNOWLEDGMENTS

This work was partially supported by the National Centre for Research and Development grant number PBS3/B7/25/2015.

REFERENCES

1. T. Wojciechowski, M. Lachowska, K. Niemczyk, “Detailed Analysis of the Space between the Semicircular Canals for the Purpose of Direct Bone Conduction Stimulation of the Inner Ear,” Audiol. Neurootol. **26**, 35-44 (2021).
2. B.E. Håkansson, M. Eeg-Olofsson, S. Reinfeldt, S. Stenfelt, G. Granström, “Percutaneous versus transcutaneous bone conduction implant system: a feasibility study on a cadaver head,” Otol. Neurotol. **29**, 1132-1139 (2008).
3. M. Kwacz, K. Niemczyk, J. Wysocki, “Round Window Membrane Motion Induced by Bone Conduction Stimulation at Different Excitation Sites: Methodology of Measurement and Data Analysis in Cadaver Study,” Ear and Hearing **40**(25), 1437-1444 (2019).
4. I. Dobrev, S. Stenfelt, C. Röösli, L. Bolt, F. Pfiffner, R. Gerig, A. Huber, J. H. Sim, “Influence of stimulation position on the sensitivity for bone conduction hearing aids without skin penetration,” Int. J. Audiol. **55**(8), 439-46 (2016).
5. S. Stenfelt, “Model predictions for bone conduction perception in the human.” Hear Res. **340**, 135-143 (2016).
6. Y. Chang, N. Kim, S. Stenfelt, “Simulation of the power transmission of bone-conducted sound in a finite-element model of the human head. Biomech. Model. Mechanobiol,” **17**,1741-1755 (2018).
7. P. Borkowski, P. Marek, K. Niemczyk, M. Lachowska, M. Kwacz, J. Wysocki, “Bone conduction stimulation of the otic capsule: a finite element model of the temporal bone”, Acta Bioeng. Biomech. **21**(3), 75-86 (2019).
8. A. L. Balboni, T. L. Estenson, J. S. Reidenberg, A. D. Bergemann, J. T. Laitman, “Assessing age-related ossification of the petro-occipital fissure: laying the foundation for understanding the clinicopathologies of the cranial base. The anatomical record. Part A, Discoveries in molecular, cellular, and evolutionary biology,” **282**(1), 38–48 (2005).
9. S. Stenfelt, S. Puria, N. Hato, R. L. Goode, “Basilar membrane and osseous spiral lamina motion in human cadavers with air and bone conduction stimuli,” Hear. Res. **181**(1–2), 131–43 (2003).
10. D.D. Greenwood, “A cochlear frequency-position function for several species-29 years later,” J. Acoust. Soc. Am. **87**, 2592–2605 (1990).

Difference in bone conduction analysis between a head model and an isolated cochlea model

Jongwoo Lim[1,a)] , Jeong-Hwan Kim[2, b)], Yong-Jin Yoon[2, c)] and Namkeun Kim[1,d)]

Author Affiliations

[1]*Department of Mechanical Engineering, Sogang University, Seoul, Republic of Korea*
[2]*Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology, Daejeon, Republic of Korea.*

Author Emails
[a)]imbellcow94@gmail.com
[b)]kjhksa17@kaist_ac.kr
[c)]yongjiny@kaist.ac.kr
Corresponding author: [d)]nkim@sogang.ac.kr

Abstract. The bone deformation and the inertia from cochlear fluid or middle ear (ME) ossicles are generally considered as two significant components causing bone conducted (BC) hearing among several components. In order to clarify the BC hearing mechanism and the important components, finite element (FE) models have been widely used. The FE model can be divided into two types, (1) isolated cochlea model composed of ME and cochlea including basilar membrane (BM) and (2) head model including the isolated cochlea structure. Using the two types of the FE model, specific pathological condition such as otosclerosis was simulated and the results were compared with clinical data. If we assume the BC hearing ability is proportional to the BM velocity in the FE simulation, the bone deformation and the inertia components were still more important factors to cause the BC hearing than cartilage near an ear canal or pressure from cerebrospinal fluid. However, when the simulated results were reasonably consistent with the clinical data, the inertia effects on BC hearing were magnified in the isolated cochlear model whereas the bone compression effect was intensified in the head model although the bone compression was not the most significant factor for BC hearing. The results suggest that care is needed not only in the analysis of BC experiments using a temporal bone but also in applying the results to patients directly.

1. INTRODUCTION

Temporal bone or cadaver head has been generally used in experimental bone-conducted (BC) hearing analysis, and the experiments have contributed to understanding of BC hearing mechanism. For example, Stieger *et al.* (2018) used the temporal bones to measure promontory velocities and intracochlear pressure which are important cochlear responses to elucidate the BC hearing mechanism[1]. Specifically, they varied the method to hold the temporal bones and investigated the effect of the holding method on BC responses represented by promontory velocities. Stenfelt *et al.* (2002) have measured the motion of the umbo and stapes footplate during BC stimulation induced by shaker using the temporal bones[2]. In the result, they showed that the stapes footplate is closely coupled to the temporal bone more than the malleus. On the other hand, there have been several experiments to measure BC responses using cadaver heads. Eeg-Olofsson *et al.* (2018) measured mechanical point impedance and promontory velocities from cadaver heads to clarify the BC stimulation position causing the largest cochlea response[3]. Furthermore, the study was expanded by measuring cochlea promontory velocity in living human, and concluded that the average promontory

velocity of living human was similar that of the cadaver head[4]. There is another study that compared promontory velocity between percutaneous and transcutaneous BC implants to elucidate the skin effect on BC hearing[5]. In addition, Dobrev *et al.* (2019) measured the 3D promontory velocities using a 3D laser doppler vibrometer (LDV) system, and concluded that the 3D velocities decreased the inter-sample variability more than the 1D velocity[6–8]. From most cadaver experiments, we could observe a common factor. Since it was physically difficult to measure cochlear responses such as basilar membrane (BM) velocity or intracochlear pressure from a cadaver head in BC hearing, most experimental results from the cadaver head in BC hearing were based on the measurement of the promontory velocities with assumption that the promontory velocity is proportional to one's BC hearing ability. Therefore, although all the experiments above were important and helpful for one's understanding of BC hearing mechanism, there has been little studies to investigate the difference in BC responses between a cadaveric head and a temporal bone from which the cochlear responses were usually measured.

