A Relationship of Tone, Consonant, and Speech Perception in Audiological Diagnosis

Woojae Han and Jont B. Allen*

Div. of Speech Pathology and Audiology, College of Natural Science, Hallym Univ., Chuncheon, Korea
*Dept. of Electrical and Computer Engineering, Univ. of Illinois at Urbana-Champaign, IL, USA
(Received April 4, 2012; revised May 14, 2012; accepted June 5, 2012)

ABSTRACT: This study was designed to examine the phoneme recognition errors of hearing-impaired (HI) listeners on a consonant-by-consonant basis, to show (1) how each HI ear perceives individual consonants differently and (2) how standard clinical measurements (i.e., using a tone and word) fail to predict these differences. Sixteen English consonant-vowel (CV) syllables of six signal-to-noise ratios in speech-weighted noise were presented at the most comfortable level for ears with mild-to-moderate sensorineural hearing loss. The findings were as follows: (1) individual HI listeners with a symmetrical pure-tone threshold showed different *consonant-loss profiles* (CLPs) (i.e., over a set of the 16 English consonants, the likelihood of misperceiving each consonant) in right and left ears. (2) A similar result was found across subjects. Paired ears of different HI individuals with identical pure-tone threshold presented different CLPs in one ear to the other. (3) Paired HI ears having the same averaged consonant score demonstrated completely different CLPs. We conclude that the standard clinical measurements are limited in their ability to predict the extent to which speech perception is degraded in HI ears, and thus they are a necessary, but not a sufficient measurement for HI speech perception. This suggests that the CV measurement would be a useful clinical tool.
Key words: Tone perception, Consonant perception, Consonant loss profile, Speech perception, Hearing impairment, Audiologic diagnosis, Audiologic test battery
ASK subject classification: Speech Production and Perception (12)

I. Introduction

It is well known that most patients with sensorineural hearing loss (SNHL) complain of difficulty in understanding speech with a hearing aid (HA) that has been fitted based on a battery of clinical tests [1-4]. While it is frequently stated by clinicians and researchers that pure-tone audiogram (PTA) does not significantly reflect hearing-impaired (HI) speech perception, given the almost ubiquitous clinical use of the half-gain rule or of NAL -R [5], clearly the clinical community is comfortable with their application of the PTA in the fitting of a HA. The usual justification of this procedure is based on the audi-

bility of the speech, captured in the expression "If you can't hear it, you cannot understand it." Unfortunately, "hearing it" does not guarantee that you will "understand it".

In the present study, we pose two key hypotheses: (1) When the HI ear cannot resolve a solitary acoustic cue (e.g., voice onset time or duration of a burst, etc.), a high error rate for only a few consonants results. That is, each HI ear would have a unique *consonant-loss profile* (CLP), defined by significant subset of consonants errors, unique to that ear (i.e., defined by the diagonal entries of the consonant confusion matrix). (2) Neither pure-tone audiograms (PTA) nor the speech recognition threshold (SRT) measurements can quantify this unique profile, because these average measures do not parse out perceptional differences at the consonant level. Consequently, HAs fitted on the basis of the PTA or SRT will necessarily

*Corresponding author: Woojae Han (woojaehan@hallym.ac.kr)
Hallym Univ., Div. of Speech Pathology and Audiology, Hallymdaehakgil 1, Chuncheon, Korea, 200-702
(Tel: 82-33-248-2216; Fax: 82-33-256-3420)

provide less benefit than HAs fitted on the basis of the few high-error consonants, as defined through the CLP. In this research paper we propose a novel speech test that we believe uniquely quantifies HI ear's idiosyncratic CLP. Only with such detailed idiosyncratic knowledge, based on speech feature loss unique to that ear, can we hope to proceed with the fitting of modern signal processing HAs. We review the pros and cons of popular clinical measurements commonly used to establish speech perception ability in HI listeners.

