Individual Variability of Hearing-Impaired Consonant Perception

Andrea Trevino, MSEE,[1] and Jont B. Allen, Ph.D.[1]

ABSTRACT

The study of hearing impaired speech recognition at the consonant level allows for a detailed examination of individual variability. Use of low-context stimuli, such as consonants, aids in minimizing the influence of some variable cognitive abilities (e.g., use of context, memory) across listeners and focuses on differences in the processing or interpretation of the existing acoustic consonant cues. We show that hearing-impaired perception can vary across multiple tokens of the same consonant, in both noise robustness and confusion groups. Within-consonant differences in noise robustness are related to differences in intensity of the consonant cue region. For a single listener, high errors can exist for a small subset of test stimuli, and performance for the majority of test stimuli can remain at ceiling. The existence of within-consonant differences in confusion groups entails that an average over multiple tokens of the same consonant results in a larger confusion group than for a single consonant token. For each consonant token, the same confusion group is consistently observed across a population of hearing impaired listeners. Quantifying perceptual differences provides insight into the perception of speech under noisy conditions and characterizes each listener's hearing impairment.

KEYWORDS: Consonant, individual differences, speech perception, confusion group

Learning Outcomes: As a result of this activity, the participant will be able to (1) describe that multiple tokens of the same consonant can be perceptually different for hearing-impaired listeners, and (2) characterize the naturally variable consonant cues that can help to explain individual differences in consonant perception.

[1]The University of Illinois Champaign Urbana, Urbana, Illinois.

Address for correspondence: Andrea Trevino, MSEE, the University of Illinois Champaign Urbana, 315 N. Orchard St. Apt. 2, Urbana, IL 61801 (e-mail: atrevin2@illinois.edu).

Individual Variability in Aided Outcomes; Guest Editor, Jason A. Galster, Ph.D.

Semin Hear 2013;34:74–85.
DOI: http://dx.doi.org/10.1055/s-0033-1341345.
ISSN 0734-0451.

Understanding the underlying individual differences in hearing-impaired (HI) consonant perception is key to the development and interpretation of speech tests. Multiple tokens of the same consonant are often considered multiple measures of the same effect. Contrary to this approach, the consonant cue literature has documented, in detail, the variability of the cues that are present in naturally produced speech.[1–7] The variability of consonant cues is quantified by analysis of the acoustical properties of each token and can be observed across speech samples that are both unambiguous and robust to noise for normal-hearing (NH) listeners. The question remains: does the natural variability across tokens of the same consonant lead to differences in HI perception?

Our research at the University of Illinois Champaign-Urbana has shown that HI perceptual differences exist across multiple tokens of a single consonant (i.e., *within-consonant*). Furthermore, HI perceptual within-consonant differences can be observed in terms of robustness to noise and confusion groups. Within-consonant differences in noise robustness can result in the observation that, for a single HI listener at a fixed signal-to-noise ratio (SNR), performance for one token of a consonant is at ceiling and another token of the same consonant reaches 100% error. For NH listeners, different tokens of a consonant can have varying noise robustness.[8,9] We show that the within-consonant differences in noise robustness for HI listeners are correlated to within-consonant differences in noise robustness for NH listeners.

Each token of a consonant has a unique subgroup of possible confusions. Although there is often overlap in the confusion groups across multiple tokens of the same consonant, some tokens can show systematic within-consonant differences. The averaged responses of HI ears to multiple tokens of a single consonant can often appear to be random guesses drawn from a large confusion group. Some of this randomness can be an artifact of averaging across tokens; smaller confusion groups are observed when slight to moderate HI subjects are examined at the token level. In addition, the token-specific confusion groups are shared across different HI listeners, implying that the subtle variability in the acoustical properties (that does not affect NH recognition) is the source of these systematic within-consonant differences in confusion groups.

Existing practice uses pure tone thresholds as the primary prescriptive measure for the HI ear. It has been widely observed that HI ears can have similar pure tone thresholds but differ in their speech perception abilities.[10–17] Differences in speech perception abilities are most commonly examined using the speech recognition threshold; we examine the variability of speech perception abilities across individual consonant tokens. We show that, in both quiet and low-noise conditions, errors can be concentrated on a small, ear-dependent subset of the test stimuli. An ear that seems "almost normal" in terms of an average error measure can, in actuality, have high error with a very small, specific group of sounds. In addition, ears with similar degrees and configurations of hearing loss and similar average consonant errors can have different individual consonants that fall into error.

When testing HI ears, the selection of the tokens is critically important. Multiple tokens of a single consonant, having acoustic cues that vary naturally in terms of intensity, frequency, and/or temporal cues, can result in different measures of hearing impairment. This natural variability could be used to advantage, but only once it has been controlled.

BACKGROUND

Consonants comprise $\sim 58.5\%$ of conversational speech.[18] Although the relative importance of consonants and vowels for HI speech perception remains uncertain,[19,20] here, we concentrate on HI consonant perception. Many previous works have examined HI consonant recognition using naturally produced speech, including Lawrence and Byers,[21] Bilger and Wang,[22] Owens,[23] Wang et al,[24] Dubno and Dirks,[25] Boothroyd,[26] Fabry and Van Tasell,[27] Dreschler,[28] Gordon-Salant,[29] and Zurek and Delhorne.[30] Overall, the effects of hearing impairment on speech perception are more severe in the presence of noise.[25,28] It has been observed that some consonant confusions are common across a variety of audiometric configurations.[23,29] In these previous studies, data analysis is performed using either an average measure (over all consonants) or with consonants grouped by distinctive features.

The work of Boothroyd and Nittrouer formulated the relationship between correct perception of low-context speech segments (e.g., phonemes) and high-context segments (e.g., words) in NH ears.[31] Follow-up studies by Bronkhorst et al greatly extended this work.[32,33] These studies demonstrate the NH individual's ability to decode high-context speech depends critically on their low-context error. When an HI listener reports that they can "hear speech but have trouble understanding it," it may be due to a small group of specific sounds being incorrectly recognized. These observations affirm the utility of studies of hearing impairment that use low-context speech.