In this study, a 3D finite element (FE) model of a whole head containing the auditory periphery was used to investigate the different effects on BC responses between a temporal bone and a cadaver head. Specifically, the model was validated by comparison of hearing loss with clinical data obtained from otosclerosis patients. In addition, through the otosclerosis condition, bone compression effects on BC hearing were investigated in both a temporal bone model and a head .

2. METHOD

2.1. Finite element model

The 3D FE model was used to implement a temporal bone and a cadaver head. The temporal bone was simulated by the isolated cochlear model consisting of middle-ear structures, cochlear fluid, outer bony shell, and BM. On the other hand, the cadaver head model consists of skin, skull, eyes, brain, and soft tissues. Specifically, the head model contains the isolated cochlear model at the proper position in the head. Although the detailed descriptions for both models are described in the previous studies, we provide a brief explanation of the model for better understanding[9–11].

The FE model of the head was generated by 3D reconstruction using the cryosectional images of a cadaver head and consists of about 258,000 and 144,000 solid- and fluid-tetrahedron elements, respectively. In the case of the isolated model, the μCT images were used to generate geometry of the auditory periphery. The isolated model was composed of about 104,000 and 507,000 solid- and fluid-tetrahedron elements. Furthermore, the BM consists of about 18,000 solid-pentahedron elements whose top and bottom surfaces are triangles. The FE models for the head and the isolated model are shown in Fig. 1(a) and (b), respectively.

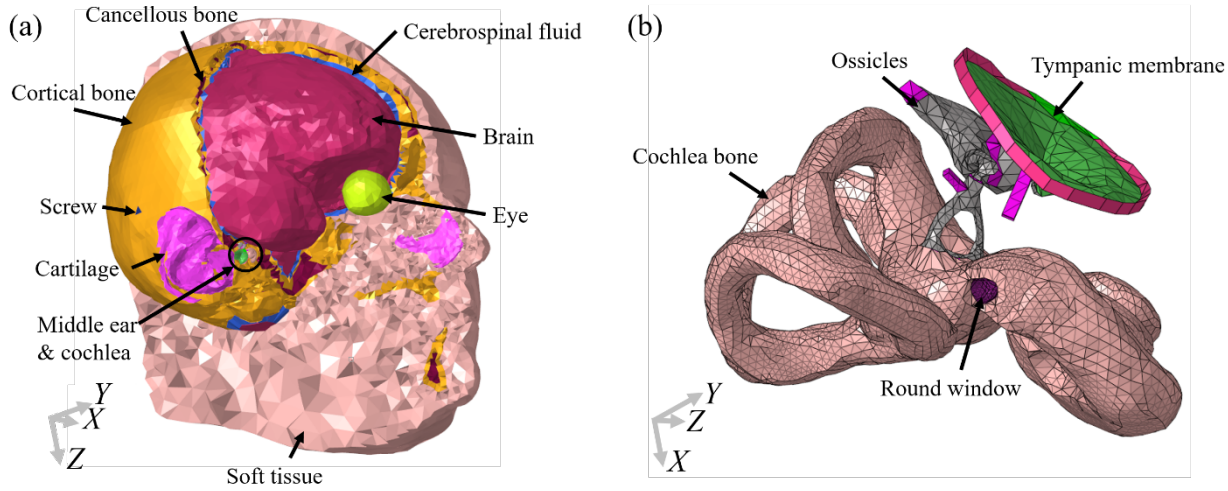


Figure 1. The FE model for the head and the isolated model. (a) The head model has eight components: soft tissue, eye, cortical bone, cancellous bone, cerebrospinal fluid (CSF), brain, and screw. (b) The auditory periphery consisting of the middle ear and cochlea. It is assumed that outer cochlear bone and end of the ligaments are fully connected with the temporal bone.

2.2. Bone-conduction stimulation

In this study, the simulation was conducted in the frequencies from 0.1 to 10 kHz with 0.1 kHz increment. The FE simulation was performed through the commercial software, ACTRAN (MSC software, Newport Beach, CA, USA). Two different BC inputs were used to implement BC hearing. The first input was assigning the sinusoidal force on the screw element near the mastoid position. Here, the screw represents the bone anchored hearing aid (BAHA) attached to a head. The sinusoidal force is applied along the longitudinal direction of the screw. The second one was assigning the sinusoidal displacement on a few outer cochlea bone nodes and end of the tympanic annulus and ligaments. The direction of the sinusoidal displacement is the same as sinusoidal force. It should be noted that the first input was applied only to the whole head model whereas the second input was implemented in the isolated cochlear model. Since the surrounding bones of the cochlea was absent in the isolated cochlear model, the bone compression effect was neglected when the isolated cochlear model was used to simulate BC hearing. Therefore, to implement the bone compression effect on the isolated cochlear model, an additional distributed pressure was applied to the outer bony surface of the isolated cochlea. Although an actual pressure wave traveling through the skull is significantly complex, the pressure wave on the cochlea in the simulation was simplified as the plane wave by assumption that the wave length is much larger than the length of the cochlea. The detailed method how to apply the additional pressure can be found in the previous study[12].

As described in the previous study, otosclerosis condition was also considered by assigning high Young's modulus like a bone to the stapes annular ligament in the simulation. Through the otosclerosis condition, we could calculate the hearing loss by the difference of the maximum BM velocity in between a normal condition and otosclerosis condition. The calculated hearing loss was used not only to verify the simulation result but also to adjust the bone compression effect on BC responses by comparison with clinical data showing Carhart's notch[13].

