Comparisons in Consonant Confusions with and without gain for the Hearing-Impaired Listeners

Yang-soo Yoon^{1*}, David M. Gooler², Jont B. Allen³, Jae-sook Gho⁴

¹Department of Speech, Language, & Hearing Sciences, Texas Tech University Health Sciences Center, Lubbock; ²Department of Speech and Hearing Science, University of Illinois - Urbana/Champaign, Champaign; ³Department of Electrical and Computer Engineering, University of Illinois - Urbana/Champaign, Urbana; ⁴Center for Counseling and Family Studies, Liberty University, Virginia

Purpose: The present study aimed determining the effect of audibility on a consonant-by-consonant perception and on perceptual confusions per hearing-impaired (HI) listener.

Methods: Six participants with sensorineural hearing loss participated. Sixteen consonant-vowel (CV) syllables with the common vowel /a/ were presented as a function of signal-to-noise ratio. Gains were computed with subject's hearing thresholds by using National Acoustics' Laboratory - Revised Compensation Rule. Then the gains were applied to each of the 16 CV syllables. Consonant confusions were measured without and with gain.

Results: We identified three levels of difficulty in CV perception regardless of applying gain: Easy- /ga/, /ka/, /ma/, /na/, /pa/, /ja/; Moderate - /da/, /fa/, /sa/, /ta/, /za/, /la/; Difficult- /ba/, /va/, /ða/, and /θa/. Enhanced audibility improved performance most for Moderate set and little for the Easy set, but created negative effect on performance for the Difficult set. The effect of gain is also listener-specific: three out of the six listeners received benefit while other three listeners experienced negative effect from CVs with gain. The confusion analysis showed that subjects who benefited had the same primary competitors between gain and no-gain conditions, while subjects who did not benefit had additional competitors with gain.

Conclusions: The preliminary results of this study indicate that audibility is one of the primary factors influencing speech recognition of HI listeners, but reduced audibility alone cannot explain the difficulty of HI listeners to understand speech-in-noise. A more affirmative conclusions can be made with further analyzing data from larger sample size.

Keywords: Audibility, Distortion, Consonant, Confusion patterns, Consonant loss profiles

Received: March 1, 2017

Revision: April 28, 2017

Accepted: April 28, 2017

Correspondence:

Yang-soo Yoon

Department of Speech, Language, and Hearing Sciences, Texas Tech University Health Sciences Center, 3601 4th St STOP 6073, Lubbock, Texas 79430, USA
Tel: +806-743-9048
Fax: +806-743-5670
E-mail: yang-soo.yoon@ttuhsc.edu

This study was not presented orally at any conferences.

This work was supported by Etymotic Research Foundation, Chicago IL

© 2017 The Korean Association of Speech-Language Pathologists

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Phatak et al. [1] and Han et al. [2] have shown that performance of consonant perception in hearing-impaired (HI) listeners is highly variable across listeners and dependent on evaluation on a consonant-by-consonant basis. Phatak et al. [1] demonstrated consonant-specificity using a consonant loss profile (CLP), which compares the difference in signal-to-noise ratio (SNR) required for the same performance between normal-hearing (NH) and HI listeners. The CLP analyses showed that a certain set of consonants (/ba/, /va/, /ða/, /θa/) elicited the poorest performance regardless of the degree of hearing loss. The CLPs of HI listeners for this set of consonants revealed that greater than approximately 20 dB SNR was required for equivalent performance with

NH listeners across SNRs. Such poor performance leads to a hypothesis that the articulation index is far less than 1 over speech frequency bands. That is, the speech spectrum is inaudible due to the higher thresholds of HI listeners. Han et al. [2] compared consonant perception error rates per HI ear before and after applying the National Acoustics' Laboratory - Revised (NAL-R) Compensation Rule. Enhanced consonant recognition was generally observed with NAL-R gain, but perception of some consonants received little or no benefit from the gain correction. However, this study did not provide analyses regarding the nature of this listener, ear, and consonant-dependent performance. In the present study, we tested the hypothesis that audibility alone cannot account for this inability of HI listeners to process consonants by comparing CLPs and confusion patterns (CPs) before and after applying prescriptive spectral NAL-R gain appropriate to the audiometric thresholds of HI listeners.

Turner and Robb [3] reported that NH listeners showed excellent discrimination of synthetic stop consonants when the bulk of the consonant stimulus spectrum was presented at supra-threshold levels, whereas HI listeners always demonstrated poorer performance than NH listeners. These results suggest that audibility deficits alone could not account for the poorer recognition of stop consonants by HI listeners. Dubno et al. [4] assessed the contribution of certain frequency regions to synthetic stop consonant perception with 18 NH listeners and a homogeneous group of 10 HI listeners with high-frequency hearing loss. They characterized differences in stop-consonant place perception among these listeners as a function of presentation level and high- and low-pass filtering. Differential filtering effects on consonant place perception were consistent with the spectral composition of acoustic cues. Differences in consonant recognition and error patterns between NH and HI listeners were not observed when the stimulus bandwidth presented to NH listeners included regions of threshold elevation typical of the HI listeners. That is, the reduction in audibility for NH listeners provided by fixed-frequency, low-pass filtering did not appropriately model changes in recognition resulting from high-frequency hearing loss.

In contrast, some investigators have reported a clear benefit of increased audibility for speech perception by HI listeners. Turner and Cummings [5] provided frequency gain (lower and higher frequency) to 10 HI listeners with various degrees of high-frequency hearing loss. They reported that adding gain to high-frequency regions (≥ 3 kHz) where the hearing

threshold was higher than 55 dB HL tended to provide little or no benefit in performance. In contrast, adding gain to lower frequency regions provided a benefit of increased audibility even for listeners with thresholds greater than 55 dB HL, including a listener with a flat, severe-to-profound hearing loss. Turner and Brus [6] measured the perception of HI listeners for nonsense syllables that were low-pass filtered as a function of frequency gain and presentation level. They reported that each HI listener benefited from lower frequency gain (< 0.9 kHz) irrespective of degree of hearing loss less than 70 dB HL. In addition, Simpson et al. [7] reported that the mean consonant scores for HI listeners improved significantly with increasing audibility of high-frequency components of the speech signal.

