Within-consonant perceptual differences in the hearing impaired ear

Andrea Trevino[a] and Jont B. Allen
Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

(Received 4 October 2012; revised 4 March 2013; accepted 23 April 2013)

The consonant recognition of 17 ears with sensorineural hearing loss is evaluated for 14 consonants /p, t, k, f, s, ʃ, b, d, g, v, z, ʒ, m, n/ + /ɑ/, under four speech-weighted noise conditions (0, 6, 12 dB SNR, quiet). One male and one female talker were chosen for each consonant, resulting in 28 total consonant-vowel test tokens. For a given consonant, tokens by different talkers were observed to systematically differ, in both the robustness to noise and/or the resulting confusion groups. Such within-consonant token differences were observed for over 60% of the tested consonants and all HI ears. Only when HI responses are examined on an individual token basis does one find that the error may be limited to a small subset of tokens with confusion groups that are restricted to fewer than three confusions on average. Averaging different tokens of the same consonant can raise the entropy of a listener's responses (i.e., the size of the confusion group), causing the listener to appear to behave in a less systematic way. Quantifying these token differences provides insight into HI perception of speech under noisy conditions and characterizes each listener's hearing impairment.
[http://dx.doi.org/10.1121/1.4807474]

PACS number(s): 43.71.Ky, 43.71.Es [JFC] Pages: 1–11

I. INTRODUCTION

Given that the primary purpose of wearing a hearing aid is to improve speech perception, it follows that a speech test should be able to provide one of the most useful measures of hearing impairment. Yet speech has not been found to be a useful tool for fitting hearing aids (Walden *et al.*, 1983; Dobie, 2011). Pure-tone thresholds remain the primary prescriptive measure for hearing aid fitting (Humes *et al.*, 1991; Dillon, 2001) despite the common clinical observation that hearing impaired (HI) ears can have similar pure-tone thresholds but differ in their speech perception abilities (Skinner, 1976; Skinner and Miller, 1983; Kamm *et al.*, 1985; Smoorenburg, 1992; Roeser *et al.*, 2007; Halpin and Rauch, 2009; Walden and Montgomery, 1975). A significant impediment to research in developing speech-based measures is the large amount of natural variability that is present in speech; this causes difficulty in identifying and acoustically characterizing the perceptually relevant cues. When the perceptual cues of the tokens that are used in a speech test are not precisely characterized, the conclusions that may be drawn are limited.

The work of Boothroyd and Nittrouer (1988) formulated the relationship between correct perception of low-context speech segments (e.g., phonemes) and high-context segments (e.g., words) in normal hearing (NH) ears. Follow-up studies by Bronkhorst *et al.* (Bronkhorst *et al.*, 1993; Bronkhorst *et al.*, 2002) greatly extended this work. These studies demonstrate that an individual's ability to decode high-context speech depends critically on their low-context error. These observations affirm the utility of studies of hearing impairment that use low-context speech segments.

Consonants comprise approximately 58.5% of conversational speech (Mines *et al.*, 1978). While the relative importance of consonants and vowels for HI speech perception remains uncertain (Hood and Poole, 1977; Burkle *et al.*, 2004), here we concentrate on HI consonant perception. Many past works have examined HI consonant recognition using naturally produced speech, including Lawrence and Byers (1969), Bilger and Wang (1976), Owens (1978), Wang *et al.* (1978), Dubno and Dirks (1982); Boothroyd (1984), Fabry and Van Tasell (1986), Dreschler (1986), Gordon-Salant (1987), and Zurek and Delhorne (1987). Overall, the effects of hearing impairment on speech perception are more severe in the presence of noise (Dubno and Dirks, 1982; Dreschler, 1986). It has been observed that listeners with similar perceptual problems can have similar audiometric configurations (Bilger and Wang, 1976) but also that some consonant confusions are common across a variety of audiometric configurations (Owens, 1978; Gordon-Salant, 1987). In addition, comparisons between the consonant recognition errors of HI listeners vs NH listeners with simulated hearing losses (noise and/or filtering applied) has shown some agreement in both errors (Zurek and Delhorne, 1987) and confusions (Wang *et al.*, 1978; Fabry and Van Tasell, 1986). In these past studies, data analysis was performed using either an average measure (over all consonants) or with consonants grouped by distinctive features. Speech measures derived from an average have been useful tools for screening and classifying those with a hearing impairment; however, they have not proven useful as prescriptive measures (Taylor, 2006; Killion and Gudmundsen, 2005).

In this work, we examine how HI perception can vary across tokens of the same consonant. Multiple tokens of the

---

[a] Author to whom correspondence should be addressed. Electronic mail: atrevin2@illinois.edu

same consonant, by different talkers or with different vowels, are often considered as multiple measures of the same effect. In contrast to this approach, the consonant cue literature has documented, in detail, the variability of the cues that are present in naturally produced speech (Baum and Blumstein, 1987; Dorman *et al.*, 1977; Herd *et al.*, 2010; Jongman *et al.*, 2000; Kurowski and Blumstein, 1987; Li *et al.*, 2010; Li *et al.*, 2012). This variability is quantified by an analysis of the acoustical properties of each individual consonant token and can be observed across speech samples that are unambiguous (no confusions) and robust to noise (error $<10\%$). Although NH listeners can correctly recognize consonants despite this variability, the question remains: Does this natural variability across tokens of the same consonant lead to differences in HI perception?

We show that HI perceptual differences exist across multiple tokens of a single consonant (which show no recognition differences for NH listeners). We refer to perceptual differences across multiple tokens of the same consonant as *within-consonant differences*. The HI within-consonant differences are observed in terms of both robustness to noise and/or confusion groups. These two types of within-consonant differences can exist independently of each other.

Within-consonant differences in noise robustness are observed over all of the HI subjects. Previous studies have shown that for individual consonant tokens, the intensity of each necessary cue region is correlated to the robustness to noise for NH listeners (Régnier and Allen, 2008; Li *et al.*, 2010; Li *et al.*, 2012; Kapoor and Allen, 2012). We test if natural variations in the intensity of the acoustic cue region that affect NH perception at low SNRs would similarly affect HI perception at higher SNRs. Although a significant correlation is observed, HI within-consonant noise-robustness differences in this study are only partially explained by the natural variations in the intensity of the necessary consonant cue region. To further examine if the variability in the acoustic properties can lead to differences in HI perception, the confusion groups of individual tokens are also analyzed.

We observe that each token has a unique subgroup of possible confusions and that these confusion groups can be different for each token of the same consonant. Thus the existing subtle differences in acoustical properties, which do not affect NH recognition, can lead to large differences in confusion groups for HI listeners. The responses of HI ears to stimuli can often appear to be "random." This study finds that such randomness can be an artifact of averaging; only when the slight-to-moderate HI subjects are examined at the token level does one observe that the subjects are self-consistent in their confusions.

When testing HI ears, the selection of the individual tokens for a perceptual experiment 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. Each token of a consonant may be considered as a sensitive probe that can provide fine-grained information about a person's hearing impairment. Thus we can use the natural variability of speech to advantage but only once we have controlled for it.

II. METHODS

A. Subjects

Nine HI subjects were recruited for this study from the Urbana-Champaign, IL, community. Both ears were tested for all listeners but one, resulting in data for 17 individual ears. All subjects reported American English as their first language and were paid to participate. IRB approval was obtained prior to the experiment. Typanometric measures showed no middle-ear pathologies (type A tympanogram). The ages of eight HI subjects ranged from 65 to 84; one HI subject (14R) was 25 yrs old. Based on the pure-tone thresholds, all ears had >20 dB of hearing loss (HL) for at least one frequency in the range 0.25-4 kHz.