3. RESULT

Figure 3 shows the hearing loss calculated by difference of BM velocity at the best frequency position in between a normal condition and the otosclerosis condition, and the hearing losses were compared with clinical data[13] which was indicated by black-dotted lines in Fig. 3(a) – (c). The black-solid and red-solid lines in Fig. 3(a) represent the cases when only inertia and inertia with bone compression were respectively used as the BC stimulation in the isolated

cochlea. In the isolated cochlea simulation, when pressures as well as inertia were applied as BC stimulation, the calculated hearing loss was similar to the clinical data within about 10 dB differences. More specifically, when only inertia was used for BC stimulation, the hearing loss difference from the clinical data was distinctly observed at low frequencies below 600 Hz. Figure 3(b) shows the hearing loss from the whole head model to which the dynamic force was applied. In this figure, the black-solid line represents the case in which Young's modulus for cortical bone and cochlear bone is 10 GPa whereas the red-solid line indicates the case in which the Young's modulus is 20 GPa. The hearing loss difference from the clinical data becomes larger from the 10 GPa case at above 1 kHz. In the case of 20 GPa of bone Young's modulus, the difference at above 1 kHz was within 10 – 15 dB. In Fig. 3(c), the red-solid and black-solid lines represent the same cases with those in Fig. 3(b). However, in Fig. 3(c), the BC stimulation is not a sinusoidal force by a sinusoidal displacement applied on a few nodes of the skull surface.

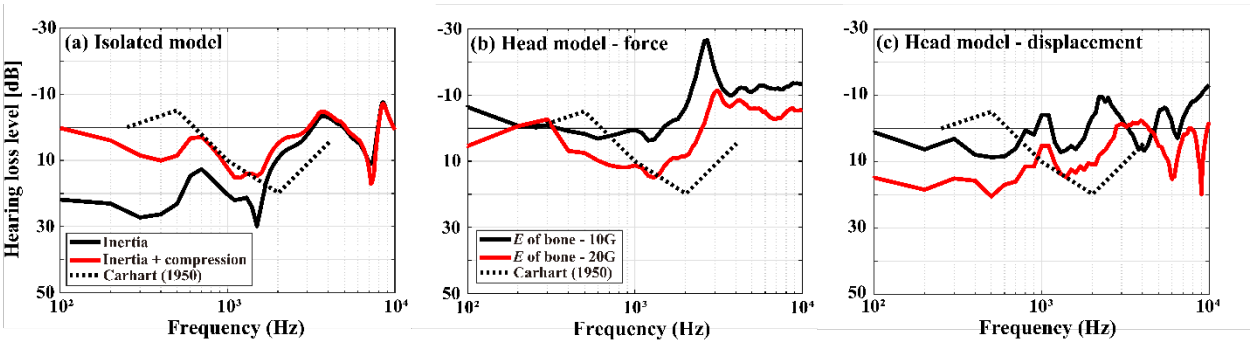


Figure 2. Comparison of the calculated hearing loss level in the simulation with clinical data. The used FE models are (a) the isolated cochlear model, (b) the full head model, and (c) the same full head model. In (a), the black-solid line represents the only inertia input whereas the red-solid line indicates the inertia and bone compression inputs. In (b) and (c), the black-solid line means the case when the Young's modulus for cortical bone and cochlear bone was 10 GPa. On the other hand, the red-solid line describes the case when the Young's modulus was 20 GPa. In all the figures, the black-dotted line represents the clinical data obtained by Carhart. The BC input is the dynamic force in (b) whereas the input is the sinusoidal displacement in (c). The displacement in (c) was applied on a few nodes of skull surface.

4. DISCUSSION

If we generally accept the existence of Carhart's notch (which is about 10 – 20 dB hearing loss at 1 – 2 kHz) from otosclerosis patients in BC hearing, the Carhart's notch can be a good standard to validate the model describing a live human head. When the isolated cochlea was used for BC simulation, only inertia BC input could not show the consistent hearing loss level with the clinical data as shown in Fig. 3(a). At the low frequencies below 600 Hz, only inertia case showed about 15 – 30 dB hearing loss whereas the clinical data showed about 0 dB hearing loss. To make up the hearing loss at the low frequency, the additional BC input, bone compression, should be considered. The inertia with bone compression could show the similar hearing loss level to the clinical data. This means that the BC experiments using an isolated cochlea with shaker as BC input can have the possibility to ignore the bone compression effect too much at low frequencies below 1 kHz.

On the other hand, the BC simulation results using a whole head with BAHA as a BC source showed consistent hearing loss levels within 10 dB differences with the clinical data at low frequencies below 1 kHz. However, at high frequencies above 1 kHz, the hearing loss levels were highly depended on the bone Young's modulus. When the Young's modulus was 20 GPa, the Carhart's notch was observed. However, the 10 GPa Young's modulus generated the hearing gain rather than hearing loss at above 1 kHz which is different trend from the clinical data. Therefore, the BC experiments using a head can have the possibility to underestimate the bone compression effect on BC hearing at the high frequencies because an old cadaver may cause the stiffened skull properties. According to the previous studies, the Young's modulus of the head was reported in ranges from 4 to 25 GPa. We need to further study through more clinical data to verify whether any BAHA users have in trouble in BC hearing at high frequencies due to the relatively low Young's modulus of the skull bone.

In addition, through Fig. 3(c), it was observed that the sinusoidal displacement in a head caused the larger bone compression effect than the same input in an isolated cochlea. It seems reasonable because the wider surface of the

skull can generate the larger bone compression effect than the isolated cochlea. However, it is still needed to further investigate the difference between combination of inertia and pressure in an isolated cochlea and that in a whole head.

5. CONCLUSION

In this study, the difference between the BC experiments using the cadaveric head and the temporal bone was studied using a 3D head FE model including the auditory periphery. When the hearing loss caused by otosclerosis for different models corresponding to BC inputs as well as geometry was compared with clinical data, the temporal bone showed different results from the cadaveric head in BC hearing. If a temporal bone with a shaker (as a BC source) was used in a BC experiment, researchers should be careful in underestimation of bone compression effect at low frequencies. On the other hand, in the BC experiment using a cadaveric head with a BAHA, mechanical properties of the skull bone can cause more significant bone compression effect on BC hearing than the temporal bone at high frequencies.

ACKNOWLEDGMENTS

This work has been supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF – 2019R1C1C1007907, & NRF – 2017M3A9E2065287).