CHARACTERIZING INDIVIDUAL TOKENS

One of the greatest impediments to both developing speech-based tests and interpreting HI speech perception data are the large amount of individual variability that exists in natural speech. Characterizing how the cues vary for tokens used in a speech test allows one to control for some of this variability.

The 3-Dimensional Deep Search (3DDS) methodology, first introduced by Feipeng Li in 2010, combines three classic NH listener psychoacoustic experiments (time truncation, high-/low-pass filtering, and noise masking) to characterize each individual token's necessary consonant cue region in terms of temporal qualities, frequency, and noise robustness.[6,7] The noise masking experiment characterizes the noise robustness of an individual token with the SNR_{90} measure, defined as the full-bandwidth SNR at which the probability of correct recognition drops below 90%. For NH listeners, the SNR_{90} correlates with the physical intensity of the specific 3DDS-isolated consonant cue region, with tokens with more intense cue regions having lower SNR_{90} values.[6,7,34]

The difference in the NH SNR_{90} values between two tokens, the NH ΔSNR_{90}, relates their robustness to noise and, therefore, can be used to relate their relative threshold audibility. When multiple tokens of the same consonant have perceptual cues that are within a similar frequency range, a measure of their relative audibility remains constant for any hearing loss configuration. Calculation of a ΔSNR_{90} value is illustrated in Fig. 1 (top), with Δ marking the difference between the SNR_{90} values of two tokens.

The NH ΔSNR_{90} values are reported for 13 example pairs of consonant tokens in Fig. 1 (bottom), with the consonants sorted along the abscissa by monotonically increasing NH ΔSNR_{90} values. This plot shows that for /g/, the token from talker 1 is more robust to noise than the token from talker 2 by 9 dB, and for /ʒ/, the token from talker 2 is more robust to noise by 10 dB. Very small differences in the noise robustness of the two tested tokens (within $\pm$ 3 dB) are observed for eight consonants, /m, t, k, ʃ, z, n, p, s/. The relation of the NH ΔSNR_{90} values to HI perception is discussed in the following section.

WITHIN-CONSONANT PERCEPTUAL DIFFERENCES

Differences in Token Noise Robustness

Noise robustness can vary across tokens of the same consonant for HI listeners. The variability in noise robustness can be observed in the probability of error as a function of SNR; the most extreme examples are when one consonant token reaches errors as high as 100%, while another token *of the same consonant* remains at ceiling correct performance (0% error).

Fig. 2 display[5] the consonant recognition error for two HI ears (Fig. 2 top left) 40L and (Fig. 2 bottom left) 34L, at four different SNRs (quiet and 0-, 6-, and 12-dB SNR in speech-shaped noise). Each subplot corresponds to one of 14 consonants and shows the error for two test tokens (one male and one female talker), and their average error. All test tokens were chosen such that they show no NH listener error ($<$ 3%) at the four tested SNRs (see "Speech Materials" section). From these plots one can see that there are several examples where one of the two tokens reaches high levels of error, and the other remains robust to noise.

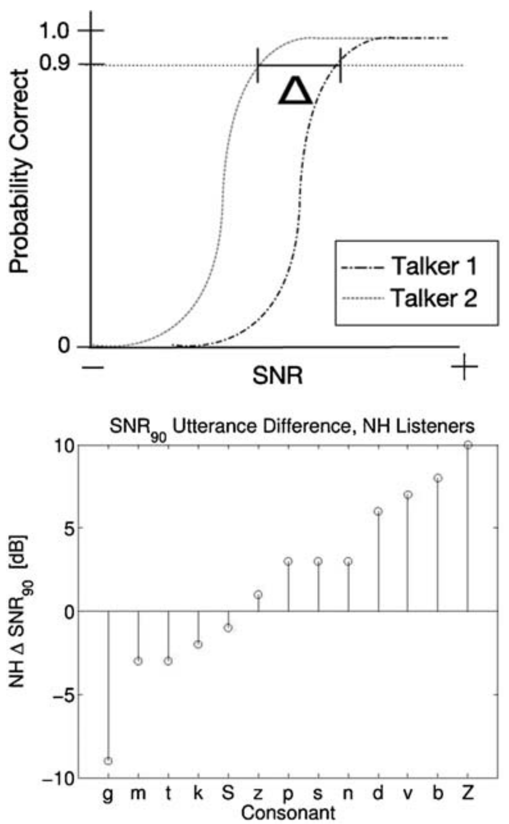

Figure 1 (Top) Illustration of probability versus SNR curves for two consonant tokens, with the difference in SNR_{90} values (ΔSNR_{90}) indicated. The SNR_{90} is defined as the SNR at which the probability of recognition drops below 90%, whereas ΔSNR_{90} quantifies the difference in noise robustness across two tokens. (Bottom) The NH ΔSNR_{90} for 13 example consonant token pairs, as computed from NH perceptual data (Table 2) in the presence of speech-weighted noise.[8] A positive NH ΔSNR_{90} indicates that the female token (talker 2) is more robust to noise, whereas a negative value indicates that the male token (talker 1) is more robust to noise. The labels *S* and *Z* correspond to /ʃ/ and /ʒ/, respectively. Abbreviations: NH, normal hearing; SNR, signal-to-noise ratio.

For example, Fig. 2 (top left) shows that ear 40L reaches ≥50% two-token average error for /b, g, m, n, v/ as noise is introduced; but when the error is analyzed on an individual token basis, one finds that the errors for /g, m/ are completely due to the female token and that the error for /v/ is completely due to the male token. Similarly, ear 34L reaches ≥50% two-token average error for /b, g, k, p, v, z/, but, across the two tokens, large differences in noise robustness exist for /k, m, s, v/.

