FUNCTIONALITY OF COCHLEAR MICROMECHANICS – AS ELUCIDATED BY UPWARD SPREAD OF MASKING AND TWO TONE SUPPRESSION

D. Sen
School of Electrical Engineering,
The University of New South Wales,
Sydney-UNSW, NSW 2052, Australia.

J. B. Allen
Electrical & Computer Engineering,
University of Illinois at Urbana Champaign,
Urbana, IL 61801, USA.

The effect of increasing the level of a low frequency masker on higher frequency probes can be observed in three paradigms: psychophysically (termed Upward Spread of Masking or USM), mechanically on the basilar membrane (Basilar Membrane Two Tone Suppression or BM-2TS) as well as neuro-physiologically at the auditory neurons (Neural Two Tone Suppression or Neural-2TS). This paper reviews various experimental USM, Neural-2TS and BM-2TS data with the aim of shedding light into the underlying physiological mechanisms in the cochlea.

INTRODUCTION

It is generally accepted that the sharp tuning observed at the characteristic frequency (CF) auditory nerve fibre can be attributed to the sharp mechanical response at the corresponding position of the basilar membrane (at the Best Place - BP). This observation has resulted in attention being focused on the basilar membrane and away from other micromechanical structures in the cochlea. However, at regions away from the BP, there is considerable evidence to suggest that the basilar membrane and auditory nerve response diverge. This is especially the case at regions basal to the BP (termed the "tail region" of the spatial response – see Fig 5b for an example of the "tail region") and introduces possibility of other physiological structures, in addition to the basilar membrane, contributing to the auditory nerve response. While most recent basilar membrane mechanical data show an excellent match to Neural responses at the CF [14], there is a clear and often marked difference in the two responses at the tail. This discrepancy seems especially pronounced in cats, where the slope of neural response in the tail region is about 1 to 4 dB/octave whereas the slope of the basilar membrane response is about 9 dB/octave [15].

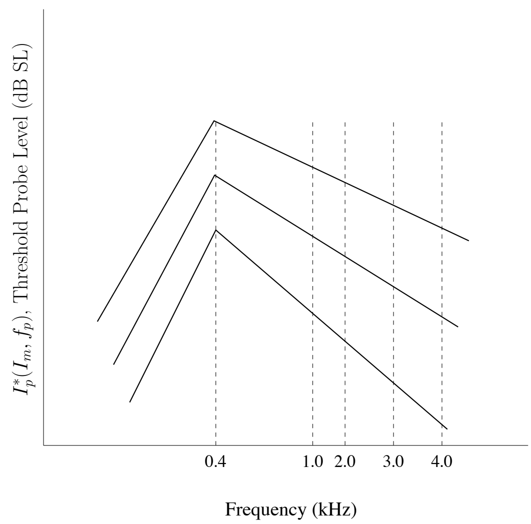
This paper reviews various experimental data that suggest that the Upward Spread of Masking (USM) effect is simply a psychophysical observation of what is also observed at the auditory nerve as Neural-2TS. This is a satisfying result as it indicates that there is little difference between what is perceived and the auditory nerve response. Similarly, if the Basilar Membrane Two Tone Suppression (BM-2TS) data were to match Neural-2TS just as exactly, this would indicate that the basilar membrane response is identical to the auditory nerve response and, by the same token, to what is perceived (USM). However, our review shows that there is discrepancy in the BM-2TS and Neural-2TS data which suggests influence of structures such as the tectorial membrane (TM), inner hair cells (IHC) and outer hair cells (OHC), collectively called micromechanical structures, that act to modify the basilar membrane response to what is observed at the auditory nerve.

The stimuli used in USM, Neural-2TS and BM-2TS is composed of two simultaneous tones of differing frequencies. While in the psychoacoustic USM literature the two tones are called "probe" and "masker", in the 2TS literature, the same stimuli are termed "CF tone" and "suppressor". In this paper, we will concentrate only on the case where the probe/CF tone is at least an octave above the masker/suppressor, placing it at the tail region of the masker (see Fig 5b) and thereby facilitating the study of characteristics at that region.