Other investigators have measured the effect of audibility in indirect ways. Hearing loss was simulated in NH listeners by raising the tone-detection threshold using frequency-shaped noise maskers and the recognition of nonsense syllables was measured as a function of SNR and overall presentation level [8-10]. Humes and Roberts [9] reported that threshold elevation was a primary determinant of speech recognition, emphasizing the importance of audibility for speech understanding in quiet for elderly HI listeners. Zurek et al. [9,10] measured the recognition of nonsense syllables in noise with both flat and high frequency gain. Their 18 HI listeners were divided into five subgroups on the basis of the configuration of their hearing loss (mild shallow-rising, mild steep-rising, mild falling, moderate rising, and moderate flat). They found that performance on nonsense consonant recognition and feature perception between HI listeners and noise-masked NH listeners was comparable. The strength of this conclusion is unclear because important perceptual differences such as performance per consonant and per HI listener at specific SNRs were obscured by averaging across performance or listener characteristics. For example, differences in performance of 10-30% on acoustic cues such as affrication, voicing, and place between HI and NH listeners were not taken into account in their conclusions [10]. Humes et al. [11] also found the performance of the HI listeners with simulated hearing loss was never worse than that of the NH listeners with simulated hearing loss. The result indicates that secondary processing deficits were not considered in the hearing loss simulation with NH listeners.

In summary, audibility is the primary factor influencing speech recognition of HI listeners [5-10], but reduced audibility alone cannot explain the difficulty of HI listeners to under-

stand speech [3,4]. Three limitations observed from the studies discussed above were that: (1) speech performance was measured in quiet, at a limited range of SNRs or at fixed SNRs, (2) the use of pooled confusion matrices (CMs) over SNRs and over NH and HI listener groups for feature perception analysis makes interpreting data difficult, and (3) the accountability for the decrease or decrement in consonant perception with NAL-R gain was not provided. In the present study, we determined that audibility alone cannot account for the inability of HI listeners to process each of the given consonants. Perceptual performance was assessed by comparing CLPs and CPs for each HI listener before and after applying prescriptive spectral NAL-R gain appropriate to the audiometric thresholds of HI listeners.

METHODS

Participants

The six paid participants with sensorineural hearing loss, were native speakers of American English, and were between 39 and 62 years old. Clinic patients (University of Illinois at Urbana/Champaign Audiology Clinic, Champaign, IL), who had given permission to be contacted for participation in research studies, were recruited on the basis of screening preexisting audiograms. Only those with a three-frequency pure-tone average (PTA: 0.5 kHz, 1 kHz, and 2 kHz) between 30 dB HL to

70 dB HL were invited to participate in the study. Participants were excluded when hearing thresholds exceeded 70 dB HL at $f > 2$ kHz as preliminary results demonstrated that percent error P_e (SNR) was $> 50\%$ on two practice blocks (120 trials/block) of the consonant recognition task in quiet. The pure tone audiometric thresholds with and without gain for all listeners are shown in Figure 1. All procedures involving human subjects were approved by the University of Illinois IRB.

Stimuli

Stimuli were composed of sixteen nonsense consonant-vowel (CV) syllables of 16 American English consonants - (/g/, /k/, /m/, /n/, /p/, /f/, /d/, /f/, /s/, /t/, /ʒ/, /z/, /b/, /v/, /ð/, and /θ/) followed by the common vowel /a/ as in “father” (Table 1) and were obtained from the Linguistic Data Consortium at the University of Pennsylvania, [12]. Nonsense syllables were used to control for speech context effects. Six stop consonants were spoken by five talkers and the remaining by another five talkers, resulting in 80 tokens ($5 \text{ talkers} \times 16 \text{ CVs}$) total. As acoustic analysis and stability of these tokens were carefully evaluated by Phatak and Allen [13], each token was level-normalized before presentation using VU-METER software [14]. The purpose of dividing syllables among talkers was to create a diversity of talkers and simultaneously shorten experiment time. The use of multiple utterances from several speakers also offers assurance about the generality of the analyses be-

Figure 1. Dotted lines represent pure-tone audiometric thresholds (i.e., no gain condition) for the six participants. Solid lines represent theoretical hearing thresholds as adjusted by the NAL-R gain procedure. The listener's ID with gender and age is also given in parentheses.

Table 1. Symbols for each CV per error set

	Easy (E) Set						Moderate (M) Set						Difficult (D) Set			
CV	ga	ka	ma	na	pa	ʃa	da	fa	sa	ta	ʒa	za	ba	va	ʂa	θa
Symbol	●	■	◆	▲	▶	◀	○	□	◇	△	▷	◁	×	+	☆	*

yond the experimental stimuli.

The CVs were presented in speech-weighted noise (SWN) as a function of SNR (-12 dB, -6 dB, 0 dB, 6 dB, 12 dB, and in Quiet {Q}). The SNR was produced by maintaining the CVs at a constant level and adjusting the level of the masking noise, as described by Phatak and Allen [13]. Stimuli were computer-controlled and generated (Mobile-Pre, M-Audio) and presented monaurally via an Etymotic™ ER-2 insert earphone. Sound levels were controlled by an attenuator and headphone buffer (TDT™ system III) so that stimuli were presented at the most-comfortable-listening level for each listener.