To quantify these observations, the difference in error between the two tokens of each consonant is calculated as a function of SNR. The *average error difference*, $\overline{\Delta P_e}$, for a given consonant is formulated as:

$$\overline{\Delta P_e} = \frac{1}{n(S)} \sum_{s \in S} \left(P_e^{T_1}(s) - P_e^{T_2}(s) \right)$$

$$S = \left\{ s \in \{ 0,\ 6,\ 12,\ Quiet \} : s \le s^* \right\}$$

where P_e is the probability of error at SNR s, s^* is the highest SNR at which more than one error is observed for either of the two tokens, and $n(S)$ indicates the number of elements in set S. Here, T_1 and T_2 indicate the tokens from talkers 1 (male) and 2 (female), respectively. $\overline{\Delta P_e}$ for each consonant is computed over the SNRs at which error is observed for at least one of the two tokens, to better capture the largest observed differences in error. If no error is observed over all SNRs, $\overline{\Delta P_e} \cong 0$.

The $\overline{\Delta P_e}$ is shown for example ears 40L and 34L in Fig. 2 (top right and bottom right). A negative $\overline{\Delta P_e}$ indicates that the male token is more robust to noise, whereas a positive value indicates that the female token is more robust. The consonants with the largest absolute $\overline{\Delta P_e}$ values for ear 40L are /g, m, v/, and /m, k, s/ for ear 34L. $\overline{\Delta P_e} = 0.4$ is marked for reference; a 0.4 error difference corresponds to at least two trials in error ($5 \le N \le 10$). The $\overline{\Delta P_e}$ values for 17 HI ears are shown in Fig. 3 (top). Large within-consonant differences is a widespread effect, with 16 of 17 ears showing at least one $\overline{\Delta P_e} > 0.4$.

We can compare the differences in token noise robustness for HI listeners to that of NH listeners with the same test tokens. The consonants in Fig. 3 (top) are plotted in the same order along the abscissa as the NH ΔSNR_{90} values, shown in Fig. 1 (bottom). This is done to determine if the token that is more robust to noise for a NH listener is also more robust for an HI listener. An increasing trend can be observed in the mean HI $\overline{\Delta P_e}$ values, similar to the trend of the NH ΔSNR_{90} values. A linear regression between the two measures is plotted in Fig. 3 (bottom); the HI $\overline{\Delta P_e}$ and NH ΔSNR_{90} values are significantly correlated ($\rho = 0.81, p < 0.001$). If

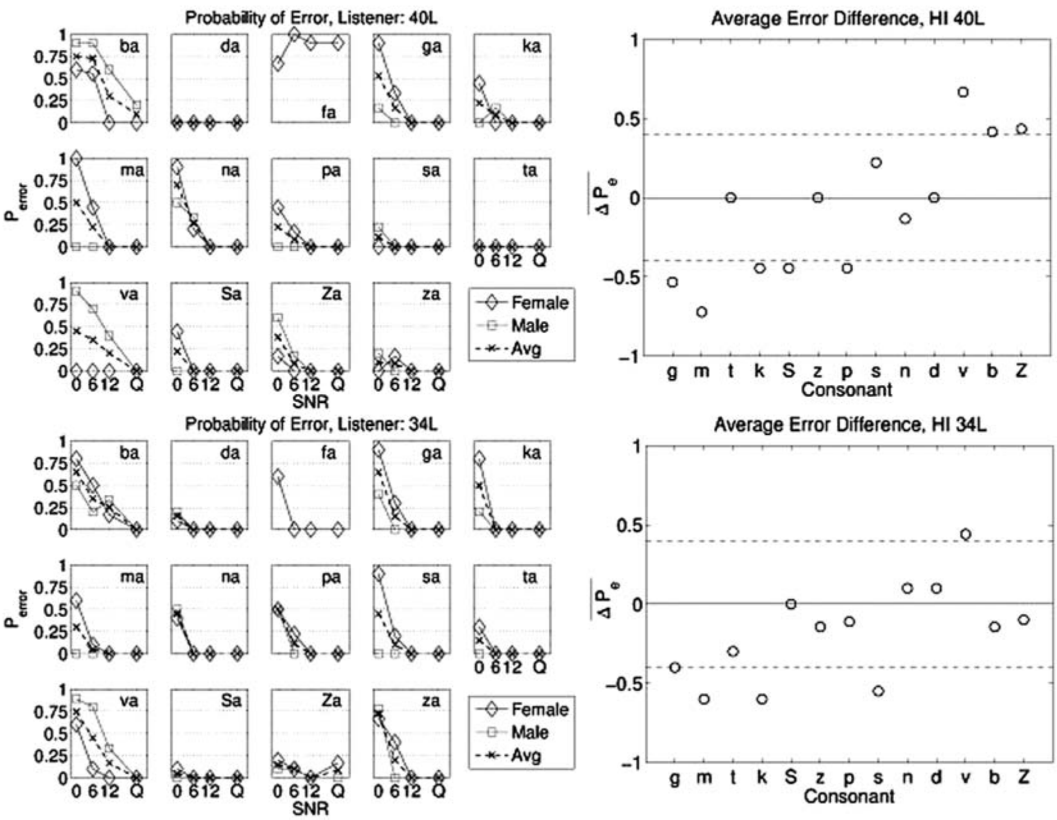

Figure 2 HI ears (top left and bottom left) 40L and 34L. Consonant recognition error as a function of SNR, each of the 14 consonant subplots shows the empirical error probability for a female token (diamond marker), a male token (square marker), and the average across the two talkers (*x* marker). Data for the male /fɑ/ not included in the analysis (top right, bottom right). $\overline{\Delta P_e}$ for each consonant; consonants are ordered along the abscissa based on the NH ΔSNR_{90} values (as in Fig. 1 at the bottom). $\overline{\Delta P_e} = \pm 0.4$ is marked for reference. The labels *S* and *Z* correspond to /ʃ/ and /ʒ/, respectively. Abbreviations: HI, hearing impaired; NH, normal hearing; SNR, signal-to-noise ratio.

the HI $\overline{\Delta P_e}$ values are computed as the average over all tested SNRs (i.e., $n(S) = 4$ fixed for all consonants), then the correlation coefficient is lower but remains significant ($\rho = 0.77$). This agreement across HI and NH listeners, in the relative noise robustness of tokens, suggests that both sets of listeners use the same primary cues, and that the audibility of these cues plays a major role in the relative noise robustness for HI listeners.