REFERENCES

1. Stieger C, Guan X, Farahmand RB, et al. Intracochlear Sound Pressure Measurements in Normal Human Temporal Bones During Bone Conduction Stimulation. *JARO - Journal of the Association for Research in Otolaryngology*. 2018;19(5):523-539. doi:10.1007/s10162-018-00684-1
2. Stenfelt S, Hato N, Goode RL. Factors contributing to bone conduction: The middle ear. *J Acoust Soc Am*. 2002;111(2):947-959. doi:10.1121/1.1432977
3. Eeg-Olofsson M, Stenfelt S, Tjellstrom A, Granstrom G. Transmission of bone-conducted sound in the human skull measured by cochlear vibrations. *International Journal of Audiology*. 2008;47(12):761-769. doi:10.1080/14992020802311216
4. Eeg-Olofsson M, Stenfelt S, Granström G. Implications for contralateral bone-conducted transmission as measured by cochlear vibrations. *Otology and Neurotology*. 2011;32(2):192-198. doi:10.1097/MAO.0b013e3182009f16
5. Håkansson B, Eeg-Olofsson M, Reinfeldt S, Stenfelt S, Granström G. Percutaneous Versus Transcutaneous Bone Conduction Implant System: A Feasibility Study on a Cadaver Head. *Otology and Neurotology*. 2008;29(8):1132-1139. doi:10.1097/MAO.0b013e31816fdc90
6. Dobrev I, Sim JH. Magnitude and phase of three-dimensional (3D) velocity vector: Application to measurement of cochlear promontory motion during bone conduction sound transmission. *Hearing Research*. 2018;364:96-103. doi:10.1016/j.heares.2018.03.022
7. Dobrev I, Sim JH, Pfiffner F, Huber AM, Röösli C. Experimental investigation of promontory motion and intracranial pressure following bone conduction: Stimulation site and coupling type dependence. *Hearing Research*. 2019;378(xxxx):108-125. doi:10.1016/j.heares.2019.03.005
8. Dobrev I, Farahmandi TS, Röösli C. Experimental investigation of the effect of middle ear in bone conduction. *Hearing Research*. 2020;395:108041. doi:https://doi.org/10.1016/j.heares.2020.108041
9. Chang Y, Kim N, Stenfelt S. The development of a whole-head human finite-element model for simulation of the transmission of bone-conducted sound. *J Acoust Soc Am*. 2016;140(3):1635-1651. doi:10.1121/1.4962443

10. Kim N, Steele CR, Puria S. The importance of the hook region of the cochlea for bone-conduction hearing. *Biophysical Journal*. 2014;107(1):233-241. doi:10.1016/j.bpj.2014.04.052
11. Lim J, Dobrev I, Röösli C, Stenfelt S, Kim N. Development of a finite element model of a human head including auditory periphery for understanding of bone-conducted hearing. *Hearing Research*. 2021;(xxxx):108337. doi:10.1016/j.heares.2021.108337
12. Kim N, Steele C, Puria S. Carhart's notch: A window into mechanisms of bone-conducted hearing. In: *Proceedings of Meetings on Acoustics*. Vol 19. ; 2013. doi:10.1121/1.4799725
13. CARHART R. CLINICAL APPLICATION OF BONE CONDUCTION AUDIOMETRY. *Archives of Otolaryngology*. 1950;51(6):798-808. doi:10.1001/archotol.1950.00700020824003

Finite-element modeling of the effect of superior canal dehiscence on intracochlear pressures in bone conduction

Junfeng Liang[1], Hamid Motallebzadeh[2,3], Sunil Puria[2,3], Xiying Guan[1, a)]

[1]*Department of Communication Sciences and Disorders, Wayne State University, Detroit, MI 48202, USA*
[2]*Department of Otolaryngology – Neck and Head Surgery, Harvard Medical School, Boston, 02114, MA USA*
[3]*Eaton-Peabody Laboratories, Massachusetts Eye & Ear, Boston, 02114, MA USA*

[a)] Corresponding author: Xiying_Guan@wayne.edu

Abstract. Superior canal dehiscence (SCD) is a pathological opening of the bone encapsulating the superior semicircular canal, causing various vestibular and auditory symptoms. One of the most debilitating symptoms associated with SCD is bone conduction (BC) hyperacusis, manifested as awareness of self-generated sounds (one's own heartbeats, voices, joint movements, eye movements) and increased sensitivity to certain external sounds and vibrations (machinery, noise experienced inside an automobile) that can be heard via BC. While BC hyperacusis is increasingly recognized by clinicians and researchers, how SCD enhances BC sensitivity is not well understood.

To better understand the underlying mechanism, we simulated the effect of SCD on intracochlear sound pressures in BC using a finite element (FE) model of a human ear. SCD was modeled by creating an opening in the bony wall in the middle of the superior canal arch. The model was first validated with air conduction (AC) excitation. BC excitation was subsequently simulated by applying rigid-body vibrations to the model along the stapes piston-motion direction and the other two orthogonal directions. The BC-evoked intracochlear sound pressures in scala vestibuli (P_{SV}) and scala tympani (P_{ST}) at the basal cochlea and the pressure difference ($P_{DIFF}=P_{SV}-P_{ST}$) were simulated under normal and SCD conditions.

The comparison between the simulated and measured intracochlear pressures in AC suggests that the damping of the cochlear channels (scala vestibuli and scala tympani) in the current model is lower than that of a real ear. In BC, the model predicts that SCD increases P_{DIFF} – the cochlear input drive – for all three excitation directions, indicating an increase in BC sensitivity. However, the simulated changes in P_{SV} and P_{ST} due to SCD in BC are inconsistent with our previous measurements and are highly dependent on the excitation direction. The discrepancy between the simulated results and the experimental data may stem from this direction-dependent nature of the SCD effect and insufficient damping of the cochlear fluid channels.

INTRODUCTION

Superior canal dehiscence (SCD) describes the pathological opening of the bone covering the superior semicircular canal (SCC). This defect is mostly found along the floor of the middle fossa near the arcuate eminence. The dehiscence results in a "mobile third window" of the inner ear, with the first and second window referring to the round and oval windows [1, 2]. This third window enables a leakage pathway in the vestibular system that diverts sound from the cochlea, resulting in a constellation of hearing and vestibular disorders [3-6].