Procedures

Each of the procedures addressed below was applied to both the no-gain and gain conditions. Participants read and signed the consent document to indicate informed consent to participate in the study. Before proceeding with the experimental tasks, the status of the ear canal and tympanic membrane were assessed using otoscopy and pure tone audiometry was performed to determine hearing thresholds and type of hearing loss for each listener. Each participant was seated in a sound-treated room (Industrial Acoustics Company) for audiometry, practice, and experimental sessions.

For both no-gain and gain conditions, stimuli were presented in a 16-alternative, forced-choice format with graphical user interface that listed the 16 CVs alphabetically. Participants were asked to identify the perceived CV from the list on the computer screen and click the icon using a mouse. A calibration button was inserted on the graphical user interface to allow for the presentation level (most-comfortable-listening level) to be determined by the subject's response to playing 10 CV syllables in quiet. In addition, pause and repeat buttons were available to allow for listener controlled rate of presentation between stimuli and repetition of stimuli (the number of repetitions was not limited).

Participants first performed a 30-minute, two-block practice session on CV identification with feedback for both no-gain and gain conditions. Stimuli were low-pass filtered (6.5 kHz cutoff frequency) and presented to the test ear in quiet.

Environmental sound to the other ear was attenuated using a foam earplug.

For both no-gain and gain conditions, the consonant identification test was administered to measure confusion matrices for CVs in SWN as a function of SNR. For each presentation, a CV and SNR were selected, and presentation was randomized from the array of 16 CVs and 6 SNR indices (including Q). SNR was selected randomly without replacement so that no SNRs were repeated before a sequence of the six levels was presented. Each individual stimulus (16 CVs × 5 talkers × 6 SNRs = 480) was repeated six times across the five talkers yielding 30 repetitions of each CV × SNR stimulus.

For both conditions, direct feedback was not provided for each CV presented. Percent correct for each block was provided on the participant's screen at the end of each block. The total number of trials and CVs already played was also provided on the participant's screen. In case the participant accidentally quit the program, all variables were saved automatically so the correct experimental conditions could be retrieved safely when the same participant logged in again. Data acquisition and processing was performed in MATLAB® (The MathWorks, Natick, MA) [15] on a Windows-based personal computer.

For each of no-gain and gain conditions, there were 480 trials in total (16 CVs × 5 talkers × 6 SNRs), evenly distributed into four blocks. The entire set was repeated six times (480 trials × 6 = 2,880 trials) in order to obtain a large enough sample size for each CV at each SNR. The row sum (30) in the confusion matrices was the total number of each CV presented at each SNR. Confusion matrices were regrouped as a function of SNR.

National Acoustics' Laboratory – Revised Compensation Rule (NAL-R)

The NAL-R procedure was considered a suitable basis for providing gain for HI listeners in the current study because it was derived from [16-18], and validated [19] and supported [20] by research and was widely accepted and used [21]. Furthermore, the main principles of NAL-R are still applied to fitting formulas used clinically (i.e., NAL-NL1).

The formula of the NAL-R procedure computes compensation gain as follows:

$$\text{Gain} = 0.05 * X * F + 0.31 * L + [-17, -8, -3, 1, 1, -1, -2, -2]$$

where,

- Hearing threshold level (L) in dB HL is taken from the HI listener's audiogram for frequencies $F = [0.25, 0.5, 1, 1.5, 2, 3, 4, 6]$ in kHz.
- X is the sum of the three hearing thresholds for $F = [.5, 1, 2]$ in kHz.
- A compensation factor is added on a frequency-specific basis in the order of the frequencies, F, listed above.

RESULTS

Mean error for three CV sets

To better understand the effects of gain on CV recognition performance, the error scores of three CV sets per listener for both gain and no-gain conditions were evaluated as shown in Figure 2. Gain condition is denoted by solid lines with filled symbols while no-gain condition is denoted by dotted lines with open symbols. The three CV sets were divided based on level of difficulty for perception: Difficulty (D) set - /ba/, /va/,

/δa/, /θa/, denoted by circles, Moderate (M) - /da/, /fa/, /sa/, /ta/, /za/, /za/, denoted by squares, and Easy (E) - /ga/, /ka/, /ma/, /na/, /pa/, /ja/, denoted by diamonds. Two general observations from Figure 2 were that: 1) the three CV sets were consistently observed with gain applied for all listeners and 2) addition of gain did not uniformly benefit performance for all CV error sets. The top panels in Figure 2 show the performance of listeners with negative effects (greater errors) of gain while the bottom panels show the performance of listeners with positive effects (fewer errors) of gain. The general pattern in the top panels was that the negative effect of gain was consistent over the CV sets and SNRs. In contrast, the general pattern in the bottom panels of Figure 2 was that the positive effects of gain were greatest for performance on the M set but the least on the D set.).

Consonant loss profiles (CLPs)

One of our primary goals was to determine the effect of audibility on consonant perception on a CV-by-CV basis. The next logical step was to determine how gain influenced recognition of syllables in each CV set for each listener. To answer this question, the CLPs, the difference in SNR required for the

Figure 2. Average error scores of the three CV error sets per listener. The error curves of performance are displayed for gain condition with solid line and for no-gain condition with dotted lines. The difficult (D) set (/ba/, /va/, /δa/, /θa/) is denoted by a circle symbol, the moderate (M) set (/da/, /fa/, /sa/, /ta/, /za/, /za/) as a square symbol, and the easy (E) set (/ga/, /ka/, /ma/, /na/, /pa/, /ja/) as a diamond symbol. The level of chance error is also shown as a dotted line at the top of each panel. Note that error scores in Q (quiet) were plotted at 18 dB SNR.

same performance between the gain condition and no gain condition, were plotted in Figures 3-5.