1.1 Current Clinical Measurements

Pure-tone audiometry is ubiquitously used to measure hearing sensitivity, to determine the degree, type, and configuration of an individual's hearing loss, and to establish either middle-ear or cochlear/auditory nerve damage [6]. Although this measurement is fast, easy to use, thus widely accepted, audiometry does not directly evaluate the ability of the HI listener to perceive speech sounds [7]. In fact, it is widely accepted that the PTA correlates poorly with HI speech perception [7-8]. Many studies have reported that for listeners with moderate-to-severe SNHL, there is *no correlation* between hearing threshold and speech perception, while others report a partial positive correlation for listeners with normal to mild SNHL [7,9]. We shall show that while an elevated threshold does predict that there will be some speech loss, it gives no diagnostic information as to the nature of that speech loss. Many studies have attempted to develop predictions of a listener's ability to understand speech on the basis of his pure-tone sensitivity. For example, Fletcher (1950) and later Smoorenburg (1992) developed a formula for predicting the HI listener's ability to perceive speech from the three-frequency average of hearing thresholds at the most important frequencies (i.e., 3-tone average (3TA) of 0.5, 1, and 2 kHz) [7,10]. They found that there was a very large across-subject variance, which depends on audiometric configuration. In particular, the 3TA had much lower (better) thresholds than speech scores for a non-flat audiogram (e.g., high-frequency ski-slope hearing loss) [8,10].

The fact that there is such loose relation between the PTA and speech perception has serious clinical ramifications.

The SRT was introduced by Plomp (1986), who defined it as the signal-to-noise ratio (SNR) at which the listener achieves 50 % accuracy for recognizing syllables, words, or sentences [11]. The SRT has been widely accepted, due to its convenience and speed, and has become a preeminent "speech test". While distinct from pure-tone audiometry, it clinically correlates well with PTA in quiet [6,12]. The SRT has three serious limitations. First, this measure evaluates a listener's speech threshold, not the ability to recognize speech. Simply said, it is a wide-band threshold test using speech, instead of narrow-band tones, quantified via a VU meter in 5-dB steps [6]. Like the PTA, the SRT has equally limited ability to predict the listener's speech recognition ability. The problem of HI speech perception is not the deficit in detection, but rather poor recognition [13]. Second, the SRT uses 20 homogeneous *spondee words* (with doubly-stressed meaningful syllables; e.g., air-plane, birth-day, cow-boy) having high context, because tests based on spondee words are easier and faster to administer than those based on sentences [6,8]. It is a problem that when the spondee words are used, patients say what they *guess*, not what they actually perceive. Third, the SRT considers only average speech scores instead of focusing on individual consonant scores. Being an average measure, it ignores valuable information about what a listener hears, that is, detailed consonant articulation scores that contain essential, even critical information about acoustic cues of the speech stimuli that the HI ear can or cannot hear. Averaged scores remove not only the wide variance of speech perception, but also the key characteristics of hearing loss.

Apart from the PTA and SRT measurements, various word/sentence tests have been used to diagnose the degree of impairment and to evaluate the benefits of HAs. These tests have become increasingly popular over the years, in part because standardized versions have become available, such as the Psychoacoustic Laboratory Phonetically Balanced monosyllabic word lists (PAL PB-50) [14], the Hearing

In Noise Test (HINT) [15], the Revised Speech Perception In Noise (SPIN-R) [16], and the Quick Speech-In-Noise test (QuickSIN) [2]. Although the tests all differ slightly in composition, research shows that a common advantage of these tests is to simulate everyday listening conditions which are realistic for measuring the speech perception ability of HI listeners [2,17]. However, these tests fail to fully reflect HI speech perception in terms of the acoustic and speech cues, because a contextual bias is inherent in these word/sentence tests [18,19]. Boothroyd (1994) clearly demonstrated that HI listeners decode consonant-vowel-consonant (CVC) based on both direct sensory evidence and indirect contextual evidence as they decode the speech sound [20]. Versfeld et al. (2000) also insisted that redundancy in speech makes hearing-impaired listeners' perceptual scores improve more than one would predict from their hearing loss [21]. As with the SRT, familiar words or topics make it even easier to understand a conversation, whereas new or unfamiliar ones make it more difficult [18,22]. Of course, the contextual linguistic skills are essential and natural in communication, but they are not appropriate in a hearing test for speech perception. Since these features allow the HI listeners to guess the target words [23,24], the test scores *do not address listeners' core and unique individual consonant errors* [25]. Thus we must separate our measures of consonant perception from the contextual effect.