The consonants /ʃ, z, p, s, n/ all have a NH $\Delta SNR_{90} \leq 3$ dB, indicating that both tokens have consonant cue regions that are nearly equal in intensity, and thus audibility. Despite this, some HI listeners can have large $\overline{\Delta P_e}$ values for these consonants. In these cases, additional signal properties (e.g., variable temporal or frequency characteristics) may play a role; one such variable signal property is the presence of *conflicting cues*,[6,7] which are naturally occurring consonant cues that are not necessary for correct perception of the target consonant and, instead, are cues for nontarget consonants.

Differences in Token Consonant Confusion Groups

Study of the consonant confusion groups for NH listeners (e.g., /b, d, g/, /p, t, k/, /m, n/) dates back to the groundbreaking work of Miller and Nicely.[35] We have shown that acoustic cue variability across different tokens of the same consonant can lead to differences in noise robustness. The natural variability of speech also can lead to within-consonant differences in confusion groups.

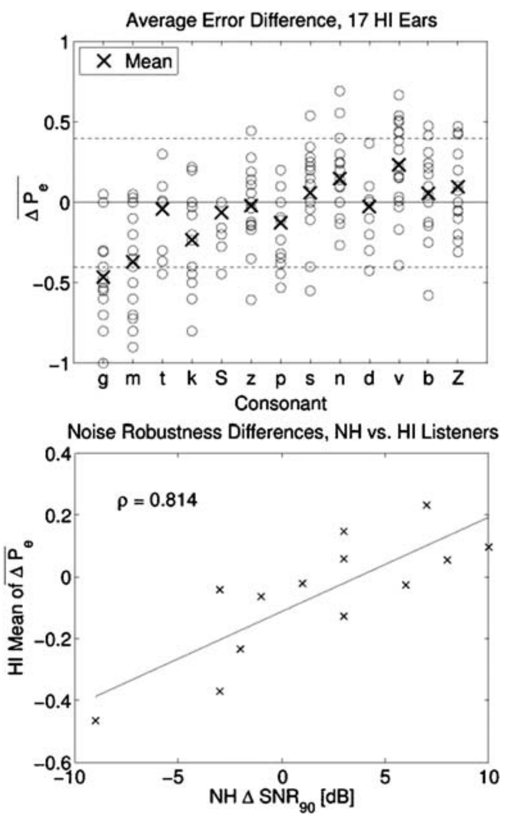

Figure 3 (Top) $\overline{\Delta P_e}$ for 17 HI ears. Each circle marker represents the value for a single HI ear, the mean across ears for each consonant is marked with an *x*. A negative $\overline{\Delta P_e}$ indicates that the male token has lower error, a positive value indicates that the female token has lower error. Consonants are ordered along the abscissa based on the NH ΔSNR_{90} values (as in Fig. 1 at the bottom). $\overline{\Delta P_e} = 0.4$ is marked for reference. The labels S and Z correspond to /ʃ/ and /ʒ/, respectively. (Bottom) Comparison and linear regression of the mean HI $\overline{\Delta P_e}$ values and the NH ΔSNR_{90} values (see Fig. 1 at the bottom); the two values are significantly correlated ($\rho = 0.81$, $p < 0.001$). Abbreviations: HI, hearing impaired; NH, normal hearing; SNR, signal-to-noise ratio.

The confusions for the two tokens of the consonant /b/ + /a/, are first analyzed in detail. Fig. 4 (top left, bottom left) shows data for six HI ears, at four SNRs. The height of each bar represents the overall probability of error, and the confusions as a proportion of this error are indicated by pattern. Individual HI listeners are separated by vertical bars that lie over the plots. For the /ba/ token from the female talker, a large proportion of the confusions are with /d, v, g/. In contrast, the /ba/ token from the male talker shows a high probability of confusions with /f, v/, over the six HI ears. The average responses for each of the two /ba/ tokens (over 17 HI ears and four SNRs) are shown as two pie charts in Fig. 4 (top right, bottom right). These pie charts show this difference in confusion groups for the two /ba/ tokens, over a population of 17 ears.

For each consonant token, the average error over 17 HI ears and four SNRs, and the confusions that comprise this error, are shown in Fig. 5. We see within-consonant differences for the confusion groups of /ba, ga, ma, sa, ʒa/. Although some confusions are shared across tokens of a consonant, distinct within-consonant differences are common. The size of the average individual token confusion group can be small (the majority of the total responses are accounted for by ≤4 confusions), indicating that the majority of the responses across the 17 HI ears and four SNRs are from the same confusion group. The HI ears make similar confusions for a given token, despite the many subject differences (e.g., degree of speech errors, hearing loss, age, gender, and background). This implies that the variable acoustic properties of each token (primary and conflicting acoustic cues) determine the possible HI confusions. If the responses to multiple tokens of a single consonant are averaged together, different confusion groups may be averaged together and cause an HI listener to appear to be guessing or more "random" in their speech perception than they actually are.

OVERVIEW OF THE ERROR DIFFERENCES FOR INDIVIDUAL EARS

An overview plot of the distribution of errors across the set of consonant test tokens allows one to visualize the underlying token errors that are often reported as a single averaged value. We have observed that large within-consonant differences exist in noise robustness/error, the same recognition data, now compared across all consonant tokens is examined here.