B. Audiometric measurements

The majority of the ears in our study have slight-to-moderate hearing loss with high-frequency sloping configurations (see Table I). One HI ear (14R) has an inverted high-frequency loss with the most hearing loss <2 kHz and a threshold within the normal range at 8 kHz. The audiometric configuration of low-frequency flat loss with high-frequency sloping loss can be modeled as a piecewise linear function of the form

$$h = \begin{cases} h_0 & \text{if } f \le f_0 \\ h_0 + s_0(\log_2(f/f_0)) & \text{if } f > f_0, \end{cases} \qquad (1)$$

where h is the hearing loss (dB) and f is frequency (kHz). The parameter f_0 estimates the frequency at which the sloping loss begins; h_0 estimates the low-frequency ($f \le f_0$) flat loss in decibels; s_0 estimates the slope of the high-frequency

TABLE I. The 17 HI ears are ordered by the average of the left and right ear h_0 values [Eq. (1)]. The model parameters estimate the flat low-frequency loss h_0 (dB), the frequency at which sloping loss begins f_0 (kHz), and the sloping high-frequency loss s_0 (dB/octave). RMS error ϵ (dB) of the model fits. The age of the listener and most comfortable level (MCL) for each ear are included. The mean and standard deviation (μ,σ) for all values are reported in the bottom row (ear 14R excluded).

HI ear	h_0	f_0	s_0	RMS ϵ	Age	MCL
44L	9	1	10	11	65	82
44R	13	1	7	7	65	78
46L	11	1.5	20	9	67	82
46R	18	3	27	7	67	82
40L	22	2	20	5	79	80
40R	18	1	11	5	79	80
36L	19	1	7	8	72	68
36R	25	1	10	4	72	70
30L	28	1.5	22	3	66	80
30R	25	1.5	27	5	66	80
32L	30	1	9	3	74	79
32R	27	1.5	14	3	74	77
34L	34	3	50	6	84	84
34R	26	1.5	26	4	84	82
01L	44	4	33	2	82	83
01R	47	3	41	4	82	82
14R	72	2	−37	3	25	89
(μ, σ)	(25, 11)	(2, 0.9)	(21, 13)	(5, 2)	(74, 7)	(79, 4)

loss in decibels/octave. The three parameters are fit to minimize the root-mean-square (RMS) error ϵ (dB). The resulting RMS ϵ values for each model fit are reported in Table I.

C. Speech materials

All stimuli used in this study were selected from the Linguistic Data Consortium Database (LDC-2005S22) (Fousek *et al.*, 2004). Speech was sampled at 16 kHz. Fourteen naturally spoken American English consonants (/p, t, k, f, s, ʃ, b, d, g, v, z, ʒ, m, n/ + /ɑ/) were used as the test stimuli. Each consonant was spoken in an isolated (i.e., no carrier phrase) consonant-vowel (CV) context, with the vowel /ɑ/. Speech samples from six female talkers and five male talkers were used (see Table IV), with two tokens selected (one male and one female talker) for each consonant, resulting in a total of 28 test tokens (14 consonants × 2 talkers = 28 tokens). The term *token* is used throughout this work to refer to a single CV speech sample from one talker.

The 28 test tokens were selected based on their NH perceptual scores in quiet and speech-weighted noise. To ensure that tokens were unambiguous and robust to noise, each token was selected based on a criteria of $\leq$3.1% error for a population of 16 NH listeners, calculated by combining results in quiet and −2 dB signal-to-noise ratio (SNR) of noise (i.e., no more than 1 error over a total $N = 32$, per token) (Phatak and Allen, 2007). Such tokens are representative of the LDC database; Singh and Allen (2012) shows, for the majority of tokens, a ceiling effect for NH listeners above −2 dB SNR. One token of /fɑ/ (male talker, label m112) was damaged in the preparation of the tokens, thus it has not been included in this analysis.

The stimuli were presented with flat gain at the *most comfortable level* (MCL) for each individual HI ear. For the majority of the HI ears the MCL was approximately 80 ± 4 dB sound pressure level (SPL) (see Table I). Two subjects (36L/R and 14R) did not choose an MCL within this range.

D. Experimental procedure

The speech was presented at 4 SNRs (0, 6, and 12 dB and quiet) using speech-weighted noise generated as described by Phatak and Allen (2007). Presentations were randomized over consonant, talker, and SNR. For each HI ear, the experiment was performed in two sessions. The first session presented each consonant eight times (four per token) at each of the 4 SNRs, resulting in 32 presentations per consonant (4 presentations × 2 tokens × 4 SNRs). The second session used an adaptive scheme to selectively increase the number of presentations, and thus the statistical power of the test. For each token, the number of session two presentations ranged from 1 to 6 at each SNR with increased presentations assigned to conditions that had produced the most error in the first session. Thus the total number presentations of each consonant ranged from $N = 40$ to 80 for each HI ear (total $N = 5$-10 over 2 sessions × 2 tokens × 4 SNRs). The Vysochanskij–Petunin inequality (Vysochanskij and Petunin, 1980) was used to verify that the number of trials were sufficient to determine correct perception within a 95% confidence interval (see appendix of Singh and Allen, 2012).

The experiment was implemented as a MATLAB graphical user interface. All of the data-collection sessions were conducted with the subject seated in a single-walled, soundproof booth with the door of the outer lab closed. The speech was presented monoaurally via an Etymotic ER-3 insert earphone. The contralateral ear was not masked or occluded. The subject chose their MCL (for non-test speech samples) before testing began. Subjects were allowed to adjust the sound level at any time during the experiment; however, none of the nine HI subjects tested chose to make such an adjustment. A practice session, with different tokens from those in the test set, was run first in order to familiarize the subject with the testing paradigm. The remaining sessions presented the randomized test speech tokens. After hearing a single presentation of a token, the subject would choose from the 14 possible consonant responses by clicking one of 14 CV-labeled buttons on the graphical user interface with the option of up to two additional token repetitions to improve accuracy. Short breaks were encouraged to reduce the effects of test fatigue. Additional experimental details are provided in Han (2011).

E. Characterizing individual tokens with normal hearing psychoacoustic data

Psychoacoustic data from classical masking, filtering and time truncation experiments can be used to characterize the consonant cues of each token in terms of intensity, frequency, and temporal properties. NH listener psychoacoustic data for the 28 test tokens (14 consonants) used in the present study were collected by Phatak and Allen (2007) and Li (2011). High-/low-pass filtering and time-truncation data allow one to identify, in each naturally variable token, the spectral time-frequency region that contains the acoustic components that are necessary for correct perception, we refer to this as the *necessary cue region* (Li *et al.*, 2010; Li *et al.*, 2012). The acoustic components that encode the primary cues fall within this necessary cue region. As an example, the necessary cue region for a /sɑ/ token would include the frication noise that contains a spectral primary cue for place and the durational primary cue for manner of articulation.

A key metric of each token's noise robustness is the SNR_{90}, defined as the full-bandwidth SNR at which the probability of NH correct recognition for that individual token drops below saturation to 90%. The lower the SNR_{90}, the more robust a token is to noise. For NH listeners, this psychoacoustic measure has been found to be significantly correlated to the physical intensity of the necessary consonant cue region, with tokens that have more intense cue regions having lower SNR_{90} values (Régnier and Allen, 2008; Li *et al.*, 2010; Li *et al.*, 2012). As discussed in Sec. II C, the NH SNR_{90} values for the selected test tokens are below the worst noise condition that was used to test HI recognition in the present study, 0 dB SNR (see Appendix). Due to natural variability of cue region intensity, the SNR_{90} values for a large number of tokens are approximately Gaussian distributed (Singh and Allen, 2012).