1.2 Nonclinical Measurements

On the other hand, various researchers have worked for nonclinical test measurements in the research setting. In 1921, Harvey Fletcher created the Articulation Index (AI), which is used in the prediction of the average phone error [10]. Although Fletcher revised the calculation for clinical application, his revised method has not been extensively used in practice because the AI provides no diagnostic or frequency dependent information. In addition, the complexity of the AI led to its disuse in clinic settings although it can be useful in choosing the gain of a hearing aid. In 1955, Miller and Nicely developed the consonant

Confusion Matrix (CM), which is a useful quantitative and visual representation, displaying the stimulus versus the response score in a table [27]. The CM allows for greater understanding of an individual's CLP (i.e., over a set of the 16 English consonants, the likelihood of misperceiving each consonant), because it gives detailed information about consonant confusions - that is, (1) which sounds an HI listener can or cannot hear (i.e., diagonal entries) and (2) which sounds are confused with other sounds (off-diagonal entries). Nevertheless, Miller and Nicely (1955)'s CM method has clinical shortcomings because it is complex, time consuming, and difficult to interpret. In 1976, Bilger and Wang studied an average consonant CM measure in 22 SNHL patients, using a combination of 16 consonants and 3 vowels averaged across SNRs [26]. They reported that HI listeners made specific consonant errors, with error patterns that depend on the degree and configuration of the hearing loss, as well as the level of noise. While measuring consonant-vowel (CV) and vowel-consonant (VC) confusions, they only reported mean scores (% correct) of 4 CV subsets. Their findings strongly suggest the need for further research into the detailed characteristics of consonant perception error, which are idiosyncratic across HI listener.

1.3 Purpose of Study

In this study, we measured individual consonant error to quantify how each SNHL listener confuses the consonants. Next, we compared these consonant percent errors (%) to the current and commonly used clinical measurements (e.g., PTA and SRT) to determine the clinical power of a confusion. Two key hypotheses were posed: (1) When a HI listener misses resolving a solitary acoustic cue (e.g., voice onset time or duration of a burst, etc.), the result is a high error rate for only a few consonants. This measurement is defined in the CLP, as quantified by a small but significant subset of consonant errors, unique to each ear and (2) Neither the PTA nor SRT measurements can quantify such a unique profile, because all average measures do not parse out perceptional differences at the consonant level.

Consequently, hearing aids fitted on the basis of the PTA or SRT necessarily provide less benefit than those fitted on the basis of a small number of high-error consonants, identified by the CLP. Only with such detailed idio-syncratic knowledge, based on speech feature loss unique to that ear, can we hope to proceed with the most beneficial fitting of modern hearing aids.

II. MATERIALS AND METHODS

2.1 Subjects

Twenty-seven HI subjects (17 females and 10 males) were recruited from the Urbana-Champaign community. All subjects were native speakers of American English. They ranged in age from 21 to 88 years (mean = 54.96 years, SD = 20.28 for all; 60.63 years for males and 49.08 years for females). Subjects were chosen based on normal middle-ear status (type A of tympanogram) and mild-to-moderate SNHL at 3TA (3-tone average in hearing thre-shold at 0.5, 1, and 2 kHz). Informed consent was obtained and approved by the Institutional Review Board of the University of Illinois at Urbana-Champaign.