An error overview plot for an individual HI ear shows the error for all test tokens, with the tokens sorted along the abscissa to create a monotonically increasing plot. This monotonically increasing plot allows one to clearly see the nature of the errors; that is, whether errors are

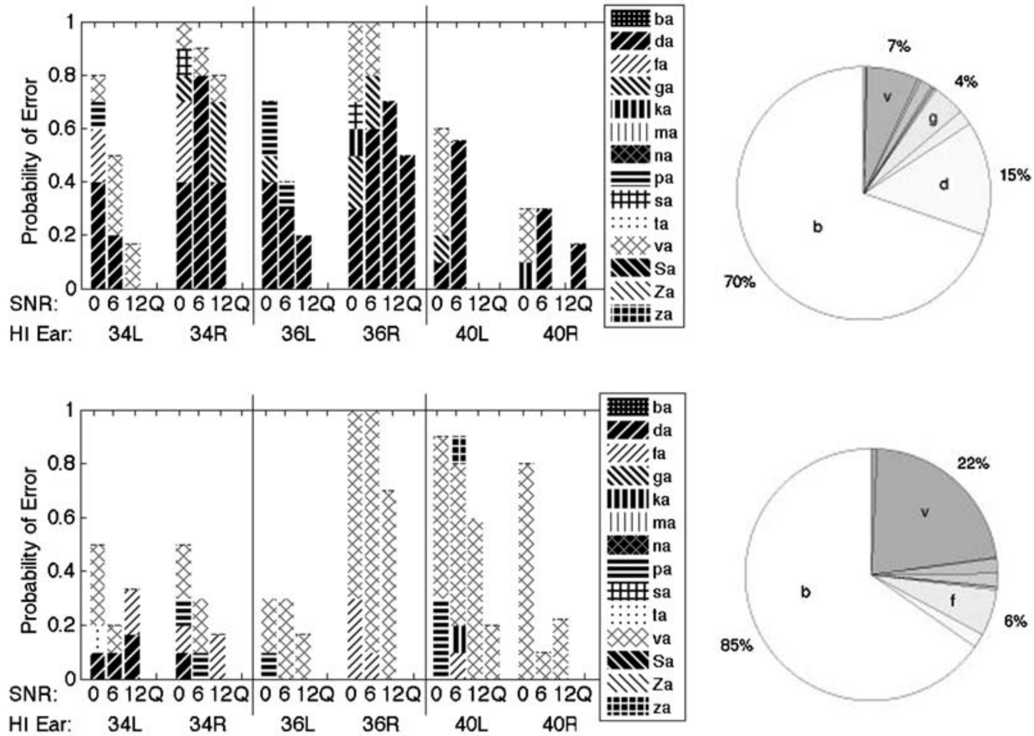

Figure 4 Probability of error and confusions for two tokens of /ba/. (Top left) Error for the female /ba/ token, data from three HI listeners (data for both ears shown) at four SNRs. Confusions as a proportion of the total error are labeled by pattern. (Top right) The proportion of responses for the female /ba/ token, averaged across 17 HI ears and four SNRs; primary confusions are with /d, v, g/. (Bottom left) Error for the male /ba/ token, data at four SNRs is from the same six HI ears as in (the top left). Confusions as a proportion of the total error are labeled by pattern. (Bottom right) The proportion of responses for the male token, averaged across 17 HI ears and four SNRs; primary confusions are with /f, v/. Abbreviations: HI, hearing impaired; SNR, signal-to-noise ratio.

widespread over all stimuli or if, instead, high errors are only observed for a specific subset of stimuli.

Overview plots of the errors across 28 test stimuli (2 talkers $\times$ 14 consonants), are shown in Fig. 6, with each line representing the data for a single HI ear. The distributions of error in quiet (Fig. 6 at the top) and in speech-shaped noise (Fig. 6 at the bottom) show that high degrees of error can be concentrated to small subsets of the total consonant stimuli. In the quiet condition, all tested HI ears with slight to mild hearing loss have no errors for the majority of consonant tokens, thus, these tokens are audible. Despite the large number of tokens with no errors, the same ears can reach $> 50\%$ error with the remaining tokens. In noise, as might be expected, higher degrees of error are observed across the test tokens; again, HI ears show a combination of some tokens with zero measured error and others with high error.

THE UNDERLYING INDIVIDUAL DIFFERENCES WITHIN THE AVERAGE CONSONANT ERROR

The average speech score, across many different stimuli, or the SNR at which the average speech score hits 50% are often used to compare HI subjects. Can HI listeners with similar pure tone thresholds and similar average speech scores in both quiet and noise have different individual consonant tokens that are perceived in error?

Fig. 7 shows data for three HI listeners (one ear per listener) who have very similar pure tone thresholds and whose average error in quiet and noise also is matched (within 10% in 6- and 12-dB SNR speech-shaped noise and in quiet). In addition, the average consonant error for all three HI ears is within 10% of NH performance at 12-dB SNR and quiet. Note that the average error for an HI ear is approximately linear on a log scale with respect to *s*, just