It follows from these findings that for two tokens of the same consonant, the difference between the NH SNR_{90} values is proportional to the difference in intensity of the necessary acoustic cue regions. Because tokens of the same consonant have perceptual cues within a similar frequency range, the NH ΔSNR_{90} can be used to relate the audibility of their necessary cue regions. For each consonant, the SNR_{90} of the token from the male talker was subtracted from that of the female talker; this measure is illustrated in Fig. 1(a) with Δ marking the difference between the two SNR_{90} values. These differences are reported for each pair of consonant tokens in Fig. 1(b) with the consonants sorted along the abscissa by monotonically increasing NH ΔSNR_{90} values. This plot shows that for /g/, the male token is more robust to noise by 9 dB, whereas for /ʒ/, the female token is more robust to noise by 10 dB. Of the selected tokens, there are small differences in the noise robustness (less than or equal to ±3 dB) of eight consonants, /m, t, k, ʃ, z, n, p, s/. The NH ΔSNR_{90} values are controlled by the selection of the experimental tokens. Although the NH SNR_{90} was controlled in the design of the experiment, the effect of NH ΔSNR_{90} on HI perception was unknown and this measure was allowed to vary from −9 to +10 [dB].

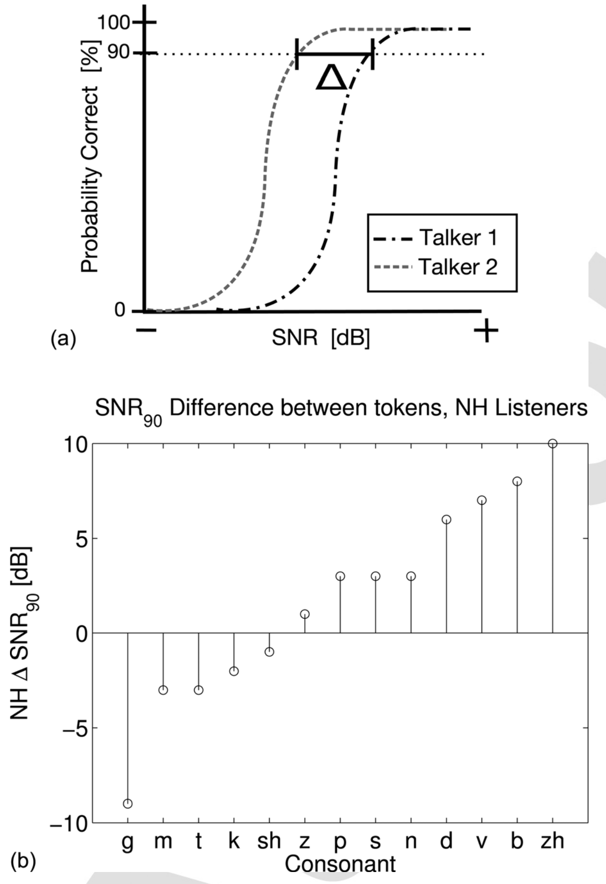


FIG. 1. (a) Illustration of Probability vs SNR curves for two tokens with the difference in SNR_{90} values (ΔSNR_{90}) indicated. The SNR_{90} is defined as the SNR at which the probability of recognition drops to 90%, while ΔSNR_{90} quantifies the difference in noise-robustness across tokens. (b) The NH ΔSNR_{90} values for each set of consonant tokens in this study, as computed from NH perceptual data in the presence of speech-weighted noise (Table IV, value for /f/ not shown). These values are computed as in the example of (a) with the male token as talker 1 and the female token as talker 2. For each consonant, a positive NH ΔSNR_{90} indicates that the female token is more robust to noise, while a negative value indicates that the male token is more robust to noise. The consonants are sorted along the abscissa by NH ΔSNR_{90}. The labels sh = ʃ and zh = ʒ.

F. Hearing impaired data analysis

For each ear, the traditional metric of average consonant error at a particular SNR, $\overline{P_e}(s)$, is computed as

$$\overline{P_e}(s) = \frac{1}{28}\sum_{i=1}^{14}\sum_{j=1}^{2} P_e(C_i, T_j, s), \qquad (2)$$

where $P_e(C_i, T_j, s)$ is the probability of error for the ith consonant C_i, jth talker T_j, at SNR s. The average is computed over all 28 tokens used in this study (14 consonants × 2 talkers = 28 tokens).

For a given consonant, the average of the token error difference, $\overline{\Delta P_e}$, is formulated as

$$\overline{\Delta P_e} = \frac{1}{n(S)}\sum_{s\in S}\left(P_e^M(s) - P_e^F(s)\right), \qquad (3)$$

$$S = \{s \in (0, 6, 12, \text{quiet}) : s \le s^*\},$$

where 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 (i.e., noise conditions) in set S. In this analysis, the probability of error for the male token $P_e^M(s)$ is always subtracted from that of the female token $P_e^F(s)$. $\overline{\Delta P_e}$ for each consonant is only computed over the SNRs below which an error is observed for at least one of the two tokens, to better separate tokens that show within-consonant differences. In the cases where no error is observed over all SNRs for both tokens, $\overline{\Delta P_e}$ is defined as zero ($\overline{\Delta P_e} \triangleq 0$).

III. RESULTS

A. Error overview

The average consonant error as a function of SNR, $\overline{P_e}(s)$, for the 17 HI ears in this study is shown in Fig. 2. The

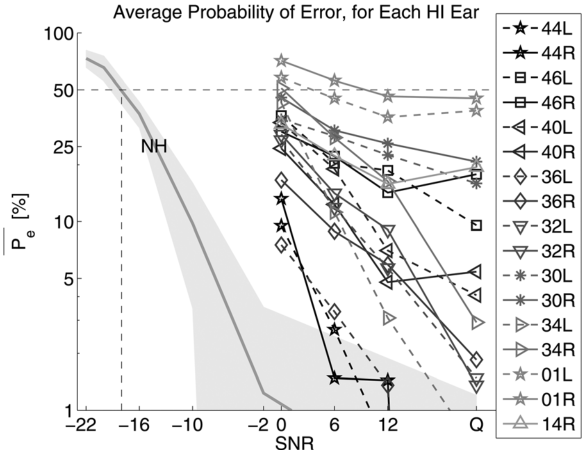


FIG. 2. Average probability of error (%) over all tested tokens for each HI ear, plotted as a function of SNR [Eq. (2)] on a log scale. Right ears (R) are shown as solid lines, left ears (L) as dashed lines. The average NH error (gray solid line) is included for reference along with a gray error region representing 1 standard deviation.

average error for 16 NH listeners, for the same set of test tokens, is overlaid in this figure for comparison. The average errors of four HI ears fall within the range of normal performance at low-noise levels (44L/R, 36L, 34L), and three HI ears reach 50% average error at 0 dB SNR (34R, 01L/R). Note that the $\overline{P_e}(s)$ for a HI ear is approximately linear on a log scale with respect to SNR, just like the error predicted by the articulation index formula (Allen, 1994).

As the inclusion of figures for all 17 individual HI ears is impractical, we examine the individual token errors for a set of representative ears, the left ears of listeners 40 and 34, in detail. Both ears have the same audiometric configuration as the majority of ears in our study, slight-to-mild low-frequency flat loss with high-frequency sloping loss (see Table I). In terms of average consonant error (Fig. 2) these two ears fall within the middle range of the tested HI ears.