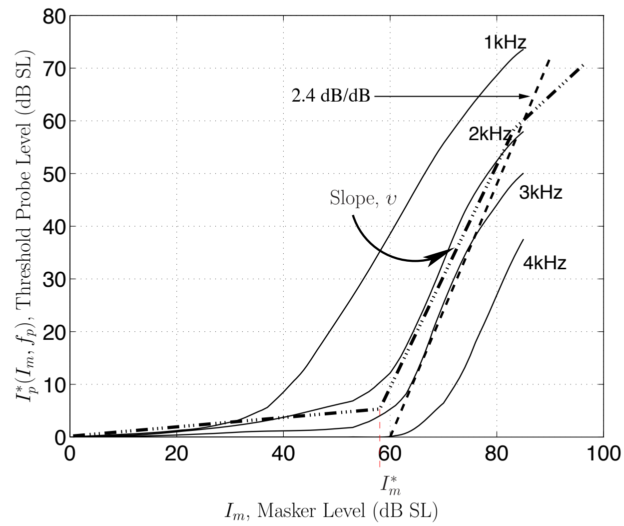
This paper will review each of the three phenomena, emphasizing their similarities, dissimilarities and the relevance of the experimental data in terms of cochlear neuro-physiology. The conclusion resulting from this phenomenological review is that micromechanical structures play a significant role and are indeed essential in explaining transduction from mechanical to auditory nerve response. Finally, we describe a simple model of the micromechanical structures which is able to account for the discrepancy between the auditory nerve and basilar membrane response at the "tail region".

UPWARD SPREAD OF MASKING (USM)

The effect of simultaneous masking is depicted in Fig 1a, where a 400 Hz masker "tone" is presented at a certain level and the corresponding level of a second "probe" at its threshold of perception is recorded. The triangular masking functions (as a function of frequency) are used to hide quantization noise below the threshold of perceptibility in "MP3" compression systems. An important characteristic of simultaneous masking is that as the level of the masker is increased, the corresponding



(A)



(B)

Figure 1. Upward Spread Masking. In (A), a cartoon depiction showing the effect of a 400 Hz tonal masker. For probe tones that have frequencies greater than the masker frequency (e.g. The 1, 2 and 3 kHz probes), the level increment required to keep the probe at its threshold of perception is greater than a proportional increase in the level of the masker. In (B) the same effect is quantified in experimental data [2] by plotting the threshold probe level versus masker level. The figure clearly shows that the probes need to be increased at a higher than proportional rate of the masker. The probes are at 1, 2, 3 and 4 kHz. The dashed line shows the maximum slope of 3 kHz probe to be about 2.4 dB/dB, meaning that every 1 dB increase in masker level will require the probe to be increased by 2.4 dB to keep it at the threshold of perception. For discussion purposes, we have modelled the effect using the dash-dot line. We will say the effect begins at masker level, I_m^*, and the strength of the effect is characterised by the slope, υ.

threshold level of a higher frequency probe tone needs to be increased more than the proportional increase in the masker level. This is clearly seen in Fig 1a for probes of 1, 2 and 3 kHz where the triangular masking curves seem to be spreading upwards – hence the term "upward spread of masking". The effect is also negligible when the probe is at a lower frequency than the masker. Fletcher [1] and Wegel & Lane [2] quantified the effect in the early 1920s. One such experimental result is shown in Fig 1b where the level of a 400 Hz masker is shown on the x-axis and the corresponding level of the probe at threshold is shown on the y-axis. The experimental results which have been verified in recent years [3] show that as the level of the masker is increased beyond 60-70 dB SPL, the higher frequency probes need to be increased by as much as 2.4 dB for every 1 dB increase in the masker (see Fig. 1B). Interestingly, this non-linear effect is difficult to make use of in typical music compression systems as the encoder typically has no indication of the volume set by the listener.