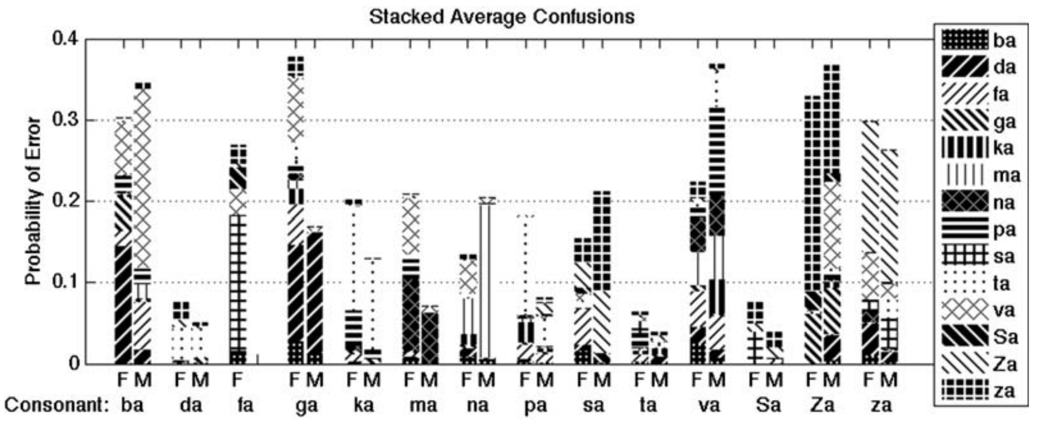

Figure 5 The probability of error averaged over 17 HI ears and four SNRs for each consonant token. The confusions that make up the error are labeled by pattern. Token differences in confusion group are seen here for /ba, ga, ma, sa, ʒa/. The labels *S* and *Z* correspond to /ʃ/ and /ʒ/, respectively. Abbreviations: HI, hearing impaired; SNR, signal-to-noise ratio.

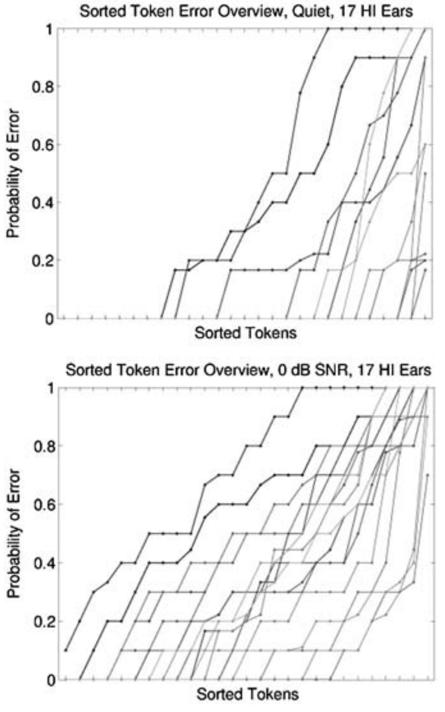

Figure 6 Sorted error overviews of the 28 test tokens (14 consonants × 2 talkers), for (top) 17 HI ears in quiet, and (bottom) 17 HI ears in 0-dB SNR speech-shaped noise. Each line is an individual HI ear at a fixed SNR. For each individual ear, the tokens are sorted along the abscissa, creating a monotonically increasing error plot. Both with and without noise, errors for each HI ear are concentrated to a small subset of the stimuli. In quiet, all ears with slight to mild hearing loss show zero error for the majority of test tokens. Abbreviations: HI, hearing impaired; SNR, signal-to-noise ratio.

like the error predicted by the articulation index formula.[36] As in the previous sections of this article, 14 consonants with 2 tokens per consonant were used as stimuli. The error for individual consonant tokens is compared with the average error for three example ears in Fig. 7 (top right, bottom left and bottom right).

Despite the similar pure tone thresholds and average errors, the individual consonants that are perceived in error are ear-dependent. Ear 32L shows errors with a large range of consonant tokens, particularly of /b, g, n, v, z/, reaching almost 100% error with the male /n, v, z/ and the female /g/ tokens. Ear 36R shows high error with only the two /b/ tokens and the male /v/. Ear 40L also has difficulty with a wider array of consonants in noise, showing consistent errors with tokens of /b, f, g, n, m, v/. We see that similar pure tone thresholds and average errors can have underlying individual variability in terms of consonant perception.

A listener with a low average error can also have underlying high errors with small, specific set of stimuli. In terms of the average, ear 36R (Fig. 7 bottom left) appears to be almost normal, falling within 10% of NH performance at 6- and 12-dB SNR and quiet. However, when the individual consonant errors are examined, ear 36R actually has very high error for both tokens of /b/ and the female token of /v/; in fact, 36R has 100% error for both /b/ tokens at 0- and 6-dB SNR. Thus, an ear that appears

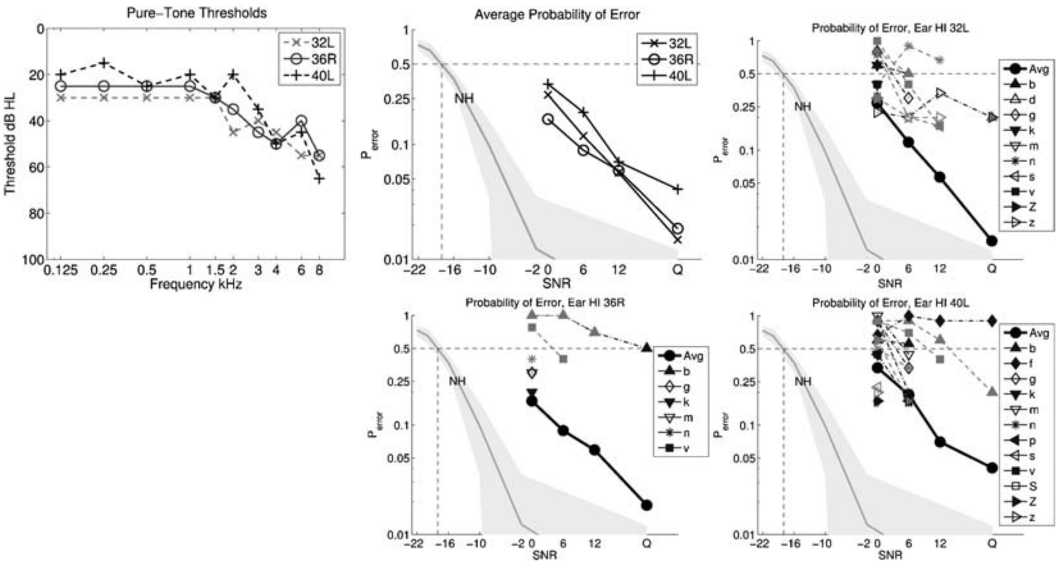

Figure 7 (Top left) The pure tone thresholds of three HI listeners (one ear per listener shown). (Top middle) Average consonant error over 28 consonant tokens (14 consonants $\times$ 2 talkers), for individual HI and NH listeners, on a log scale (shaded region for NH listener data denotes one standard deviation). The HI data are shown for 0-, 6-, and 12-dB SNR and quiet. (Top right, bottom left and right) Individual and average token errors for HI ears 32L, 36R, and 40L. For the individual consonant tokens, dash-dot black lines indicate a token from a female talker, dashed gray lines indicate a token from a male talker. Random errors $\leq 1/N$ are not shown to reduce clutter. The labels *S* and *Z* correspond to /ʃ/ and /ʒ/, respectively. Abbreviations: HI, hearing impaired; NH, normal hearing; SNR, signal-to-noise ratio.