An overview of the individual token errors for each of these two HI ears (40L and 34L) is presented in Fig. 3. Each plot shows the sorted error over all test tokens at each SNR. The tokens are sorted along the abscissa to create a monotonically increasing error distribution. This sorted distribution allows one to clearly visualize the proportion of tokens that contribute to the overall average error and the degree of error for each token. In the lower noise conditions, no error is observed for the majority of the tested tokens, while a small subset of the tokens can show high degrees of error. Such a concentration of error to only a few tokens is observed across all of the slight-to-mild HI ears in the study. For ear 40L [Fig. 3(a)], only three tokens show error at 12 dB SNR; at the worst noise condition, 0 dB SNR, 16 of 27 (≈59%) of the tokens have non-zero error. Ear 34L [Fig. 3(b)], also has a small subset of test tokens that account for all of the error at low-noise levels (6 and 12 dB SNR and quiet). Although a small number of tokens are incorrectly recognized at low-noise levels, a high degree of error can be associated with these tokens.

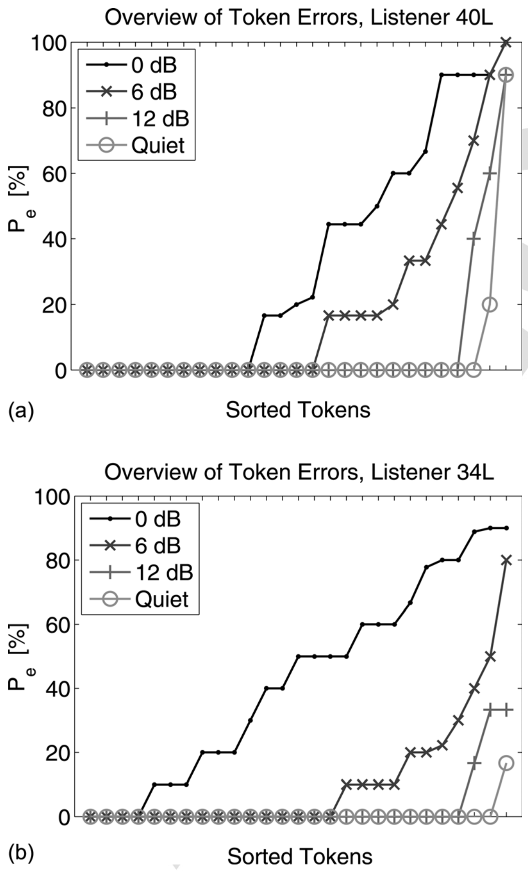


FIG. 3. Distribution of error for (a) ear 40L and (b) ear 34L at each of the four noise conditions. The abscissa corresponds to the 27 test tokens, sorted for each SNR such that the error increases monotonically; thus, the sort order can vary across ears and SNRs.

Cases such as these, where a small subset of tokens have high error while the remaining majority of tokens are recognized normally (i.e., without error), are misrepresented by a single overall average. In the following section, the variability of error across tokens of the same consonant is examined.

B. Within-consonant differences—robustness to noise

The noise robustness of a token is quantified by the threshold SNR at which significant errors are first observed. Here, we examine within-consonant differences in robustness to noise by analyzing the variability of error across tokens of the same consonant. The most extreme example of this token error difference for a HI ear is where one token of a consonant has no error at any tested SNR while the other token of the consonant reaches errors as high as 100%. As described in Sec. II, each token in the experiment was selected to be robust to at least −2 dB of noise for NH listeners (see Appendix). Thus for the HI ears, observations of zero error at the 0, 6, and 12 dB SNR and quiet conditions is equivalent to "normal" performance.

The consonant recognition error as a function of SNR for both talker tokens [$P_e^M(s)$ and $P_e^F(s)$] and the average across the two talkers is displayed in 14 sub-plots (one for each consonant) for ears 40 L and 34 L in Figs. 4(a) and 4(c), respectively. Ear 40 L reaches ≥50% two-talker average error for /b, g, m, n, v/, as noise is introduced; when the error is analyzed at the token level, one finds that the error for /g, m/ is completely due to the female token and that the error for /v/ is completely due to the male token. Ear 34L reaches ≥50% two-talker average error for /b, g, k, p, v, z/, as noise is introduced. The largest differences in noise robustness for ear 34L are for tokens of /k, m, s, v/. For this ear, almost all of the average error for /k, m, s/ can be attributed to errors with only the female token. For /v/, the male token is recognized with no error in only the quiet condition, while the female token is robust down to 6 dB SNR. Thus, for both ears 40L and 34L, one can observe large differences in the noise robustness of tokens of the same consonant. Although the acoustical differences across these tokens are small enough for them to be recognized as the same consonant by NH listeners, they are appreciable enough to make a difference in HI perception.

To quantify this observation, the token error difference is calculated as a function of SNR. These values are then used to compute the average of the token error difference, $\overline{\Delta P_e}$ [Eq. (3)], shown for ears 40L and 34L in Figs. 4(b) and 4(d). A negative $\overline{\Delta P_e}$ indicates that the male token is more robust to noise, while a positive value indicates that the female token of a consonant is more robust to noise.