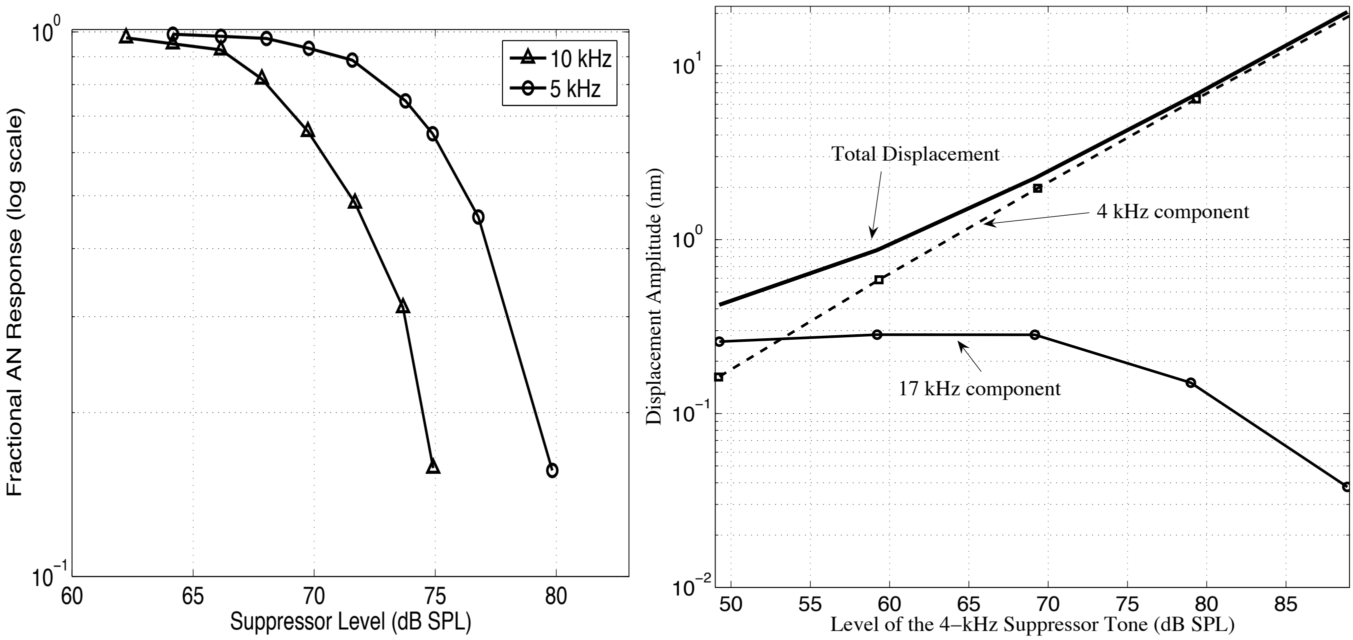
As indicated in Fig 1B, we will use two parameters to facilitate discussion of the USM effect: υ and I_m^*. The first parameter, υ, is the slope of the masking as a function of masker level, also known as the Growth of Masking (GOM). The parameter, υ, thus describes the "strength" of the masking. The * on the second parameter, I_m, indicates that the masker intensity is at the threshold level or the lowest level of the masker which causes masking of the probe. As such, I_m^* is a function of the masker frequency, f_m, and probe frequency, f_p. When the probe tone is thus at least an octave higher than the masker tone, (i.e $f_p > 2f_m$), slope υ is at its highest at about 2 to 3 dB/dB [3,4]. Similarly, I_m^*, can be estimated to be between 55-65 dB SPL in those same studies.

NEURAL TWO TONE SUPRESSION

Neural-2TS [5, 6] is the effect where the neural discharge rate of a CF auditory nerve fibre (hereafter termed probe neuron) is observed to reduce during the presence of a suppressing tone of different frequency (see Fig 2A). The CF tone is typically introduced at 6 to 10 dB above threshold. The ability of the suppressor tone to reduce the discharge rate of the probe neuron depends mainly on the suppressing tone's intensity and on its frequency relative to the CF. The curves in Fig 2A are clearly similar to the curves in Fig 1B once we account for the fact that, in USM, we need to increase the probe to account for its underlying suppression, while in Neural-2TS we are not attempting to raise the CF tone at all but just witnessing its suppression.