Figure 3 shows the CLPs for the D set. Listener S1 showed a negative effect of gain on recognition of / δ a/ and /ba/, as did

Figure 3. Consonant Loss Profiles for the D set (/ba/, /va/, / δ a/, / θ a/) between no-gain and gain per listener. Note that the 45-degree solid line indicates equal performance between experiments at a given SNR.

Figure 4. Consonant loss profiles for the M set (/da/, /fa/, /sa/, /ta/, /za/, /za/) between no-gain and gain per listener.

Figure 5. Consonant loss profiles for the E set (/ga/, /ka/, /ma/, /na/, /pa/, /fa/) between no-gain and gain per listener.

listener S2 on recognition of /va/ and /θa/. Listener S3 showed no clear effect of gain. Listener S4 experienced a negative effect of gain on recognition of /ða/ and /θa/, and listener S5 showed a positive effect of gain on recognition of /ba/ and /va/. Listener S6 showed a negative effect of gain on recognition of /ba/ and /ða/ but positive effect on recognition of /va/. The patterns of the CLP skewed toward the right, meaning that the effect of gain became positive as SNR increased. Benefits of all three listeners from gain were highly sound- and listener-specific. The performance of listeners who demonstrated negative effects of gain on certain CVs did not have globally negative effects of gain for all CVs. Overall, this result indicates that (1) audibility is not a major determinant for discriminating the D set and (2) gain can produce additional distortion, resulting in poorer performance due to higher entropy or uncertainty for CV selection.

Figure 4 shows CLPs for the M set. Listener S1 had a CLP close to the diagonal, meaning that perception of the M set was less dependent on audibility. For S2, the performance for three sounds /fa/, /da/, and /za/ decreased with gain, whereas that improved for /za/. The CLPs for two syllables, /sa/ and /ta/, were skewed toward the right below the diagonal, meaning that gain helped discrimination of these sound at SNRs > 0dB. For S3, only /sa/ was below the diagonal line

(benefit from gain), whereas all others were above the diagonal. The CLPs for S4 showed that curves for five syllables except the fricative /fa/ were just below the diagonal and showed benefit from gain; the perception of /sa/, /ta/, /za/, and /za/ largely improved with increasing SNR. Similarly, five sounds were below the diagonal except /fa/ for S5. The recognition of /sa/, /ta/, and /za/ mainly improved, but the recognition of /fa/ was reduced. The patterns of the CLPs for S6 were different from the other two listeners. The curves markedly shifted to the right of the diagonal as SNR increased for all sounds except /fa/, implying that perceptual cues become audible after approximately 6 dB SNR. Recognition of the fricative /fa/ was often affected negatively by gain for all three listeners. This result suggests that the perceptual correlate for /fa/ was altered by gain, resulting in poorer perception. In summary, the negative effect of gain on the M set was evident for S2 and S3, whereas gain had a positive effect on syllable recognition for S4 and S5. Performance of listener S1 was less dependent on gain, and S6 was sensitive to noise. Recognition of the fricative /fa/ was negatively influenced by gain even for listeners who showed improved performance with gain for most syllables.

Figure 5 shows CLPs for the E set. As represented in Figure 2, positive or negative effects of gain were mild for all listeners. A

negative effect of gain on recognition for all six syllables was seen for S2 and S3, whereas a positive effect of gain was evident for S5. The CLPs for the other three listeners remained adjacent to the diagonal, indicating that the perception of the E set was less dependent on gain. This analysis suggests that either set E was less dependent on gain or HI listeners reached the limits of their abilities to extract perceptual cues for the set.

Confusion patterns (CPs)

In the previous section, we determined the degree to which recognition of individual syllables showed benefit, no change, or deficit as a result of NAL-R gain. In this section, we attempted to answer the following question: does each syllable have the same or different perceptual confusions under no gain and gain conditions? For each listener, we selected *four consonants with highly contrasting CPs*. In Figures 6-8, the CPs are presented for the three listeners who demonstrated a negative effect of gain on syllable recognition (i.e., top panels in Figure 2). The CPs for the three listeners who demonstrated a positive effect of gain on syllable recognition (bottom panels

in Figure 2) are shown in Figures 9-11. In the CP figures, each target CV is identified above the panels. The no-gain condition is presented in the first row the gain condition in the second row. The percent correct for the target is displayed as a thick curve with symbols, while error probabilities for major competitors of the target consonant are represented by thin curves with labels. Only data points of CVs that competed with recognition of the target consonant above the chance level, denoted by the dotted horizontal line, are plotted.

Figure 6 shows the CPs for S1 for two D set sounds (/b/ and /δa/) and one from both the M set (/fa/) and the E set (/pa/). The perception of the target /ba/ did not improve at higher SNR for the gain condition (bottom panel), compared to the no-gain condition, as two sounds /va/ and /θa/ still accounted for above 40% of the error combined. Two syllables, /va/ and /fa/, competed against the target /ba/ under no-gain conditions but became distinct from the target at 6dB SNR. The percent recognition of the target /δa/ was below chance level in the gain condition, while the recognition of the target was far above chance, around 40% correct at SNR > 0dB. Two primary competitors, /ba/ and /θa/, were most often selected

Figure 6. Confusion patterns for four selected syllables measured from S1 listener. Each target is identified above the panels. The no-gain condition and gain conditions are presented in the first row and the second row, respectively. For all panels, the percent correct for the target is displayed as a thick curve with symbols, and error probabilities for major competitors are represented by thin curves with symbols and labels. Only data points of CVs that competed above the chance level, denoted by the dotted horizontal line, are plotted.

Figure 7. Confusion patterns of the four selected syllables for listener S2. Each target identified above the panels. The basic description is the same as described in Figure 6.