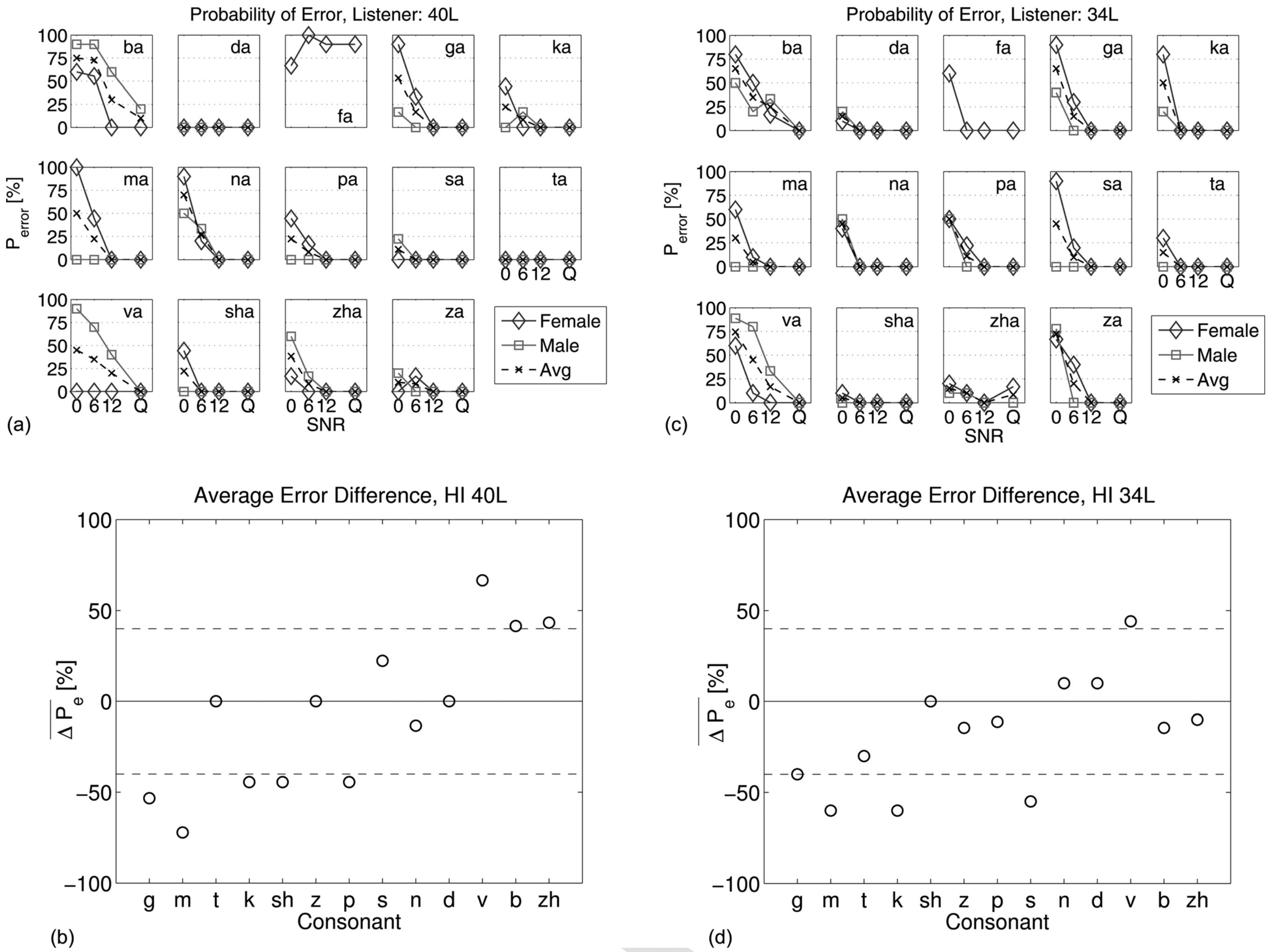


FIG. 4. Top left and right: Consonant recognition error as a function of SNR, for HI ears (a) 40L and (c) 34L. Each subplot shows the data for one consonant; plots display the error for the female token (diamond marker), male token (square marker), and the average across the two tokens (x marker, dashed line). Bottom left and right: $\overline{\Delta P_e}$ for each consonant [Eq. (3)], for HI ears (b) 40L and (d) 34L. Consonants are ordered along the abscissa based on the NH $\Delta\mathrm{SNR}_{90}$ values (as in Fig. 1). $\overline{\Delta P_e} = \pm 40\%$ is marked for reference. The labels sh = ʃ, zh = ʒ, and a = ɑ.

$\overline{\Delta P_e} = 40\%$ is marked for reference. The minimum number of experimental presentations for an token at a given SNR is $N = 5$, thus a 40% error difference corresponds to two trials in error, which is significantly different ($\alpha = 0.05$) from NH performance (Singh and Allen, 2012). The consonants with the largest average error differences for ear 40 L are /g, m, v/ and /m, k, s/ for ear 34 L. The consonants are ordered along the abscissa by the NH $\Delta\mathrm{SNR}_{90}$ values, as shown in Fig. 1(b). This is done to determine if the token of a consonant that is more robust to noise for a NH listener would also be more robust for a HI listener. Overall, there is some agreement as a rough increasing trend can be observed in Figs. 4(b) and 4(d).

The NH SNR_{90} has been found to significantly correlate with the intensity of the time-frequency region that contains the primary consonant cues (Régnier and Allen, 2008; Li *et al.*, 2010; Li *et al.*, 2012; Kapoor and Allen, 2012). Thus the NH $\Delta\mathrm{SNR}_{90}$ relates the difference in intensity of the NH consonant cue regions. If HI perception was completely dependent on audibility/intensity of the primary consonant cues that NH listeners use, then the tokens of a consonant that are more robust to noise for NH listeners (i.e., lower NH SNR_{90}s) would also be more robust to noise for HI listeners. The $\overline{\Delta P_e}$ values for all 17 HI ears are shown in Fig. 5(a); the consonants along the abscissa are in the same order as Fig. 1(b). Overall, large token error differences is a widespread effect, with 16 of 17 HI ears showing at least one average token error difference >40%. A clear increasing trend can be observed in the mean HI $\overline{\Delta P_e}$ values, similar to the trend of the NH $\Delta\mathrm{SNR}_{90}$ values. A linear regression between the two measures is plotted in Fig. 5(b); the HI $\overline{\Delta P_e}$ and NH $\Delta\mathrm{SNR}_{90}$ values are significantly correlated ($\rho = 0.81$, $p < 0.001$).

Despite this strong relationship, a notable amount of individual variability can be observed in the data of Fig. 5(a). Tokens that are almost identically noise robust for a NH listener can show large $\overline{\Delta P_e}$ values for a HI ear. As an example, the two tokens of /z, p, s/ have NH $\Delta\mathrm{SNR}_{90} \leq 3$ dB, indicating that the two tokens have necessary cue regions that are nearly equal in intensity. Yet there are individual HI ears for which a $\overline{\Delta P_e} > 50\%$ is observed for /z, p, s/. In such cases, additional signal properties, perhaps the presence of *conflicting cues* (Li *et al.*, 2010; Li *et al.*, 2012; Kapoor and Allen, 2012) or variations of the primary cues to which the HI ears could be sensitive, may play

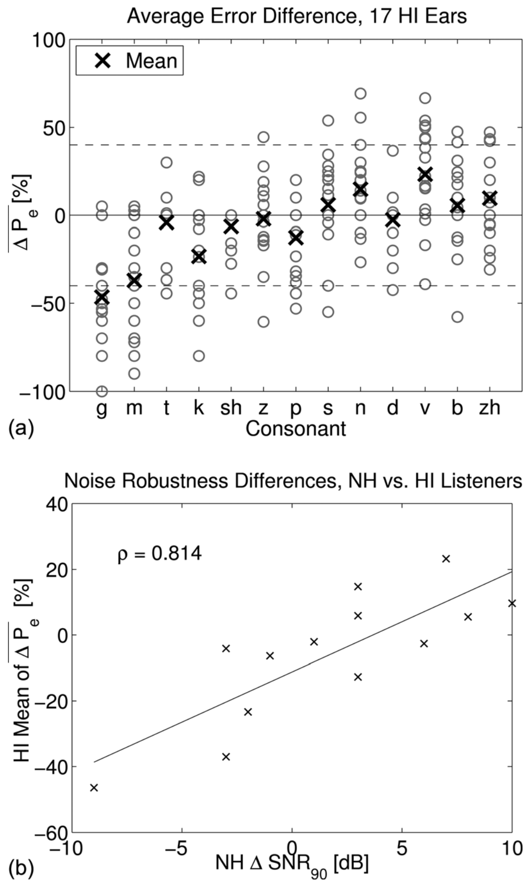


FIG. 5. (a) $\overline{\Delta P_e}$ for all HI ears. Each point represents the value for a single HI ear, the mean across ears for each consonant is marked with an 'x'. A negative $\overline{\Delta P_e}$ indicates that the male token has lower error, a positive value indicates that the female token has lower error. Consonants are ordered along the abscissa based on the NH $\Delta\mathrm{SNR}_{90}$ values (as in Fig. 1). $\overline{\Delta P_e} = 40\%$ is marked for reference. (b) Comparison and linear regression of the mean $\overline{\Delta P_e}$ values and the NH $\Delta\mathrm{SNR}_{90}$ values (see Fig. 1), the two values are significantly correlated ($\rho = 0.81$, $p < 0.001$). The labels sh = ʃ and zh = ʒ.

a role. To better understand the HI within-consonant differences, we next examine the consonant confusions.

C. Within-consonant differences—confusion groups

The common NH listener confusion groups for English consonants were established by Miller and Nicely (1955) (e.g., /b, d, g/, /p, t, k/, /m, n/). When analyzing HI speech perception, some of these same confusion groups are observed. In this section, we investigate the extent of within-consonant differences in terms of the confusion groups.

The confusions for each of the two tested /bɑ/ tokens are first analyzed in detail. The confusion matrices for these two /bɑ/ tokens are shown in Tables II(a) and II(b), for six HI ears (34L/R, 36L/R, 40L/R) at each of the four tested SNRs (0, 6, and 12 dB SNR and quiet). For the female /bɑ/ [Table II(a)], although the HI ears have different degrees of error at different SNRs, one can observe frequent /d, g, v/ confusions. For the male /bɑ/ [Table II(b)], the primary confusions are instead with /v, f/. Similar differences in confusion groups for the two /bɑ/ tokens are observed across all of the tested HI ears. The average responses over all 17 HI ears as a function of SNR are shown for the female and male /bɑ/ tokens in Tables II(c) and II(d).