An interesting characteristic of Neural-2TS is shown in Fig 3A where the dotted line shows a normal threshold tuning curve. The three straight line curves represent the presence of a CF tone (at three different levels) and the corresponding level required by secondary tones (across the frequency range) to increase the discharge rate. The three open circles represent levels required by a suppressor tone to reduce the discharge rate back to threshold discharge rates when the CF tone is present at the three different levels. The three circles are clearly at or below the threshold levels required to excite the CF neuron – meaning that even though the three tones are individually unable to excite the CF-neuron, they are able to suppress the discharge rate of the CF/Probe tone. This means the suppressor is non-excitatory but suppressive at the CF neuron.

Suppression is most effective when the suppressor frequency is lower than CF, and we shall thus concentrate on



(A) (B)

Figure 2: 2TS characteristics as a function of suppressor level. (A) Neural-2TS data from [5] showing average discharge rate from a single auditory nerve fibre with a characteristic frequency (CF) of 17.8 kHz as a function of suppressor level. The data is normalized by the discharge rate of the neuron when no suppressor is present. (B) BM-2TS data from Geisler [8]. Here the probe is at 17 kHz while the suppressor is at 4 kHz. The total response shown by the straight line actually increases with increasing suppressor level. However, when the 17 kHz component is extracted, it clearly reduces as a function of suppressor level.

the case when the suppressor frequency is at least an octave lower than the probe CF. This constraint is in essence identical to the constraint we placed on our USM of $f_p > 2f_m$, if we identify the CF tone frequency with f_p and the suppressor frequency with f_m.

Using the above analogy, we can compare I_m^* of the USM effect with suppressor threshold or the minimum level of the suppressor required to reduce the discharge rate of the CF neuron. This has been studied extensively by Fahey & Allen [7] and shown to be 65 dB SPL (± 5 dB) above 1 kHz. This is strikingly similar to the USM threshold, I_m^*, discussed in the previous section. Further, the Rate of Suppression (RoS), which is the slope of the auditory nerve discharge rate as a function of suppressor level, has been studied extensively [3,6] and found to be a maximum of 2.4 dB/dB. The RoS can be directly compared with USM's GOM or υ.

These similarities inevitably lead to the conclusion that USM and Neural-2TS are closely related and are possibly the same phenomena just observed psychophysically (in the case of USM) or in neural discharge rate measurements (in the case of Neural-2TS). This conclusion is given further credence when one also observes that, for both phenomena, the maximum effect occurs when the suppressor/masker frequency is below that of the probe/CF tone frequency.

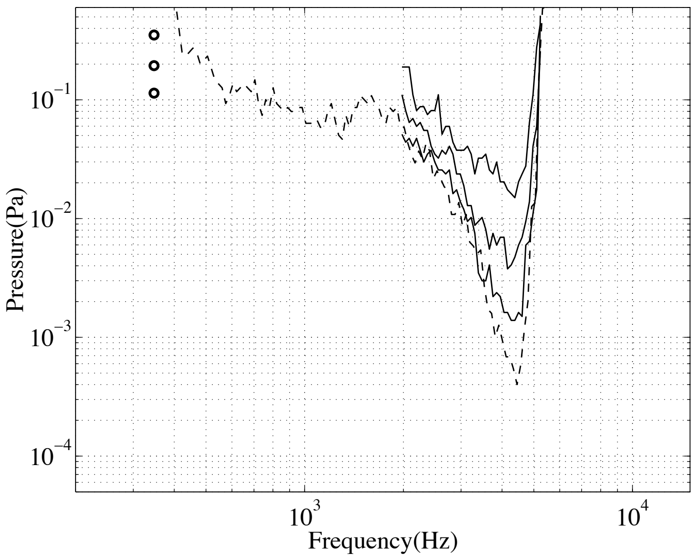
BASILAR MEMBRANE TWO TONE SUPPRESSION (BM-2TS)

2TS is observed mechanically on the basilar membrane (BM-2TS). This is clearly seen in Fig 2B where the stimulus consists of a 17 kHz probe/CF tone along with a 4 kHz masker/suppressor tone. While the total RMS motion of the basilar membrane at the BP of the 17 kHz increases with increasing suppressor level, the 17 kHz component of the motion is seen to reduce. Quantitatively, this is dramatically different from Neural-2TS, where the total discharge rate is typically reduced when a suppressor tone is present simultaneously.