"almost normal" in terms of the average speech score can actually have a severe problem with a very specific type of speech sound.

SUMMARY

The natural within-consonant variability always should be considered when interpreting the results of an HI consonant test. HI within-consonant perceptual differences are observed in terms of both noise robustness and confusion groups for tokens that show no error and are unambiguous for NH listeners. Although tokens are similar enough to be categorized as the same consonant by NH listeners, existing subtle natural variations in signal properties can lead to systematic differences in HI perception. Characterizing the primary and any conflicting perceptual cues of test tokens is thus critically important to understanding HI perception of naturally spoken consonants.

The within-consonant recognition differences in noise robustness across two tokens of a consonant can be as large as a difference in error probability of 100% at a fixed SNR. Agreement between NH and HI ears in the relative noise robustness of tokens indicates that signal properties, in part the relative intensity and thus audibility of the consonant cues, are a source of noise robustness differences. Other varying signal properties, such as variations in temporal and frequency characteristics of the primary consonant cues and the presence of conflicting consonant cues, may also play a role in these differences.

Within-consonant differences in the confusion groups are observed across a population of 17 HI ears. When the confusions of HI ears are examined on a token (as opposed to a consonant) level, one observes that they are much more self-consistent in their responses (i.e., the confusion groups are smaller). A population of 17 HI ears all tended to respond from a single limited token-dependent confusion group, suggesting that the acoustic properties of each token determine the confusions. Analysis of which acoustic cues lead to particular confusions has the potential to provide new insight into the speech perception strategies that are being used by HI listeners.

We observe that listeners with slight to moderate hearing loss have difficulty with only a small subset of stimuli; for some ears, this holds true even at the 0-dB SNR noise condition. These observations are particularly relevant to clinical interpretations of average speech measures. Systematic errors with a particular type of speech signal can lead to difficultly in understanding words that contain these consonants; for natural conversation, the need for correct perception of every phoneme decreases as context is introduced via meaningful words and sentences.[31] This provides a plausible explanation for why hearing impairment deficits are more difficult to observe in tasks that provide context.

Even when HI listeners have similar pure tone thresholds and average speech errors, perception of individual consonant stimuli can be widely different. Further exploration of the sources of these individual differences will improve our understanding of how HI listeners process and interpret naturally spoken consonant cues. Methods need to be developed to apply these findings in a hearing aid paradigm.

HI Subjects

The majority of HI subjects in this article (16 of 17 ears) have approximately flat loss below 1 to 3 kHz with high-frequency sloping loss. One ear (14R) has the most hearing loss < 2 kHz, with an inverted high-frequency loss and a pure tone threshold within normal ranges at 8 kHz. Tympanometric measures showed no middle ear pathologies for all ears. Pure tone averages are listed in Table 1.

Table 1 17 HI Ears Ordered by the Average of the Left and Right Ear 3-Tone (0.5, 1, 2 kHz) PTA, with Age of the Listener and MCL for Each Ear

HI Subject Data			
HI ear	PTA	Age	MCL
44L	10	65	82
44R	15	65	78
46L	8	67	82
46R	17	67	82
40L	22	79	80
40R	23	79	80
36L	27	72	68
36R	28	72	70
30L	30	66	80
30R	27	66	80
32L	35	74	79
32R	27	74	77
34L	32	84	84
34R	28	84	82
01L	45	82	83
01R	47	82	82
14R	73	25	89

Abbreviations: HI, hearing impaired; MCL, most comfortable level; PTA, pure tone average.

Speech Materials

The LDC-2005S22 database labels for the test tokens, along with the NH SNR_{90} values, are listed in Table 2. All SNR_{90} values are calculated by linear interpolation between measurements taken at −22, −20, −16, −10, and −2 dB.

Table 2 Male and Female Talker Labels for Each CV Token, Along with the Corresponding NH SNR_{90} Values

CV	M Talker	SNR_{90} (dB)	F Talker	SNR_{90} (dB)	CV	M Talker	SNR_{90} (dB)	F Talker	SNR_{90} (dB)
bɑ	m112	−2	f101	−10	pɑ	m118	−14	f103	−17
dɑ	m118	−7	f105	−13	sɑ	m120	−10	f103	−13
fɑ	m112[a]	−5[a]	f109	−12	ʃɑ	m118	−16	f103	−15
gɑ	m111	−12	f109	−3	tɑ	m112	−17	f108	−14
kɑ	m111	−13	f103	−11	vɑ	m118	−3	f101	−10
mɑ	m118	−14	f103	−11	ʒɑ	m107	−7	f105	−17
nɑ	m118	−4	f101	−7	zɑ	m118	−17	f106	−18

Abbreviations: CV, consonant-vowel; F, female; M, male, NH, normal hearing, SNR, signal-to-noise ratio.
[a]This token was not included in the data analysis.

ACKNOWLEDGMENTS

This work would not have been possible without the dedication of Professor Woojae Han, who collected the hearing-impaired data and conducted an initial analysis, as partial fulfillment of her Ph.D.[37] We would also like to thank the members of the HSR research group for their many critical discussions. This study was conducted with the support of NIH Grant #RDC 009277A and a generous grant from Research in Motion.