Figure 8. Confusion patterns of the four selected syllables for listener S3. Each target identified above the panels. Figure description is the same as described in Figure 6.

Figure 9. Confusion patterns of the four selected syllables for listener S4. Each target identified above the panels. Figure description is the same as described in Figure 6.

Figure 10. Confusion patterns of the four selected syllables for listener S5. Each target identified above the panels. Figure description is the same as described in Figure 6.

Figure 11. Confusion patterns of the four selected syllables for listener S6. Each target identified above the panels. Figure description is the same as described in Figure 6.

as alternatives under the no-gain condition, whereas multiple sounds such as /na/, /da/, and /ja/ competed under the gain condition. Providing audible speech (gain) introduced more confusion with new competing syllables for this target.

The recognition of the M set /fa/ was negatively influenced by gain. The primary competitor under no-gain conditions was /ba/, which showed a steady error of 20% across SNRs. Under the gain condition /ba/ continued to compete with the target, but accounted for an increasing percentage of error with SNR until error approached 40% in quiet. This increase in confusion was the main reason for the poor performance and the consequence of the gain provided. The pattern of performance curves for the E set /pa/ and a primary competitor /ta/ were similar under the no gain and gain conditions, but confusion from the competitor under the gain condition was not resolved to levels below chance with any SNR. Thus, the primary competitors were the same for the targets /ba/, /fa/, and /pa/, but the confusions from the competitors increased as SNR increased in the gain condition, suggesting that gain provided audible cues shared by the targets.

Figure 7 presents CPs obtained from listener S2 for two D set sounds (/va/ and /θa/) and two M set sounds (/fa/ and

/za/) to demonstrate a negative effect of gain. The target /va/ was clearly discriminated from competitors /ba/ and /fa/ at SNR=6 dB under the no-gain condition, but under the gain condition the same competitors contributed a steady error of about 35% above 6 dB SNR. For the target /θa/, the major competitor /fa/ produced similar levels of error for both conditions. Under the gain condition, two additional consonants, /da/ and /ða/, added to the confusions, resulting in extremely poor recognition of the target /θa/ even in quiet. The recognition score for target /fa/ increased with SNR and reached its asymptote (almost 100% correct) at 6 dB SNR in the no-gain condition, but the performance was below 50% in the gain condition because of confusions with /ba/ and /pa/. Similarly, the target /za/ had two competitors under the gain condition, resulting in roughly 30% error across SNR, whereas no confusions were found under the no-gain condition. Thus, the major competing sounds were the same, but the degree of confusions became stronger or new competitors were added for targets /va/ and /θa/ under gain conditions. For targets /fa/ and /za/, there were no major confusions under the no-gain condition, but major competitors /ba/, /sa/, and /za/ became prominent under the gain condition.

Figure 8 shows the CPs for three D set sounds (/ba/, /va/, and /ða/) and an M set sound (/fa/) for listener S3. For three targets /ba/, /ða/, and /fa/, the common characteristics of the confusions were that a primary competitor was the same for both conditions, but an additional competitor was added, increasing errors under the gain condition, and leading to poorer discrimination. For example, /ða/ was the primary competitor for the targets /ba/ and /fa/ under no gain conditions, but a new competitor was added for each (/fa/ for /ba/ and /pa/ for /fa/) under gain conditions. Similarly, the main competitor for the target /ða/ was /θa/ for both gain and no gain conditions, but two additional syllables /na/ and /va/ generated more errors under gain conditions. For the target /va/, the confusions from the common competitor /ða/ produced errors that remained above chance for both conditions, but the error produced by the competitor was higher under gain conditions. In the latter case, providing gain did not influence the internal structure of perceptual confusions, but yielded a greater degree of confusion with the primary competitor.

The CPs for the three listeners who demonstrated positive effects of gain on CV recognition in noise are shown in Figures 9-11. For listener S4, a D set sound (/ba/) and three M set sounds (/ta/, /za/, and /za/) are evaluated (Figure 9). The CPs for the target /ba/ demonstrates target perception with less dependence on gain. The poor recognition score of the target (except for in quiet) was almost identical for both conditions with the same primary competitor /fa/, suggesting that gain did not help the unique perceptual cues of /ba/ become audible. Confusion patterns for the three targets /ta/, /za/, and /za/ demonstrate how gain can help resolve consonant confusions (Figure 9). The primary competitor for the target /ta/ under no gain conditions was /ka/ and secondarily /pa/, but only /ka/ remained to a much lesser degree under the gain conditions, leading to a constant recognition score of 75% correct above 0 dB SNR. For the target /za/, the CV /ja/ was confused with the target on more than 80% of presentations above 0 dB SNR under the no-gain condition, whereas it was clearly rejected as a competitor at -6 dB SNR under the gain condition. For the target /za/, multiple competitors (primarily /va/) were prominent under the no-gain condition, but no primary competitors were evident under the gain condition. Thus, spectral gain clearly provided perceptual cues for target discrimination, especially for the M set sounds for listener S4.

The effects of gain on CPs of three M set sounds (/fa/, /sa/, and /ta/) and an E set sound (/pa/) for listener S5 are shown

in Figure 10. The confusion patterns for the target CV /fa/ (Figure 10, first column) demonstrate the negative effect of gain on CV recognition. Overall, providing gain benefited recognition of M set CVs, but the benefit was not equally distributed across all consonants. Under both conditions, the plosive /pa/ was a common competitor for the target /fa/, and the error evoked by the competitor was high. The effect of gain vs. no gain condition differed as a function of increasing SNR in that the confusion was minimally reduced under the no-gain condition, but no reduction of error was observed for the gain condition. The result was extremely poor recognition of /fa/ under the gain condition even in quiet. Interpretation of this data implies that common perceptual cues for the target /fa/ and for /pa/ were made more audible by gain. In contrast, for the targets /sa/, /ta/, and /pa/, the positive effect of gain on CV recognition was clearly marked by the changes in CPs and improved target recognition scores. Multiple competitors were found for the no-gain conditions, but the competitors were clearly rejected in the gain conditions. This suggests that the discrimination of these targets requires higher audibility but less clarity, or that there is less distortion of auditory function in noise for this listener.