The confusion matrices for all test tokens (averaged across all 17 HI ears and SNRs) are shown in Table III. Here we can again see the differences in confusion groups for the two /bɑ/ tokens, but we also observe within-consonant differences for the average confusion groups of /gɑ, mɑ, sɑ, ʒɑ/. Although some confusions are shared across multiple tokens of the same consonant, distinct within-consonant differences can be observed in the confusions.

The size of the confusion groups observed in the averages can be small, indicating, in those cases, that the majority of the responses across all HI ears and noise conditions are drawn from the same confusion group. These similar confusions across HI ears are observed despite the many subject differences, including degree of hearing loss, age, gender, and background. This consistency across HI ears implies that the acoustic properties of each token (i.e., variable primary and conflicting acoustic cues) are responsible for the HI confusions. When the confusion groups for multiple tokens of a consonant are different, as in the case of these two /bɑ/ tokens, averaging the data from these tokens causes HI listeners to appear more "random" (higher entropy) in their speech perception than they actually are.

D. Repeatability

A pilot experiment was conducted approximately a year before the main experiment reported on in this study (Han, 2011). This pilot experiment collected consonant recognition data from 46 HI ears, including 16 of the 17 HI ears in this study. The speech materials of the pilot experiment were also drawn from the LDC database with 16 consonants in a consonant-vowel context (/p, t, k, f, s, ʃ, b, d, g, v, z, ʒ, m, n, θ, ð/ + /ɑ/) and six tokens per consonant. Of the 28 tokens that are reported on in this study, 17 were also tested in the pilot experiment. Consonant recognition was tested at the same SNRs as in this study but with only two presentations at each SNR per token. Presentations were randomized over consonant and talker but not SNR. The pilot experiment was conducted with an identical setup (observers, graphical user interface, location) as the present study.

The data for tokens that are common with the pilot experiment can be used to provide a measure of the repeatability. The average error for 16 HI ears across the two experiments is significantly correlated ($\rho = 0.83$, $p < 0.001$), indicating reasonable test-retest reliability of this consonant recognition test.

IV. SUMMARY

HI ears can have large perceptual differences for tokens of the same consonant. Such differences are observed in both their robustness to noise and confusion groups.

Consistent differences in the noise-robustness of tokens of the same consonant are observed for the majority of the tested HI ears. These differences can be observed to the extreme that one token of a consonant has no errors at the worst noise condition of 0 dB SNR while the other token of the same consonant reaches 100% error at equal or better SNRs. The average

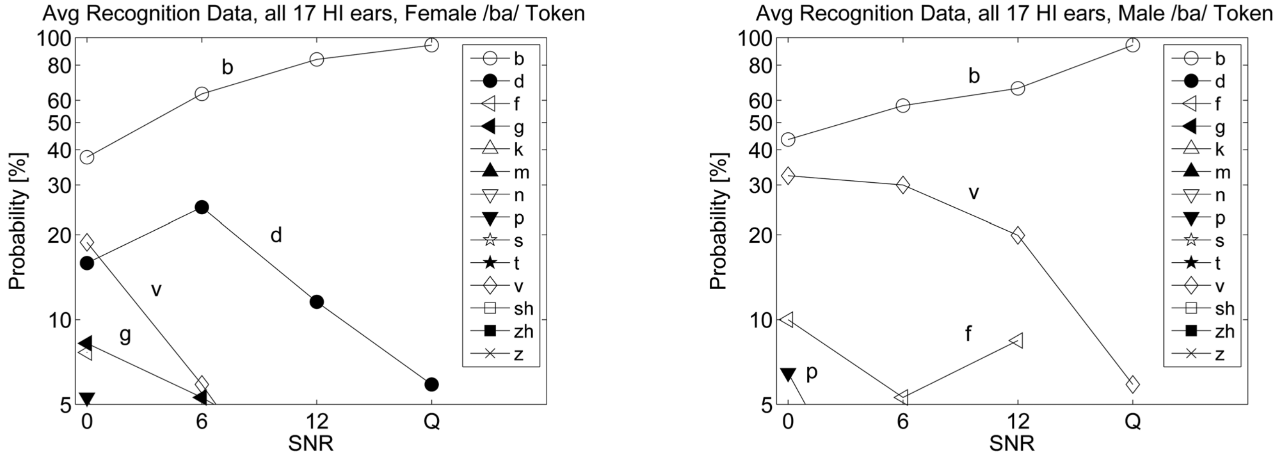
TABLE II. (a) Confusion matrix for the female /bɑ/ token, data from six HI ears (34L/R, 36L/R, 40L/R), at each SNR (dB). (b) Confusion matrix for the male /bɑ/ token, data from the same six HI ears (34L/R, 36L/R, 40L/R), at each SNR (dB). For both confusion matrices, the highest probability confusion in each row is highlighted in bold, and probabilities of 0% are removed to reduce clutter. (c) The recognition data for the female token, averaged across all 17 HI ears; primary confusions are with /d, v, g/. (d) The recognition data for the male token, averaged across all 17 HI ears; primary confusions are with /f, v/. The labels sh = ʃ, zh = ʒ, and a = ɑ.

Ear	SNR	b	d	f	g	p	v	s, k
34L	Q	100						
	12	83					**17**	
	6	50	20				**30**	
	0	20	**40**	20		10	10	
34R	Q	100						
	12	20	**40**		30		10	
	6	10	**80**				10	
	0		**40**	30	10		10	10
36L	Q	100						
	12	80	**20**					
	6	60	**30**			10		
	0	30	**40**		10	20		
36R	Q	50	**50**					
	12	30	**70**					
	6		**60**		20		20	
	0		**30**		20		**30**	20
40L	Q	100						
	12	100						
	6	44	**56**					
	0	40	10		10		**40**	
40R	Q	83	**17**					
	12	100						
	6	70	**30**					
	0	70					**20**	10

(a) Female bɑ token

Ear	SNR	b	d	f	g	p	v	k, t, z
34L	Q	100						
	12	66	**17**	**17**				
	6	80	**10**				**10**	
	0	50	10				**30**	10
34R	Q	100						
	12	83		**17**				
	6	70				10	**20**	
	0	50	10	10		10	**20**	
36L	Q	100						
	12	83			**17**			
	6	70			**30**			
	0	70			**20**		10	
36R	Q	100						
	12	30					**70**	
	6			10			**90**	
	0			30			**70**	
40L	Q	80					**20**	
	12	40					**60**	
	6	10		10			**60**	20
	0	10				30	**60**	
40R	Q	100						
	12	78					**22**	
	6	90					**10**	
	0	20					**80**	

(b) Male bɑ token



(c)

(d)

token error difference, $\overline{\Delta P_e}$ [Eq. (3)], can be used to quantify this difference in noise robustness. Comparing the $\overline{\Delta P_e}$ values for all HI ears [Fig. 5(a)] shows that across the HI ears one of the two tokens can be consistently more robust to noise than the other. Specifically, this figure shows that the male token of /g, m, k/ is consistently more robust to noise than the female one, and the female token of /n, v/ is consistently more robust to noise than the male one. This shows that a physical property of the signal makes one token more noise robust than the other for HI listeners. To investigate possible signal properties that

TABLE III. A confusion matrix showing the average response (%) for each token (average taken over the 17 HI ears and 4 SNRs). Each row contains data for a single token. Confusion probabilities >5% are highlighted in bold, and probabilities <2% are not shown. *F*, *M* subscripts denote tokens from female and male talkers.