Most recent studies [8, 9] of BM-2TS have found the Rate of Suppression (RoS) to be about 1 dB/dB. In a study by Ruggero [10], the maximum RoS was found to be approximately 1.42 dB/dB (measured using iso-velocity analysis). There is thus a very large discrepancy in RoS between the BM-2TS and Neural-2TS (1 vs. 2.4 dB/dB). The difference between a 1 and 2.4 dB/dB suppression amounts to 10 dB and 24 dB of suppression for a 10 dB change in suppressor level (more than a factor of 4 deviation for every factor of 3 change in level).

The masker threshold, I_m^*, in USM plays a similar role to the suppressor threshold (lowest suppressor level that causes suppression at the CF or BP). While both USM and Neural-2TS is characterised by I_m^* of about 55-65 dB SPL, in BM-2TS, the suppressor must be more than 80 dB SPL before it suppresses the probe tone [8,9]. It is interesting to note that both the USM and the Neural-2TS effects are almost over when the masker/suppressor level is about 80 dB SPL (see Fig 1B).

Figure 2B shows the results from a BM-2TS study [9] where the frequency of the probe tone is 26 kHz and the suppressors (represented by open symbols) are at about 3.5, 7.5 and 20 kHz. Again, just as in the Neural-2TS study (Figure 2A), the probe tone is placed at 3 different levels (filled symbols) and the open circles represent the levels of the suppressors required to suppress the basilar membrane motion. We clearly see from this study that the suppressor levels are almost 18 dB on average above the threshold levels required



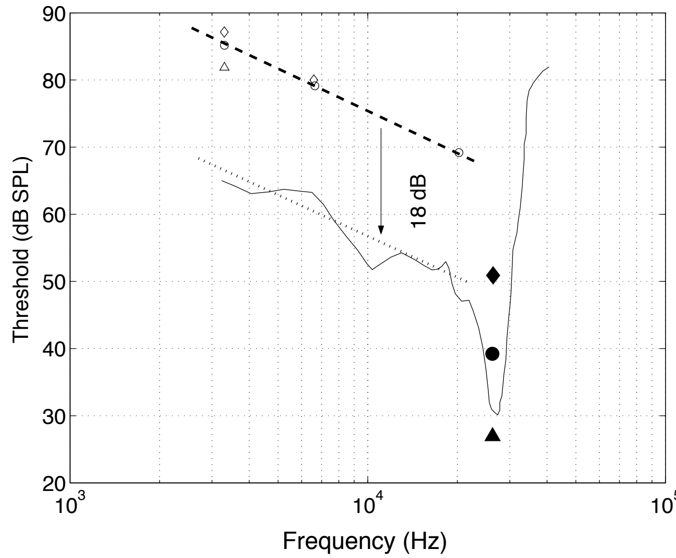


Figure 3. Two Tone Suppression characteristics as a function of frequency. (A) Neural 2TS data from [7]. The open circles show the level and frequency of the suppressor and the corresponding reduction in respond at the CF. The dotted line is for the case when there is no suppressor. Note that even though the suppressor tones excite the CF neurons (they are below the threshold tuning curve), they are able to suppress the CF tone. (B) Basilar membrane 2TS data from Cooper [9] (Figure 3D). The open symbols show the level and frequency of the suppressors required to suppress the tone near BF shown by filled symbols. The solid line show the levels of various tones required to evoke a constant amplitude (0.1 nm) vibration of the basilar membrane at the BP of the probe tone. Note that the low side suppressors are far above (approximately 18 dB) the levels required to produce threshold vibrations at the BP. The suppressor levels are also significantly higher than the minimum suppressor levels (around 65 dB SPL) required for Neural-2TS.

to produce the tuning curve given by the straight line. This is quite a sharp contrast from the suppressor levels in Neural-2TS which, as we discussed in the previous section, were actually below the tuning curve.

The discrepancies between BM-2TS and USM/Neural-2TS of (i) differences in suppressor threshold levels, (ii) differences in rate of suppression, and (iii) the fact that in neural-2TS unlike BM-2TS, the neural response is typically lower than the probe alone provides sufficient evidence to suggest that the two phenomena are quite different. Each of these pose a serious problem for theories which suggest that neural response is directly related to basilar membrane mechanics. The next section reviews existing models of 2TS as well as providing a plausible and elegant conceptual model that does not contradict modern views of cochlear mechanics and is yet able to resolve the discussed discrepancies.