For listener S6 (Figure 11), the common characteristics of CPs of one D set sound (/va/), two M set sounds (/za/ and /za/), and one E set sound (/na/) were that primary competitors were the same, but the degree of confusions varied under the two listening conditions. Extreme confusion for target /za/ was evident. This listener heard the target as /ja/ over 90% of presentations across SNRs for both conditions. This confusion may be due to the distortion of the cochlear signal, precluding the potential benefit of high-frequency gain to resolve the target. For the target /va/, two syllables, /ba/ and /fa/, competed under both conditions. The confusions were reduced to 30% and 10% of errors in quiet under the gain condition. For targets /za/ and /na/, confusions from major competitors remained high as SNR increased under no-gain conditions, but these confusions began to be resolved above 0 dB SNR under the gain condition. Thus, the effect of gain was sound-specific and listener-specific. In summary, the effect of gain on consonant perception, especially in the presence of noise maskers, cannot be generalized.

DISCUSSION & CONCLUSIONS

The present study demonstrated that the benefit of providing improved audibility to HI listeners was both consonant and

listener specific. Increasing the audibility of speech was beneficial most for the M set (/da/, /fa/, /sa/, /ta/, /za/, /za/) and little for the E set (/ga/, /ka/, /ma/, /na/, /pa/, /ja/) and but not for the D set (/ba/, /va/, /ða/, /θa/). Improved audible speech provides benefits to three listeners (the top three panels of Figure 2) but not to the other three listeners (the bottom three panels of Figure 2). The CLP analyses (Figures 3-5) identified that the listeners who generally showed negative effects of gain did so consistently for targets /ba/, /fa/, /ga/, and /ka/ but benefited from gain on certain consonants such as /sa/, /ta/, /za/, and /za/. The CP analyses (Figures 6-11) provided the nature of benefit or little or no benefit from gain. First, the primary competitors were the same between gain and no-gain conditions, but the error rates were different. Second, the primary competitors were different, or additional competitors were added to the confusion with gain, but the error rates were relatively constant.

The hypothesis we wanted to test was that audibility was directly correlated with consonant perception, especially for the D set. The presentation level was set at the most comfortable level for each listener (range 85-92 dB SPL depending on CV syllables), and the gain was provided based on sensitivity so that CV syllables were surely audible across frequency bands, that is, articulation index=1. Thus, the speech spectrum was on average 20-45 dB above the listener's corrected hearing threshold (See Figure 1). For all six listeners, the error scores of the D set were not significantly different in quiet between no-gain and gain conditions (Figure 2); hence, lack of audibility is not the only factor for poor recognition of the D set.

Plomp and colleagues modeled hearing loss with two theoretical parameters, an audibility component (attenuation of all sounds entering the ear) and a distortion component (distortion of these sounds); a first-order model was verified experimentally [22-28]. In this model, the class-audibility component is exclusively correlated to a threshold shift (audibility) in quiet that is easily compensated by amplification, whereas the class-distortion component is independent of audibility but exclusively dependent on SNR (distortion component is not compensated by amplification). Thus, it is possible that cochlear signals for the D set were highly distorted, mapping irrelevant perceptual cues.

In contrast, the CLP of the M and E sets (Figures 4,5) were relatively parallel to the diagonal but shifted downward to the right until reaching threshold, suggesting that the acoustic properties of the M and E sets were less distorted by the cochlea or more dependent on the A component [26]. Our com-

parisons of CP between HI and NH listeners confirmed that HI listeners had different or additional competitors from NH listeners for the D set but almost identical competitors for the M and E sets [13]. These results indicate that multiple acoustic signals and their cochlear/neural correlates were similarly mapped between normal and abnormal auditory systems for the M and E set, whereas they were mapped differently for the D set because of the distortion of cochlear signals.

The major finding of the present study was the idiosyncratic benefit of providing audible speech in noise. Three listeners showed benefit from gain and three others did not. This finding is consistent with many previous studies [5,6,9,10,29-32] but exists in contrast to the conclusion of Turner and Henry [33], who reported that providing audible speech resulted in increase in consonant recognition scores irrespective of the degree and configuration of hearing loss. However, this increase in performance was very small ($\leq 5\%$). The difference in experimental conditions may account for this inconsistency. In the study by Turner and Henry [33], multiple-talker babble noise was used as a masker with a fixed SNR (+9 dB SNR for HI listeners), and the final stimulus was low-pass filtered with various cutoffs to modify the amount of spectral information. In addition, a single talker, and CV and VC, were used to produce target sounds.

Listener S1 among the H group listeners is the only one who did not benefit from frequency-dependent gain. The corrected HL of S1 was 35 dB HL at $f < 2$ kHz, around 50 dB HL between 2-4 kHz, and > 60 dB HL at $f > 4$ kHz (See Figure 1). A possible explanation for the lack of gain-dependent benefit in this case is that differences in thresholds at lower (< 1.5 kHz) and higher (> 4 kHz) frequencies might limit the ability to use audible cues provided by gain. Skinner [32] found that an imbalance in thresholds between lower and higher frequencies was one of the primary factors influencing word recognition. Another explanation is that adding gain to the spectral bands where the listener has greater than moderate hearing loss for $f \geq 4$ kHz is less effective at transmitting useful cues to listeners. Turner and Cummings [5] measured nonsense recognition with 10 HI listeners and concluded that little or no benefit in performance was provided for listeners with thresholds higher than 55 dB HL at $f > 3$ kHz. In contrast, adding gain to lower frequency regions for a listener with a flat, severe-to-profound hearing loss provided a benefit with increasing audibility, even for listeners with thresholds greater than 55 dB HL. Turner and Brus [6] also reported that each HI listener benefited from lower frequency gain (< 0.9 kHz) irrespective

of the degree of hearing loss less than 70 dB HL. Fewer HI listeners benefited from higher frequency gain (>2.8 kHz) on nonsense syllables that were low-pass filtered, as a function of frequency gain and of presentation level.