One of the disorders associated with SCD is increased sensitivity to bone-conducted sounds. This disorder is referred as bone conduction (BC) hyperacusis, manifesting as autophony (hearing one's own voice as loud or distorted); pulsatile tinnitus; perception of self-generated internal sounds (audible eye movements, footsteps, chewing, bowel movements); and hypersensitivity to external vibrations transmitted through the body (machinery noise experienced inside a car) [7-10]. It is believed that BC hyperacusis results from an altered BC inner-ear mechanism due to the SCD. A few BC inner-ear models, including lumped-element models[11-13] and finite-element models[14, 15], have been created to better understand the BC mechanism and explain BC hyperacusis. While they predicted that SCD increases BC hearing, those models were not fully validated due to the lack of inner-ear measurements at the time they were published.

Recently, we measured BC-evoked intracochlear pressures in the scala vestibuli (P_{SV}) and scala tympani (P_{ST}) at the basal cochlea of human cadaveric ears under normal and SCD conditions[11]. The study revealed that SCD increased P_{SV} by ~10 dB on average below 1 kHz, but increased P_{ST} only by up to 4 dB on average. We created a lumped-element model based on the previous work by Stenfelt[16] and were able to replicate the experimental results. However, due to its intrinsic limitation, the model is unable to simulate the spatial distribution of sound pressure or the flow of sound in the inner ear for BC, which could help understand how sound is generated within the inner ear in BC. In addition, the circuit network, the components, and the parameters of the components in the lumped-element model were determined not only by the simplified inner-ear geometry but also by the direction of the BC vibration. Thus, the lumped-element model is not useful in studying stimulation direction dependency for the intracochlear pressures and SCD's influence, which is important because BC excitation can be three-dimensional.

In this paper, we presented our recent development of a human ear FE model to study the SCD effects on BC-elicited intracochlear pressures in three BC excitation directions. We validated the model using our previous air-conduction intracochlear pressure measurements. Then we simulated SCD effects on intracochlear pressures in BC and compared the simulated results with our BC experimental results. The development of this FE model is an important step toward fully understanding the mechanisms underlying BC and BC hyperacusis.

METHOD

The geometry of the previous human ear FE model created by Kim et al.[14] was extracted from Actran and slightly modified. As shown in Fig.1, the model in the present study

consists of the ear canal, the ossicular chain, the supporting tenons, the soft tissues connecting the ossicular joints, a coiled cochlea and the SCCs. The cochlea ducts were simplified with only two compartments: scalar vestibuli (SV) and scalar tympani (ST). The organ of Corti was simplified as basilar membrane alone, separating SV and ST. The membranous labyrinth in the vestibule was merged with the perilymph fluid space.

Compared with the original model, the tympanic membrane (TM), the basilar membrane (BM), and the round window membrane (RWM) were modified as thin shells. The variation of the thickness and the width of the BM was applied using the values reported in the literature. The stapes tendon was removed to mimic the experimental preparation for intracochlear pressure measurements[11,17]. COMSOL Multiphysics 5.6 (COMOL, Inc, Burlington, MA, US) was used to mesh the model geometry, assign the material properties, apply the loading and the boundary conditions, implement the simulation, and post-process the computational results.

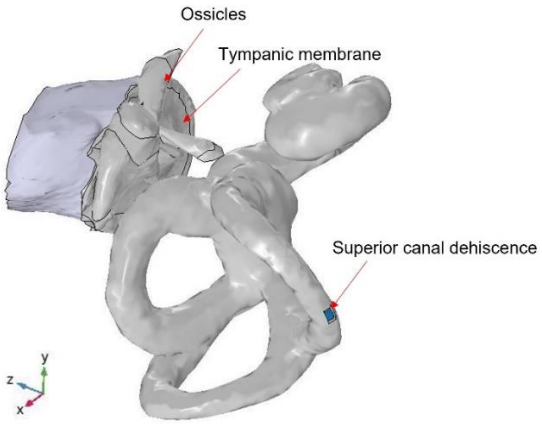


FIGURE 1. The FE model of a human ear.

Tetrahedral elements with maximum size of 2.42 mm were used to mesh the geometry. The TM, BM and the RWM were meshed with triangular shell elements with total number of 146,006, respectively. The mesh was automatically generated with COMSOL Multiphysics code. The ear canal was also filled with tetrahedral elements that represents air.

The material properties of the tissues in the model were from three sources: 1) tissue properties measured and reported in the literatures; 2) material properties used in the previous FE models; and 3) previous parametric studies that reconcile numerical results with the experimental counterpart. Specifically, the ossicles and the bony shell of the inner ear were considered as bone material with isotropic elastic properties. All the soft tissues were modeled as linear viscoelastic materials. The TM was modeled as anisotropic material with Young's modulus different in radial, circumferential, and through-thickness directions. The constitutive model for the BM was orthotropic with transverse (out-of-plane) elasticity much higher than the other two directions. The transverse Young's modulus of the BM also varies from base to apex to simulate the passive tuning effect of

the organ of Corti. Material properties of water was used for the inner-ear fluid. The values for the material properties of all the components were listed in Table 1.

TABLE 1. Material properties used in the FE model

Component	Young's modulus (Mpa)	Density (kg/m3)	Loss factor
Malleus[18]	1.4×10^4	2.39×10^3	0
Incus[18]	1.4×10^4	2.15×10^3	0
Stapes[18]	1.4×10^4	1.1×10^3	0
Malleus/incus joint[19]	7	1.2×10^3	0.15
Incus/stapes joint[18]	0.44	1.2×10^3	0.15
Tensor tympani[19]	5	1.2×10^3	0.3
Ligaments[19]	2	1.1×10^3	0.15
Tympanic membrane[20]	140, 75, and 20*	1.1×10^3	0.1
Tympanic annulus[21]	0.6	1.2×10^3	0.5
Stapes annular ligament[22]	0.7	1.2×10^3	0.3
Round window	1	1×10^3	0.15
Cochlear bone	1.4×10^3	2.39×10^3	0
Cochlear fluid[24]	--	1×10^3	--
Basilar membrane[22]	6.5~5.5, $0.2\sim1\times10^3$**	1×10^3	0.3

*in radial circumferential, and thickness direction; ** in longitudinal and transverse direction

Two series of simulations were conducted using the model: First, the SCD effects on AC-elicited intracochlear pressures were simulated. Second, the SCD effects on BC-elicited intracochlear pressures were simulated in three BC stimulation directions.