MODELLING AND DISCUSSION

The prevailing explanation [8,13] for BM-2TS is that the nonlinear OHC response is saturated by the high level suppressor. Another explanation for BM-2TS is that it is an epiphenomenon of the half octave shift of the basilar membrane response (or "migration") towards the base as a function of intensity [11]. It can be imagined that, as the response moves towards the base with increasing level, the response at the BP decreases, giving the illusion of suppression. While both of these explanations model BM-2TS quite well, they are unable to explain the Neural-2TS and BM-2TS discrepancies discussed in the previous sections. However, if we accept that BM-2TS and Neural-2TS are inherently different, then either of the above explanations suffices and all that is required is: (i) an explanation for Neural-2TS and (ii) why an increasing basilar membrane response at the BP (when both the suppressor and CF tone is present – see Fig. 2B) is not reflected by an increasing discharge rate at the CF auditory neuron (which actually decreases).

Towards a solution to the second problem (ii) above, we suggest that the basilar membrane response is modified in its transduction path to the auditory nerve fibres [9]. The transduction path depicted in Fig 4, clearly shows that there is ample scope for the micromechanics (TM, Cilia, OHCs, IHCs) to modify the basilar membrane response. If the modification is of a high-pass nature, it will act to attenuate the basilar membrane response only in the area basal to the CF providing the required solution to the problem while not contradicting the prevailing view that basilar membrane and auditory nerve tuning at the CF are identical.

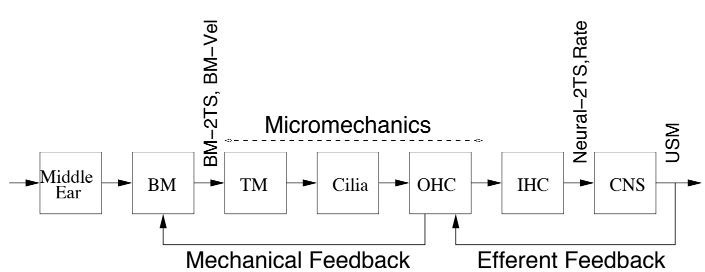


Figure 4. A block diagram depicting the transduction path in auditory physiology. This also is a block diagram for the computational model in the paper.