REFERENCES

1. Baum SR, Blumstein SE. Preliminary observations on the use of duration as a cue to syllable-initial fricative consonant voicing in English. J Acoust Soc Am 1987;82:1073–1077
2. Dorman M, Studdert-Kennedy M, Raphael L. Stop-consonant recognition: release bursts and formant transitions as functionally equivalent, context-dependent cues. Atten Percept Psychophys 1977;22:109–122
3. Herd W, Jongman A, Sereno J. An acoustic and perceptual analysis of /t/ and /d /flaps in American English. J Phonetics 2010;38:504–516
4. Jongman A, Wayland R, Wong S. Acoustic characteristics of English fricatives. J Acoust Soc Am 2000;108(3 Pt 1):1252–1263
5. Kurowski K, Blumstein SE. Acoustic properties for place of articulation in nasal consonants. J Acoust Soc Am 1987;81:1917–1927
6. Li F, Menon A, Allen JB. A psychoacoustic method to find the perceptual cues of stop consonants in natural speech. J Acoust Soc Am 2010; 127:2599–2610
7. Li F, Trevino A, Menon A, Allen J. A psychoacoustic method for studying the necessary and sufficient perceptual cues of fricative consonants in noise. J Acst Soc Amer 2012;132:2663–2675
8. Phatak SA, Allen JB. Consonant and vowel confusions in speech-weighted noise. J Acoust Soc Am 2007;121:2312–2326
9. Singh R, Allen JB. The influence of stop consonants' perceptual features on the Articulation Index model. J Acoust Soc Am 2012;131:3051–3068
10. Skinner M. Speech intelligibility in noise-induced hearing loss: effects of high-frequency compensation. Doctoral Dissertation; 1976
11. Skinner MW, Miller JD. Amplification bandwidth and intelligibility of speech in quiet and noise for listeners with sensorineural hearing loss. Audiology 1983;22:253–279
12. Kamm CA, Dirks DD, Bell TS. Speech recognition and the Articulation Index for normal and hearing-impaired listeners. J Acoust Soc Am 1985; 77:281–288
13. Smoorenburg GF. Speech reception in quiet and in noisy conditions by individuals with noise-induced hearing loss in relation to their tone audiogram. J Acoust Soc Am 1992;91:421–437
14. Roeser R, Valente M, Hosford-Dunn H. Audiology: Diagnosis. New York, NY: Thieme Medical Publishers; 2007
15. Halpin C, Rauch SD. Clinical implications of a damaged cochlea: pure tone thresholds vs information-carrying capacity. Otolaryngol Head Neck Surg 2009;140:473–476
16. Walden BE, Montgomery AA. Dimensions of consonant perception in normal and hearing-impaired listeners. J Speech Hear Res 1975;18: 444–455
17. Killion M, Niquette P. What can the pure-tone audiogram tell us about a patients SNR loss. Hear J 2000;53:46–53
18. Mines MA, Hanson BF, Shoup JE. Frequency of occurrence of phonemes in conversational English. Lang Speech 1978;21:221–241
19. Hood JD, Poole JP. Improving the reliability of speech audiometry. Br J Audiol 1977;11:93–101
20. Burkle T, Kewley-Port D, Humes L, Lee J. Contribution of consonant versus vowel information to sentence intelligibility by normal and hearing-impaired listeners. J Acoust Soc Am 2004;115:2601
21. Lawrence DL, Byers VW. Identification of voiceless fricatives by high frequency hearing impaired listeners. J Speech Hear Res 1969;12:426–434
22. Bilger RC, Wang MD. Consonant confusions in patients with sensorineural hearing loss. J Speech Hear Res 1976;19:718–748
23. Owens E. Consonant errors and remediation in sensorineural hearing loss. J Speech Hear Disord 1978;43:331–347
24. Wang MD, Reed CM, Bilger RC. A comparison of the effects of filtering and sensorineural hearing loss on patients of consonant confusions. J Speech Hear Res 1978;21:5–36
25. Dubno JR, Dirks DD. Evaluation of hearing-impaired listeners using a Nonsense-Syllable Test. I. Test reliability. J Speech Hear Res 1982; 25:135–141
26. Boothroyd A. Auditory perception of speech contrasts by subjects with sensorineural hearing loss. J Speech Hear Res 1984;27:134–144
27. Fabry DA, Van Tasell DJ. Masked and filtered simulation of hearing loss: effects on consonant recognition. J Speech Hear Res 1986;29:170–178
28. Dreschler WA. Phonemic confusions in quiet and noise for the hearing-impaired. Audiology 1986;25:19–28
29. Gordon-Salant S. Consonant recognition and confusion patterns among elderly hearing-impaired subjects. Ear Hear 1987;8:270–276

30. Zurek PM, Delhorne LA. Consonant reception in noise by listeners with mild and moderate sensorineural hearing impairment. J Acoust Soc Am 1987;82:1548–1559
31. Boothroyd A, Nittrouer S. Mathematical treatment of context effects in phoneme and word recognition. J Acoust Soc Am 1988;84:101–114
32. Bronkhorst AW, Bosman AJ, Smoorenburg GF. A model for context effects in speech recognition. J Acoust Soc Am 1993;93:499–509
33. Bronkhorst AW, Brand T, Wagener K. Evaluation of context effects in sentence recognition. J Acoust Soc Am 2002;111:2874–2886
34. Régnier MS, Allen JB. A method to identify noise-robust perceptual features: application for consonant /t/. J Acoust Soc Am 2008;123:2801–2814
35. Miller G, Nicely P. An analysis of perceptual confusions among some English consonants. J Acoust Soc Am 1955;27:338
36. Allen J. How do humans process and recognize speech? Speech and audio processing. IEEE Transactions on 1994;2:567–577
37. Han W. Methods for Robust Characterization of Consonant Perception in Hearing-Impaired Listeners [Ph.D. thesis]. Urbana-Champaign, IL: University of Illinois; 2011