	b	d	f	g	k	m	n	p	s	t	v	ʃ	ʒ	z
b_F	70	**15**	2	4							**7**			
b_M	65	2	**6**			2		2			**22**			
d_F		93								4				2
d_M		95								4				
f_F			73						**17**		3	3		2
g_F	3	**12**	**5**	62	2					2	**8**			2
g_M		**15**		83										
k_F					80			4		**13**				
k_M					87					**11**				
m_F						79	**9**	2			**7**			
m_M						93	**6**							
n_F						4	86				4			
n_M						**19**	80							
p_F			2		3			82		**12**				
p_M								92		3				
s_F	2		4						84				3	3
s_M									79				**8**	**12**
t_F								2	2	93				
t_M										96				
v_F	3	2	4			4	4				78			2
v_M			4	4		**5**	**5**	**11**		4	63			
ʃ_F									4			92		2
ʃ_M												96		2
ʒ_F				**6**								2	67	**24**
ʒ_M		3		**6**							**11**		63	**13**
z_F		4									**6**		**16**	70
z_M									4	2	2		**16**	74

could lead to such differences in noise robustness, we have considered a perceptual measure of the acoustic cue region intensity, the NH SNR$_{90}$.

For each token, the NH-listener necessary acoustic cue region can be isolated in time-frequency space with a combination of time-truncation and high-/low-pass filtering psychoacoustic experiments (Phatak and Allen, 2007; Li *et al.*, 2010; Li *et al.*, 2012). The NH SNR$_{90}$ has been found to significantly correlate with the intensity of this NH necessary cue region. Thus the difference in NH SNR$_{90}$ values can be used to relate the intensity of the NH necessary cue region across tokens. The NH ΔSNR_{90} values are compared to the means of the HI $\overline{\Delta P_e}$ values in Fig. 5(b). A significant correlation of $\rho = 0.81$ between the two measures demonstrates that the variable acoustic properties that make a token more robust to noise for NH listeners also, generally, affect perception for HI listeners.

Going beyond the error, an analysis of the confusion groups reveals additional within-consonant differences; we have found that tokens of the same consonant can have different confusion groups for HI listeners. We observe confusion group differences for the selected tokens of /b, g, m, s, ʒ/ across all of the HI ears in this study. When examined on a token (as opposed to a consonant) level, one observes that HI ears are much more consistent in their confusions.

V. CONCLUSIONS

In this study, we analyze HI consonant recognition using a low-context stimulus task with four speech-weighted noise conditions. The majority of HI ears have slight-to-moderate hearing loss with a high-frequency sloping audiometric configuration.

For each HI ear, fewer than half of all tested tokens show errors (Fig. 3) in the low-noise conditions. Despite a small number of ear-specific tokens that are in error, the degree of error for these tokens can be large ($\geq 80\%$). The average error as a function of SNR, $\overline{P_e}(s)$, is insensitive to large degrees of error for only a small subset of the test tokens.

NH-listener data from psychoacoustic tests (e.g., masking, filtering, time-truncation) can be used to characterize naturally variable consonant tokens. Generally, filtering data can be used to identify the necessary frequency range, time-truncation data can identify acoustic components that are necessary for temporal/durational cues, and the resulting threshold SNR$_{90}$ from a noise-masking experiment is correlated to the intensity of the necessary acoustic cue region. In addition, the acoustic elements that encode conflicting cues (for non-target consonants) can be identified with the same filtering and time-truncation experiments. In general, NH-listener psychoacoustic data can be used to characterize the perceptually relevant information of variable acoustic cues (e.g., the necessary frequency range for correct perception) and test for their effect on HI perception. In this article, we use the characterization provided by the NH-listener noise-masking data to explore the role of cue region intensity in HI perception.

For NH listeners, the noise robustness of a sound is correlated to the intensity of the acoustic components within the necessary cue region. We find that the within-consonant differences in noise robustness for HI ears are correlated to the noise robustness of consonants for NH listeners (Fig. 5). This supports the hypothesis that the acoustic cues that are necessary for NH listeners are also necessary for the HI listeners, although they may not be sufficient. Thus, just as selective amplification of the NH cue region can manipulate the noise robustness of tokens for NH listeners (Kapoor and Allen, 2012), similar selective amplification might make a token more noise robust for HI listeners. For cases where the relative noise robustness of tokens for NH and HI listeners are inconsistent, other signal properties besides the intensity of acoustic cues (e.g. within-consonant variability of the primary cues or the presence of conflicting cues) must play a role.

Within-consonant differences in confusion groups are observed. When the HI ears make an error, they collectively draw from a limited token-dependent confusion group (Tables II and III). Despite the many differences across HI ears (hearing loss, age, gender), the token-specific confusion groups are observed consistently. These consistencies over different HI ears require that the acoustic properties of each token define the possible confusions; this also implies that these HI ears, despite their many differences, use similar cues when making confusions. If each HI ear used different cues or interpreted the cues in an ear-dependent way, then such consistencies in the confusions across ears would not be observed. When, due to a hearing impairment, the primary cues are distorted or missing, remaining conflicting cues may be a source of the consistent token-specific confusions. Further analysis of the acoustic cues that lead to particular confusions has the potential to provide increased insight into the speech perception strategies that are being used by HI listeners.

Within-consonant perceptual differences for HI listeners are observed for sounds that are noise robust and unambiguous for NH listeners. Although the tokens are identified as the same consonant by NH listeners, subtle natural variations in signal properties can lead to systematic differences in HI perception. Averaging different token-specific confusion groups of a consonant can cause a HI listener to appear more random in their responses than they really are. In terms of entropy, averaging recognition data for multiple tokens with identical amounts of error but different confusion groups will produce higher-entropy results than would be obtained if calculated for the individual tokens.

The results suggest that when a HI listener reports that they can "hear speech but have trouble understanding it," it may be due to consistent errors with only a subset of phonemes. Multiple tokens of a single consonant have naturally variable cues, leading to varying measures of hearing impairment. These natural variations in signal properties may also affect NH consonant recognition when the speech signal is degraded (e.g., noisy, filtered). Characterizing the primary and any conflicting perceptual cues of test tokens is thus critically important to the design and interpretation of HI speech tests.

ACKNOWLEDGMENTS

This work would not have been possible without the dedication of Professor Woojae Han, who collected the hearing impaired data and conducted the first analysis, as partial fulfillment of her Ph.D. (Han, 2011). 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 009277 A and a generous grant from Research in Motion (RIM).

APPENDIX: TEST TOKENS

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

TABLE IV. For each consonant-vowel token (CV), the male (M) and female (F) talker labels are listed, along with the corresponding NH SNR$_{90}$ values (dB). The /fɑ/ from talker m112 is marked with a * to indicate that this token was not included in the data analysis.

CV	M Talker	SNR$_{90}$	F Talker	SNR$_{90}$	CV	M Talker	SNR$_{90}$	F Talker	SNR$_{90}$
bɑ	m112	−2	f101	−10	pɑ	m118	−14	f103	−17
dɑ	m118	−7	f105	−13	sɑ	m120	−10	f103	−13
fɑ	m112*	−5*	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

Allen, J. B. (**1994**). "How do humans process and recognize speech?" IEEE Trans. Speech Audio Process. **2**, 567–577.

Baum, S. R., and Blumstein, S. E. (**1987**). "Preliminary observations on the use of duration as a cue to syllable-initial fricative consonant voicing in English," J. Acoust. Soc. Am. **82**, 1073–1077.

Bilger, R. C., and Wang, M. D. (**1976**). "Consonant confusions in patients with sensorineural hearing loss," J. Speech Hear. Res. **19**, 718.