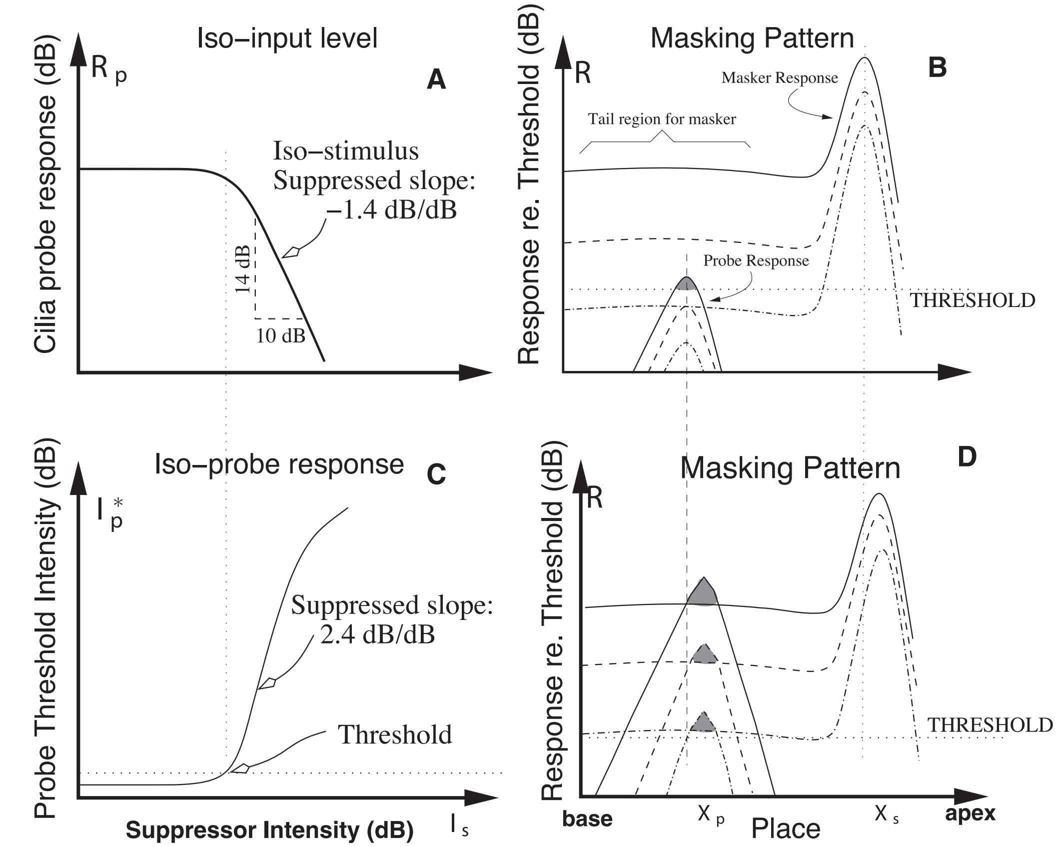


Figure 5. A cartoon explaining the physiological mechanism for USM and 2TS.

Another related observation in Neural-2TS, that requires explanation, is that low level suppressors while non-excitatory at the CF, are able to suppress the CF probe. This can be explained if we assume that OHCs are slightly more sensitive than IHCs. In this case, the OHCs will initiate the suppression even though the IHC neurons will not respond at the threshold of suppression, as observed.

Finally, we need an explanation of Neural-2TS (or USM). This is done phenomenologically in Fig 5. In Fig. 5A, we assume that the masker's tail response is able to suppress the probe at a rate of up to 1.4 dB/dB (Fig. 5A). The tail response is known to increase linearly with level, as shown in Figure 5B. In the USM experimental paradigm, the probe is required to overcome the effect of its suppression as well as the linear growth of the tail. This is shown in Fig 5D. In order for it to overcome both these effects, it has to be increased at a rate of (as much as) 1.4 + 1=2.4 dB/dB as shown. This will then agree with the experimental observation of a maximum 2.4 dB/dB GOM in USM. In USM, asking the subject to increase the probe tone such that it is just perceivable is the psychophysical equivalent of iso-discharge rate of the nerve fibres at the CF of the probe. Neural-2TS observations, which use iso-rate measurements to calculate the growth of suppression [6], also display this maximum rate of growth of about 2.4 dB/dB. The conclusion from this phenomenological model is that the underlying physiological suppression producing both Neural-2TS and USM must be at a rate of 1.4 dB/dB (or 1 dB/dB lower than the iso-response observation).

To test the above hypothesis, we have incorporated a simple micromechanical model into a hydromechanical macro-mechanical model of the cochlea. The TM is modelled

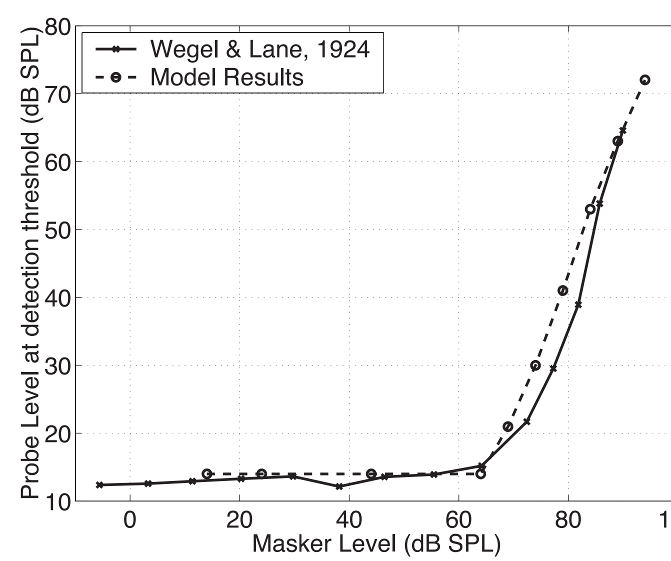


Figure 6. USM prediction from a computational model. The output from the model is compared to data from Wegel & Lane [2].