In the AC simulation, sound pressure of 90 dB was delivered at the entrance of the ear canal in the form of a plane wave. The harmonic analysis was implemented to solve the equilibrium equations at frequency from 100 Hz to 10 kHz. The velocity of the stapes (V_{STAP}), P_{SV}, and P_{ST} were computed.

In BC stimulation, a rigid-body vibration with velocity magnitude of 1×10^{-4} m/s was applied on the boundary of the TM annular ring, external ends of the tendons and the bony shell of the inner ear. The rigid-body excitation was applied in three orthogonal directions with the z direction aligned approximately with the stapes piston-motion direction (Fig 1). P_{SV}, and P_{ST} in each BC excitation direction were computed.

To simulate the SCD, a rectangular area with the size of 1 mm by 0.5 mm in the bony shell of the superior canal was removed. A zero-pressure boundary condition was applied to fluid surface within the dehiscence in both AC and BC simulations.

RESULTS

Air-conduction results

Fig. 2 shows the measured and simulated P_{SV}, P_{ST}, and P_{DIFF} normalized to V_{STAP} in AC under the control and SCD condition. The solid black and red lines are the medians of the measurements made from six ears before and after an SCD was created in the previous

study (Guan et al. 2020). The dashed black and red lines are simulated results for the normal and SCD conditions, respectively. The upper panels show the magnitude and the lower panels the phase.

As shown in Fig. 2A and 2B, the measured P_{SV} in normal ears (solid black line) varies little with frequency and its phase is close to zero, suggesting the cochlear impedance is predominantly resistive. SCD decreases P_{SV} mainly below 1000 Hz and the lower the frequency the greater the reduction (solid red line vs solid black line). The simulated P_{SV} in the normal ear (dotted black line) shows a dip below 1000 Hz, inconsistent with the measurement. The phase of the simulated P_{SV} starts at -0.25 cycle at 100 Hz, lower than the measurement, suggesting that the damping of the cochlea is lower than that in a real ear. The discrepancy may result from the simplification of cochlea from three compartments to two compartments, where the cross-section area of scala vestibuli is increased thus its damping is reduced. In the simulation, SCD decreases P_{SV} below 3000 Hz, and the decrease is substantial at 100-300 Hz (dotted black line vs dotted red line).

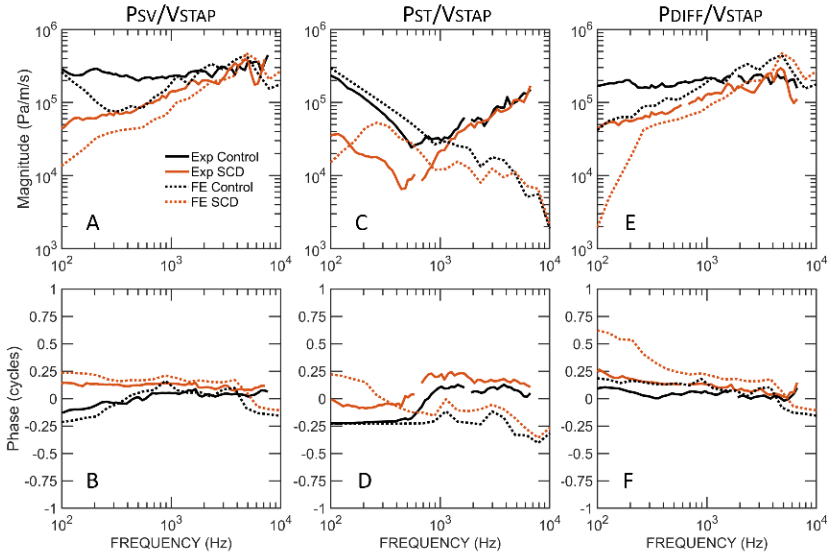


FIGURE 2. The measured and simulated magnitudes and phases of P_{SV}, P_{ST}, and P_{DIFF} in AC. The solid black and red lines represent the experimental results before and after SCD was created, respectively[11]. The dotted black and red lines represent the simulated results under the normal and SCD conditions, respectively.

P_{ST} normalized to the stapes velocity in AC are shown in Fig. 2C and 2D. The measured P_{ST} in normal ears (solid black line) decreases with frequency below 500 Hz and increases with frequency above 500 Hz. Its phase stays at -0.25 cycles below 500 Hz and increases to about +0.2 at high frequencies. In normal ears, below 500 Hz P_{ST} is determined predominantly by the stiffness of the round window; above 500 Hz P_{ST} is determined by the mass of the fluid between the pressure sensor and the round window. SCD decreases P_{ST} below 1000 Hz in the experiments (solid black line vs solid red line). The simulated P_{ST} in the normal ear decreases with frequencies and its phase is close to -0.25 cycles from 100 Hz to 10 kHz (dotted black line). This is because in the simulation P_{ST} was represented by the pressure on the round window membrane, and the membrane's mass is much less

than the fluid's mass in the experiments. In the simulation, SCD decreases P_{ST} below 3 kHz; and a sizeable decrease takes place below 300 Hz (dotted black line vs dotted red line).

P_{DIFF} normalized to the stapes velocity in AC are shown in Fig. 2E and 2F. In the experiments, SCD decreases P_{DIFF} below 1000 Hz and the lower the frequency the greater the reduction (solid black line vs solid red line). In the simulation, SCD decreases P_{DIFF} below 2000 Hz, and the reduction is substantial below 300 Hz (dotted black line vs dotted red line).

Bone-conduction results

The dotted lines in Figs 3-5 show the SCD-induced changes in the magnitude and phase of P_{SV}, P_{ST}, and P_{DIFF} simulated by the FE model for BC excitation in the x, y, and z direction, respectively. The solid lines in those figures are the medians of the changes in P_{SV}, P_{ST}, and P_{DIFF} due to SCD measured from five ears[11]. During the measurements the BC excitation was produced by a Baha (BP 100, Cochlear) mounted on isolated temporal bones, vibrating approximately in parallel with the ear canal[11].