Boothroyd, A. (**1984**). "Auditory perception of speech contrasts by subjects with sensorineural hearing loss," J. Speech Hear. Res. **27**, 134.

Boothroyd, A., and Nittrouer, S. (**1988**). "Mathematical treatment of context effects in phoneme and word recognition," J. Acoust. Soc. Am. **84**, 101–114.

Bronkhorst, A., Bosman, A., and Smoorenburg, G. (**1993**). "A model for context effects in speech recognition," J. Acoust. Soc. Am. **93**, 499–509.

Bronkhorst, A. W., Brand, T., and Wagener, K. (**2002**). "Evaluation of context effects in sentence recognition," J. Acoust. Soc. Am. **111**, 2874–2886.

Burkle, T. Z., Kewley-Port, D., Humes, L., and Lee, J. H. (**2004**). "Contribution of consonant versus vowel information to sentence intelligibility by normal and hearing-impaired listeners," J. Acoust. Soc. Am. **115**, 2601.

Dillon, H. (**2001**). "Prescribing hearing aid amplification," in *Hearing Aids* (Thieme Medical Publishers, New York), Chap. 10.

Dobie, R. A. (**2011**). "The AMA method of estimation of hearing disability: A validation study," Ear Hear. **32**, 732–740.

Dorman, M. F., Studdert-Kennedy, M., and Raphael, L. J. (**1977**). "Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent, context–dependent cues," Atten. Percept. Psychophys. **22**, 109–122.

Dreschler, W. (**1986**). "Phonemic confusions in quiet and noise for the hearing-impaired," Int. J. Audiol. **25**, 19–28.

Dubno, J. R., and Dirks, D. D. (**1982**). "Evaluation of hearing-impaired listeners using a nonsense-syllable test. I. Test reliability," J. Speech Hear. Res. **25**, 135.

Fabry, D. A., and Van Tasell, D. J. (**1986**). "Masked and filtered simulation of hearing loss: Effects on consonant recognition," J. Speech Hear. Res. **29**, 170.

Fousek, P., Grezl, F., Hermansky, H., and Svojanovsky, P. (**2004**). "New nonsense syllables database-analyses and preliminary asr experiments," in *Proceedings of International Conference on Spoken Language Processing (ICSLP)*, 2004–29.

Gordon-Salant, S. (**1987**). "Consonant recognition and confusion patterns among elderly hearing-impaired subjects," Ear Hear. **8**, 270.

Halpin, C., and Rauch, S. D. (**2009**). "Clinical implications of a damaged cochlea: Pure tone thresholds vs information-carrying capacity," Otolaryngol. Head Neck Surg. **140**, 473–476.

Han, W. (**2011**). "Methods for robust characterization of consonant perception in hearing-impaired listeners," Ph.D. thesis, University of Illinois, Urbana-Champaign.

Herd, W., Jongman, A., and Sereno, J. (**2010**). "An acoustic and perceptual analysis of /t/ and /d/ flaps in American English," J. Phonetics **38**, 504–516.

Hood, J. D., and Poole, J. P. (**1977**). "Improving the reliability of speech audiometry," Br. J. Audiol. **11**, 93–102.

Humes, L. E. (**1991**). "Understanding the speech-understanding problems of the hearing impaired," J. Am. Acad. Audiol. **2**, 59–69.

Jongman, A., Wayland, R., and Wong, S. (**2000**). "Acoustic characteristics of English fricatives," J. Acoust. Soc. Am. **108**, 1252–1263.

Kamm, C. A., Dirks, D. D., and Bell, T. S. (**1985**). "Speech recognition and the articulation index for normal and hearing-impaired listeners," J. Acoust. Soc. Am. **77**, 281–288.

Kapoor, A., and Allen, J. B. (**2012**). "Perceptual effects of plosive feature modification," J. Acoust. Soc. Am. **131**, 478–491.

Killion, M. C., and Gudmundsen, G. I. (**2005**). "Fitting hearing aids using clinical prefitting speech measures: An evidence-based review," J. Am. Acad. Audiol. **16**, 439–447.

Kurowski, K., and Blumstein, S. E. (**1987**). "Acoustic properties for place of articulation in nasal consonants," J. Acoust. Soc. Am. **81**, 1917.

Lawrence, D. L., and Byers, V. W. (**1969**). "Identification of voiceless fricatives by high frequency hearing impaired listeners," J. Speech, Lang. Hear. Res. **12**, 426.

Li, F. (**2011**). "Perceptual cues of consonant sounds and impact of sensorineural hearing loss on speech perception," Ph.D. thesis, University of Illinois, Urbana-Champaign.

Li, F., Menon, A., and Allen, J. B. (**2010**). "A psychoacoustic method to find the perceptual cues of stop consonants in natural speech," J. Acoust. Soc. Am. **127**, 2599–2610.

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," J. Acoust. Soc. Am. **132**, 2663–2675.

Miller, G. A., and Nicely, P. E. (**1955**). "An analysis of perceptual confusions among some English consonants," J. Acoust. Soc. Am. **27**, 338–352.

Mines, M. A., Hanson, B. F., and Shoup, J. E. (**1978**). "Frequency of occurrence of phonemes in conversational English," Lang. Speech **21**, 221–241.

Owens, E. (**1978**). "Consonant errors and remediation in sensorineural hearing loss," J. Speech Hear. Disord. **43**, 331.

Phatak, S. A., and Allen, J. B. (**2007**). "Consonant and vowel confusions in speech-weighted noise," J. Acoust. Soc. Am. **121**, 2312–2326.

Régnier, M. S., and Allen, J. B. (**2008**). "A method to identify noise-robust perceptual features: Application for consonant /t/," J. Acoust. Soc. Am. **123**, 2801–2814.

Roeser, R. J., Valente, M., and Hosford-Dunn, H. (**2007**). *Audiology: Diagnosis* (Thieme Medical Publishers, New York), p. 289.

Singh, R., and Allen, J. B. (**2012**). "The influence of stop consonants perceptual features on the articulation index model," J. Acoust. Soc. Am. **131**, 3051–3068.

Skinner, M. W. (**1976**). "Speech intelligibility in noise-induced hearing loss: Effects of high-frequency compensation," Program in Audiology and Communication Sciences, Washington University School of Medicine.

Skinner, M. W., and Miller, J. D. (**1983**). "Amplification bandwidth and intelligibility of speech in quiet and noise for listeners with sensorineural hearing loss," Int. J. Audiol. **22**, 253–279.

Smoorenburg, G. F. (**1992**). "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. **91**, 421–437.

Taylor, B. (**2006**). "Predicting real-world hearing aid benefit with speech audiometry: An evidence-based review," Ph.D. thesis, Central Michigan University.

Vysochanskij, D. F., and Petunin, Y. I. (**1980**). "Justification of the 3σ rule for unimodal distributions" Theory Probab. Math. Stat. **21**, 25–36.

Walden, B. E., Holum-Hardegen, L. L., Crowley, J. M., Schwartz, D. M., and Williams, D. L. (**1983**). "Test of the assumptions underlying comparative hearing aid evaluations," J. Speech Hear. Disord. **48**, 264.

Walden, B. A., and Montgomery, A. A. (**1975**). "Dimensions of consonant perception in normal and hearing-impaired listeners," J. Speech Hear. Res. **18**, 444.

Wang, M. D., Reed, C. M., and Bilger, R. C. (**1978**). "A comparison of the effects of filtering and sensorineural hearing loss on patterns of consonant confusions," J. Speech Hear. Res. **21**, 5.

Zurek, P. M., and Delhorne, L. A. (**1987**). "Consonant reception in noise by listeners with mild and moderate sensorineural hearing impairment," J. Acoust. Soc. Am. **82**, 1548–1559.