as a transmission line, terminated by the cilia (m_c, k_c, and r_c modeling the mass, stiffness and damping respectively). The micromechanical model acts to attenuate the response basal to the CP on the basilar membrane tail response. In the model, the flat tail of the low frequency masker (which grows linearly with masker level) changes the stiffness of the basilar membrane at the place of the probe/CF tone, which in turn suppresses the probe actuating the Neural-2TS/USM effect. The serial consecutive depiction of the model (see Fig. 4) is only broken by the feedback due to the OHC motility, which is able to change the impedance of the cochlear partition and explains various non-linear phenomena (including Neural-2TS and others such as the half-octave shift of the basilar membrane and OAE effects). The high-pass micromechanical filter is also able to convert the 9dB/octave slope of the basilar membrane tail response to the almost place invariant (flat) neural tail, accounting for the discrepancy observed in [14]. Figure 6 shows results from the computational model as described here. The model is clearly able to predict the effect of USM (and therefore Neural-2TS) while not contradicting any existing theoretical and experimental observations.

REFERENCES

1. Fletcher, H., "Physical measurements of audition and their bearing on the theory of hearing", Journal of the Franklin Institute. 196 (3), 289–326, (1923).
2. Wegel, R and Lane, C., "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, (1924).
3. Bacon, S., Bodern, L., Lee, J. and Repovsch, J., "Growth of simultaneous masking for $f_m < f_s$: Effects of overall frequency and level", J. Acoust. Society Amer., 106 (1), 341–350, (1999).
4. Zwicker, E., and Fastl, H., "Psychoacoustics: Facts and Models", Springer, Berlin, 1990.
5. Abbas, P., Sachs, M., "Two–tone suppression in auditory–nerve Fibers: Extension of a stimulus–response relationship", J. Acoust. Soc. Amer. 59 (1), 112-122, (1976).
6. Delgutte, B., "Two-tone suppression in auditory-nerve Fibres: Dependence on suppressor frequency and level", Hearing Research 49, 225-246, (1990).
7. Fahey, P.F. and Allen, J. B., "Nonlinear phenomena as observed in the ear canal and at the auditory nerve", J. Acoust. Soc. Amer. 77 (2), 599–612, (1985).
8. Geisler, C. and Nuttall, A., "Two-tone suppression of basilar membrane vibrations in the base of the guinea pig cochlea using "low-side" suppressors", J. Acoust. Soc. Amer. 102 (1), 430–440, (1997).
9. Cooper, N., "Two-tone suppression in cochlear mechanics", J. Acoust. Soc. Amer. 99 (5), 3087–3098, (1996).
10. Ruggero, M., Robles, L. and Rich, N., "Two-tone suppression in the basilar memnbrane of the cochlea: Mechanical basis of auditory-nerve rate suppression", J. of Neurophys. 68 (4), 1087–1099, (1992).
11. McFadden, D., "The curious half–octave shift: Evidence of a basalward migration of the traveling–wave envelope with increasing intensity", in: R. Salvi, D. Henderson, R. Hamernik, V. Coletti (Eds.), Applied and Basic Aspects of Noise-Induced Hearing Loss, Plenum, (1986).
12. Geisler, C.D., Yates, G.K., Patuzzi, R.B. and Johnstone, B.M., "Saturation of outer hair cell receptor currents causes two tone suppression", Hearing Research 44, 241–256, (1990).
13. Patuzzi, R.B., Yates, G.K. and Johnstone, B.M., "Outer hair cell receptor current and sensorineural hearing loss", Hearing Research 42, 47–72, (1989).
14. Ruggero, M., Narayan, S., Temchin, A. and Recio, A., "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. Nat. Acad. of Sciences 97 (2), 11744–11750, (2000).
15. Sen, D. and Dedhia, B., "Slope of the Human Auditory Nerve Response at the Tail", Midwinter Meeting of the Association for Research in Otolaryngology (ARO), (2003).
16. Allen, J.B and Sen, D., "An unified theory of two tone suppression and the upward spread of masking", in Proceedings of the 16th International Congress on Acoustics (ICA) and 135th meeting of the Acoustical Society of America, (1998)