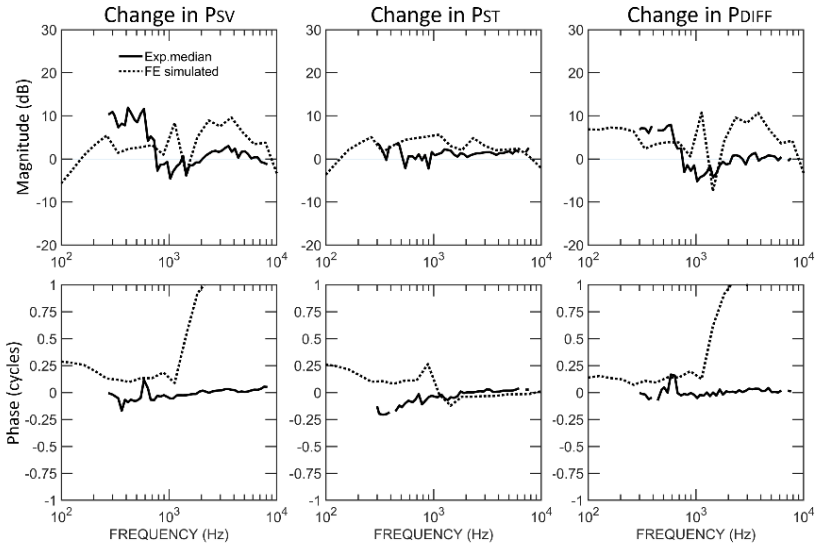


FIGURE 3. The simulated changes in the magnitude and phase of P_{SV}, P_{ST}, and P_{DIFF} due to SCD when the model is under BC excitation in the x direction (dotted lines). The solid lines represent previous measurements[11].

The simulated results show that SCD effects on P_{SV}, P_{ST}, and P_{DIFF} in BC depends on the direction of BC stimulation. When the model undergoes BC stimulation in the x direction (Fig. 3), SCD increases P_{SV}, P_{ST}, and P_{DIFF} over almost the entire frequency range. The highest increases in P_{SV} and P_{DIFF} are 8-10 dB, occurring at 2-4 kHz. The increase in P_{ST} is less than 6 dB.

When the model undergoes BC excitation in the y direction (Fig. 4), SCD decreases P_{SV} at 100-250 Hz, increases P_{ST} around 300 Hz but decreases it near 800 Hz. As a result, P_{DIFF} is increased below 400 Hz by as much as 17 dB.

When the model is excited in the z direction (Fig. 5), SCD increases P_{SV}, P_{ST}, and P_{DIFF} over all the frequencies except around 300 Hz. The increase in P_{SV} and P_{DIFF} below 300 Hz is greater than that at other frequencies. The largest increase in P_{ST} takes place at 100 Hz and 800 Hz.

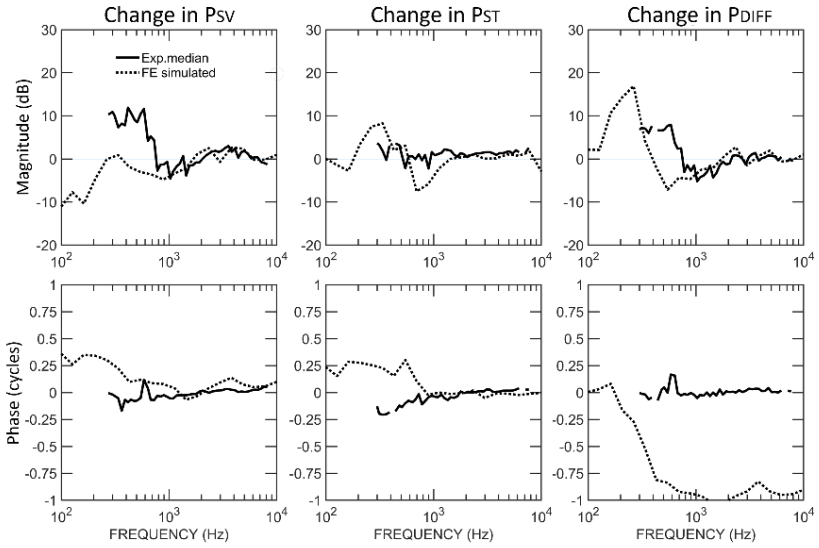


FIGURE 4. The simulated changes in the magnitude and phase of P_{SV}, P_{ST}, and P_{DIFF} due to SCD when the model is under BC excitation in the y direction (dotted lines). The solid lines represent previous measurements[11].

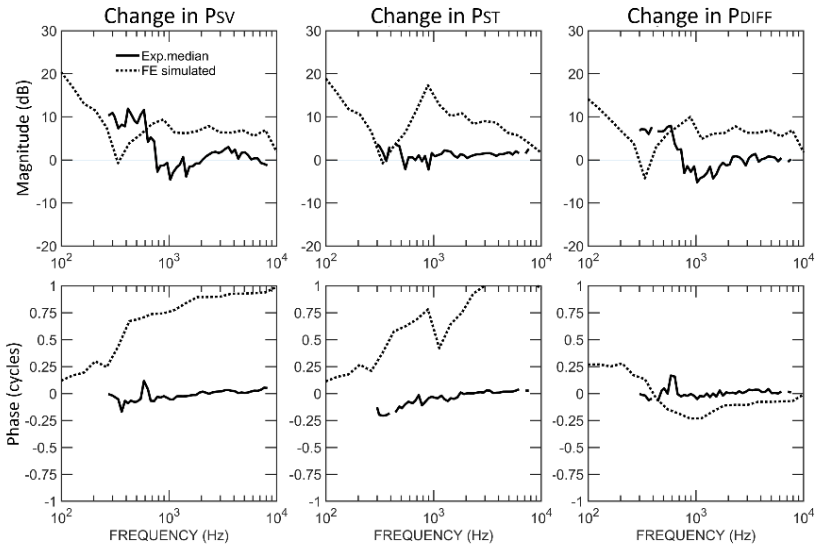


FIGURE 5. The simulated changes in the magnitude and phase of P_{SV}, P_{ST}, and P_{DIFF} due to SCD when the model is under BC excitation in the z direction (dotted lines). The solid lines represent previous measurements[11].

Although the model predicts that SCD increases P_{DIFF} – the cochlear input drive – in all three BC excitation directions, the simulated changes in P_{SV}, P_{ST}, P_{DIFF}, for any of the three excitation directions, do not simultaneously match the experimental data.