For the two listeners who showed negative effects of gain on syllable recognition (S2 and S3), the corrected hearing threshold was mild hearing loss or normal hearing range for important speech frequencies (1 kHz to 4 kHz). The reason for this negative effect of gain might be related to the presentation level of syllables. None of our listeners complained about presentation level because it was determined from their most comfortable level while listening to 10 CVs in quiet. However, during the experiment, as both target sounds and noise masker (speech-weighted noise) were amplified, the level of the final stimuli could have been greater than most comfortable level and produced distortion, especially for some of high frequency sounds. Summers and Cord [34] showed that NH listeners deteriorated performance with increasing presentation level than HI listeners in the frequency range of 1.5 kHz-4 kHz. Our two listeners had a normal range of hearing thresholds in that band after amplification; thus, they might have experienced listening conditions that produced distortion. Others also reported that poorer performance at higher sound levels was even increased in the presence of competing noise [35,36]. The most comfortable level was set before the experiment by presenting 10 CVs in quiet. It is possible that providing gain introduced additional distortions, making irrelevant, but possibly common, perceptual cues audible; which explains why listeners had more uncertainty and increased confusions.

In summary, audibility is one of the primary factors influencing speech recognition of HI listeners [7,9-10,31], but reduced audibility alone could not account for the difficulty of HI listeners to understand speech [3,-4]. It should also be noted that factors other than audibility and distorting components - such as nonlinear loudness shifts, poor frequency resolution, and poor temporal resolution - can produce poor speech perception of HI listeners in noise.

Our results demonstrate that effects of gain were CV syllable and listener dependent. Three listeners benefited from gain and showed thresholds of 20 to 50 dB HL between 0.25 and 4 kHz. The other three listeners who were negatively affected by gain showed mild or normal hearing between 1 kHz and 4 kHz but steeply sloping high-frequency hearing loss between 0.5 and 4 kHz. The gain effect was not consistent with CV syllables between and within the groups that benefited and

those that did not benefit from gain. Of the three listeners who benefited from gain, listener S4 showed a positive or at least neutral (CLPs being around diagonal) effect of gain on all syllables except /θa/ and /fa/, and S5 showed a clearly positive effect of gain on all syllables except /δa/, /fa/, and /ga/, S6 showed clearly positive effect of gain on most syllables except /ba/ and /fa/, and a negative effect on recognition of /da/, /sa/, and /za/ at SNR <0 dB, but positive above that SNR. Of the three listeners who showed overall negative effects of gain on CV recognition, listener S1 showed a highly negative effect from gain on two D set sounds /δa/ and /ba/, but a nearly neutral effect on the M and the E sets. S2 showed negative effects of gain on most syllables except /δa/ and /za/, and S3 showed a negative effect of gain on all syllables except /sa/.

The comparisons of the CLP with NH listeners per CV set showed that audibility was not a major factor for poorer performance on the D set, but deficits in performance might result from a combination of audibility and distortion components. Because the CLP compared with those of NH listeners on the M and E sets were parallel to the diagonal with downward shifts at higher SNRs, these results suggest that processing of these two CV sets was less affected by distortion, but was influenced by listener's poorer thresholds.

The CP analysis showed two major patterns. First, the primary competitors were the same between gain and no-gain conditions, but the error rates were different. Second, the primary competitors were different, or additional competitors were added to the confusion with gain, but the error rates were relatively constant. For listeners whose performance was impaired by gain, additional competitors were often added to the confusions when gain was provided. For listeners who benefited from gain, the major competitors were less frequently selected under gain conditions, and minor competitors found in no-gain conditions were also resolved.

Thus, the preliminary results from the present study suggest that audibility is one of the primary factors influencing speech recognition of HI listeners, but reduced audibility alone cannot explain the difficulty of HI listeners to understand speech in noise. Similarly, amplification alone cannot restore the speech recognition ability of most HI listeners to the level of NH listeners. Even though the current study focused on analyses of subject-by-subject performances and of consonant-by-consonant recognition, it is needed to obtain additional measure from larger size of populations to draw a definitive conclusion.

ACKNOWLEDGEMENTS

We would like to thank our participants for their time and effort. We would also like to thank Allison Coltisor and Taylor Myers for their editorial assistance. This work was supported by Etymotic Research Foundation, Chicago IL.