DISCUSSION

Since the first case of SCD reported by Minor in 1998, there have been a myriad of research focusing on the physiological cause, the symptoms, and treatments of the disease. The mechanism of SCD-induced AC hearing loss (i.e. the air-bone gap) has been studied using lumped-element models and FE models. It is accepted that SCD acts as a third window and thus provides a low-impedance shunting pathway for AC sounds, decreasing the sound pressures in SV and ST at low frequencies in AC. Our present FE model shows similar results – SCD decreases P_{SV}, P_{ST}, and P_{DIFF} at low frequencies in AC. But the comparison between the simulation and measurements in AC also suggests that the overall damping of the cochlea in the model is insufficient. One explanation is that the scala tympani was merged with SV, increasing the cross-section area of the SV and thus reducing the damping of the channel. In addition, the organ of Corti, which could contribute to the damping in the cochlea, is simplified as the BM. The lack of the scala tympani and other structures in the organ of Corti may be the reason that the damping of the cochlea is inadequate in the present model.

While AC hearing loss due to SCD has been well studied, the SCD-induced BC hyperacusis remains not well understood. Clinical studies show that the thresholds of BC decrease in patients with SCD[8]. The simulated results in the present study suggest that SCD increases the cochlear input drive (P_{DIFF}) in BC for all three excitation directions, supporting the audiometric findings and the BC hyperacusis symptoms. But the frequency dependence of SCD-induced change in P_{SV}, P_{ST}, and P_{DIFF} predicted by the model is inconsistent with the previous temporal bone study[11]. Since the SCD effect on BC might depend on the excitation direction, we hypothesize that the BC stimulation direction in the previous experiments may not align with any of the three excitation directions in the present model, contributing to the discrepancies. Furthermore, the low damping of the cochlea in the model may also contribute to difference between the simulation and the measurements.

CONCLUSION

We developed a FE human ear model to understand the SCD effect on BC hearing. While this model is able to simulate SCD-induced hearing loss for AC and hearing enhancement for BC (indicated by the changes in P_{DIFF}), there are discrepancies between the simulated intracochlear pressures (P_{SV} and P_{ST}) and their experimental counterparts, possibly contributed by the lack of the membranous labyrinth in the model and the difference in the BC excitation direction between the simulation and the experiments. In future, we will add the membranous labyrinth to the model and test other BC stimulation directions.

ACKNOWLEDGEMENTS

The present study is supported by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health under Award Number NIH/NIDCD R21DC017251.

REFERENCES

1. L. B. Minor, Am J Otol 21 (1), 9-19 (2000).
2. L. B. Minor, D. Solomon, J. S. Zinreich and D. S. Zee, Arch Otolaryngol Head Neck Surg 124 (3), 249-258 (1998).
3. S. N. Merchant and J. J. Rosowski, Otol Neurotol 29 (3), 282-289 (2008).
4. S. N. Merchant, J. J. Rosowski and M. J. McKenna, Adv Otorhinolaryngol 65, 137-145 (2007).
5. A. A. Mikulec, M. J. McKenna, M. J. Ramsey, J. J. Rosowski, B. S. Herrmann, S. D. Rauch, H. D. Curtin and S. N. Merchant, Otol Neurotol 25 (2), 121-129 (2004).
6. S. A. Mueller, D. Vibert, R. Haeusler, A. Raabe and M. Caversaccio, Eur Arch Otorhinolaryngol 271 (6), 1369-1374 (2014).
7. A. Banerjee, A. Whyte and M. D. Atlas, Clin Otolaryngol 30 (1), 9-15 (2005).
8. K. Brantberg, L. Verrecchia and M. Westin, Audiol Neurootol 21 (6), 365-371 (2016).
9. B. K. Ward, Y. Agrawal, E. Nguyen, C. C. Della Santina, C. J. Limb, H. W. Francis, L. B. Minor and J. P. Carey, Otol Neurotol 33 (8), 1386-1391 (2012).
10. B. K. Ward, J. P. Carey and L. B. Minor, Front Neurol 8, 177 (2017).
11. X. Guan, Y. S. Cheng, D. J. Galaiya, J. J. Rosowski, D. J. Lee and H. H. Nakajima, Sci Rep 10 (1), 16564 (2020).
12. J. J. Rosowski, J. E. Songer, H. H. Nakajima, K. M. Brinsko and S. N. Merchant, Otol Neurotol 25 (3), 323-332 (2004).
13. J. E. Songer and J. J. Rosowski, J Acoust Soc Am 122 (2), 943-951 (2007).
14. N. Kim, C. R. Steele and S. Puria, Hear Res 301, 72-84 (2013).
15. J. Lim, I. Dobrev, C. Roosli, S. Stenfelt and N. Kim, Hear Res, 108337 (2021).
16. S. Stenfelt, Hear Res 340, 135-143 (2016).
17. C. Stieger, X. Guan, R. B. Farahmand, B. F. Page, J. P. Merchant, D. Abur and H. H. Nakajima, J Assoc Res Otolaryngol 19 (5), 523-539 (2018).
18. R. Z. Gan, B. Feng and Q. Sun, Ann Biomed Eng 32 (6), 847-859 (2004).
19. K. Homma, Y. Shimizu, N. Kim, Y. Du and S. Puria, Hear Res 263 (1-2), 204-215 (2010).
20. K. N. O'Connor, H. Cai and S. Puria, J Acoust Soc Am 142 (5), 2836 (2017).
21. T. Koike, H. Wada and T. Kobayashi, J Acoust Soc Am 111 (3), 1306-1317 (2002).
22. N. Kim, C. R. Steele and S. Puria, Biophys J 107 (1), 233-241 (2014).
23. J. Liang, D. Nakmali, R. Z. Gan, H. Lu and C. Dai, Otol Neurotol 42 (2), 319-326 (2021).
24. E. Levin, S. Muravchick and M. I. Gold, Anesth Analg 60 (11), 814-817 (1981).
25. J. J. Rosowski, W. Chien, M. E. Ravicz and S. N. Merchant, Audiol Neuro-Otol 12 (4), 265-276 (2007).