REFERENCES

- Phatak SA, Yoon YS, Gooler DM, Allen JB. Consonant recognition loss in hearing impaired listeners. *J Acoust Soc Am*. 2009;126(5), 2683-2694.
- Han WJ, Allen JB. A relationship of speech perception and amplification for hearing-impaired listeners. *J Korean Academy of Audiology*. 2011;7:133-144.
- Turner CW, Robb MP. Audibility and recognition of stop consonants in normal and hearing-impaired subjects. *J Acoust Soc Am*. 1987;81(5):1566-1573.
- Dubno JR, Dirks DD, Ellison DE. Stop-consonant recognition for normal-hearing listeners and listeners with high-frequency hearing loss. I: The contribution of selected frequency regions. *J Acoust Soc Am*. 1989;85(1):347-354.
- Turner CW, Cummings KJ. Speech audibility for listeners with high-frequency hearing loss. *Am J Audiol*. 1999;8(1):47-56.
- Turner CW, Brus SL. Providing low- and mid-frequency speech information to listeners with sensorineural hearing loss. *J Acoust Soc Am*. 2001;109(6):2999-3006.
- Simpson, A., McDermott, H. J., & Dowell, R. C. Benefits of audibility for listeners with severe high-frequency hearing loss. *Hear Res*. 2005;210(1-2):42-52.
- Humes LE, Roberts L. Speech-recognition difficulties of the hearing-impaired elderly: The contributions of audibility. *J Speech Hear Res*. 1990;33(4):726-735.
- Zurek PM, Delhorne LA. Speech reception in noise by hearing-impaired Listeners. *J Acoust Sec Am. Suppl*. 1986;1 79:S23.
- Zurek PM, Delhorne LA. Consonant reception in noise by listeners with mild and moderate sensorineural hearing impairment. *J Acoust Soc Am*. 1987;82(5):1548-1559.
- Humes LE, Dirks DD, Bell TS, Kincaid GE. Recognition of nonsense syllables by hearing-impaired listeners and by noise-masked normal hearers. *J Acoust Soc Am*. 1987;81(3):765-773.
- Fousek P, Svojanovsky P, Grezl F, Hermansky H. (2004). "New nonsense syllables database analyses and preliminary ASR experiments," in *Proceedings of the International Conference on Spoken Language Processing (ICSLP)*, October 4-8, Jeju, South Korea, http://www.isca-speech.org/archive/interspeech_2004. Viewed 3/26/07.
- Phatak S, Allen BJ. Consonant and Vowel Confusions in Speech-Weighted Noise. *J Acoust Soc Am*. 2007;121(4):2312-2326
- Lobdell BE, Allen JB. A model of the VU (volume-unit) meter, with speech applications. *J Acoust Soc Am*. 2007;121(1):279-285.
- MATLAB (2002). "The language of technical computing," The MathWorks, Inc., Natick, MA.
- Byrne D. Effects of bandwidth and stimulus type on most comfortable loudness levels of hearing-impaired listeners. *J Acoust Soc Am*. 1986;80(2):484-493.
- Byrne D. Effects of frequency response characteristics on speech discrimination and perceived intelligibility and pleasantness of speech for hearing-impaired listeners. *J Acoust Soc Am*. 1986;80(2):494-504.
- Byrne D, Murray N. Predictability of the required frequency response characteristic of a hearing aid from the pure-tone audiogram. *Ear Hear*. 1986;7(2):63-70.
- Byrne D, Cotton S. Evaluation of the national acoustic laboratories' new hearing aid selection procedure. *J Speech Hear Res*. 1988;31(2):178-186.
- Byrne D. (1993). Implications of National Acoustic Laboratories' (NAL) research for hearing aid gain and frequency response selection strategies. In: Studebaker GA, Hochberg I, eds. *Acoustical Factors Affecting Hearing Aid Performance*. Boston: Allyn and Bacon, 119-131.
- Martin FN, Champlin CA, Chambers JA. Seventh survey of Audiometric practices in the United States. *J Am Acad Audiol*. 1998;9:95-104.
- Duquesnoy AJ, Plomp R. The effect of a hearing aid on the speech-reception threshold of hearing-impaired listeners in quiet and in noise. *J Acoust Soc Am*. 1983;73(6):2166-2173.
- Dreschler WA, Plomp R. Relation between psychophysical data and speech perception for hearing-impaired subjects. *J Acoust Soc Am*. 1980;68(6):1608-1615.
- Dreschler WA, Plomp R. Relation between psychophysical data and speech perception for hearing-impaired subjects. *J Acoust Soc Am*. 1985;78(4):1261-1270.
- Festen JM, Plomp R. Relations between auditory functions in impaired hearing. *J Acoust Soc Am*. 1983;73(2):652-662.
- Plomp R. Auditory handicap of hearing impairment and the limited benefit of hearing aids. *J Acoust Soc Am*. 1978;63(2):533-549.
- Plomp R, Mimpen AM. Improving the reliability of testing the speech reception threshold for sentences. *Audiology*. 1979;18(1):43-52.
- Plomp R, Mimpen AM. Speech-reception threshold for sentences as a function of age and noise level. *J Acoust Soc Am*. 1979;66(5):1333-1342.
- Ching TYC, Dillon H, Byrne D. Speech recognition of hearing-impaired listeners: Predictions from audibility and the limited role of high-frequency amplification. *J Acoust Soc Am*. 1998;103:1128-1140.
- Hogan CA, Turner CW. High-frequency audibility: Benefits for hearing-impaired listeners. *J Acoust Soc Am*. 1998;104:432-441.
- Rankovic CM. An application of the articulation index to hearing aid fitting. *J Speech Hear Res*. 1991;34(2):391-402.
- Skinner MW. Speech intelligibility in noise-induced hearing loss: Effects of high-frequency compensation. *J Acoust Soc Am*. 1980;67(1):306-317.
- Turner CW, Henry BA. Benefit of amplification for speech recognition in background noise. *J Acoust Soc Am*. 2002;112(4):1675-1680.
- Summers V, Cord MT. Intelligibility of speech in noise at high presentation levels: effects of hearing loss and frequency region. *J Acoust Soc Am*. 2007;122(2):1130-1137.

35. Speaks C, Karmen JL, Benitez L. Effect of a competing message on synthetic speech identification. *J Speech Hear Res.* 1967;10:390-396.
36. Studebaker GA, Sherbecoe RL, McDaniel DM, Gwaltney CA. Monosyllabic word recognition at higher-than-normal speech and noise levels, *J Acoust Soc Am.* 1999;105:2431-2444.