The etiologies of individual subjects varied, in terms of the degree and configuration of hearing loss. Of the 27 subjects, 21 had symmetrical bilateral, 4 had asymmetrical bilateral, and 2 had unilateral hearing loss. Of these, ear-by-ear, 10 ears had flat audiograms, with 3 mild, 4 mild-to-moderate, and 3 moderate SNHL. Another 16 ears showed high-frequency SNHL varying in the degree of impairment, with 8 mild, 6 moderate, and 2 moderate-to-severe in hearing loss. A mild-to-moderate high frequency SNHL was present in 11 ears, with a ski-slope loss at either 1 or 2 kHz. The following atypical configurations were also included: 2 ears with low-frequency hearing loss, 2 with cookie-bite (middle-frequency) hearing loss, 3 with reversed cookie-bite (low- and high-frequencies) hearing loss, and 4 with mild hearing loss accompanied by a notch at 4 kHz.

2.2 Speech Stimuli

Isolated English CV syllables were chosen from the Linguistic Data Consortium (LDC) 2205S22 database [27], spoken by eighteen native speakers of American-English. The CV syllables consisted of sixteen consonants (six stops /p, b, t, d, k, g/, eight fricatives /f, v, s, ʃ, z, ʒ, ð, θ/, and two nasals /m, n/) followed by the /a/ vowel [28]. Only using /a/ vowel was allowed to control other possible variables and to see the consonant perception (or consonant loss profile) in hearing impaired listeners, which is the most common problem in the population. All stimuli used were digitally recorded at a sampling rate of 16 kHz. They were presented monaurally in quiet and at five different SNRs (+12, +6, 0, -6, -12 dB) in speech-weighted noise. The presentation level of the syllables was set to the subject's most comfortable level (MCL) initially, and then adjusted so that the CVs were equally loud independent of SNR.

2.3 Experimental Procedure

The procedures for the CV measurements were very similar to those used in a previous study by Phatak et al. [27]. While sitting in front of a computer in the sound booth and listening a test stimulus presented by Matlab program through an inserted earphone (ER-2), all subjects had one practice session consisting of 10 syllables in quiet to familiarize each subject with the experiment. Subjects were asked to identify each presented consonant of the CV syllable, by selecting 1 of 16 software buttons on a com-puter screen, each labeled with an individual consonant sound. A 'noise only' button was available for the subjects to specify if they heard *only noise*. A pronunciation for each consonant was provided using an example word below its button to avoid possible confusions from any orthographic similarity between consonants. The subjects were allowed to hear each utterance a maximum of three times before making their decision. Once a response was entered, the next syllable was automatically presented after a short pause. The experiment took a total of 1 to 1.5 hours per ear.

Each syllable presentation was randomized with respect to consonants and speakers, but not across SNRs. The test proceeded from the easiest to the most difficult noise conditions - quiet first, followed by +12 to -12 dB SNR. This was done in order to gradually increase the difficulty from the onset, so that subjects were not pushed beyond their limits in terms of performance level. A maximum of 1152 trials were presented (16 consonants × 6 utterances × 2 presentations × 6 different noise conditions) to every subject. When the score was less than or equal to 3/16 (18.75 %, or three times chance) for any given consonant, the consonant was not presented at subsequent (lower) SNRs.
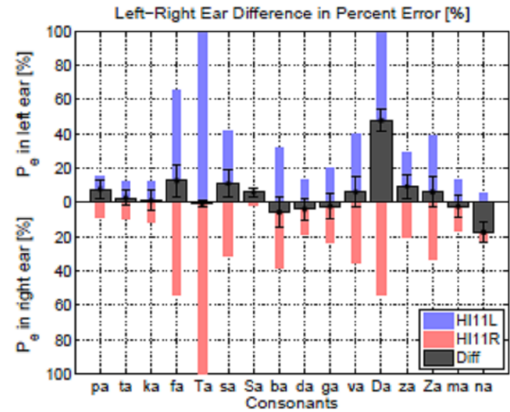
III. RESULTS

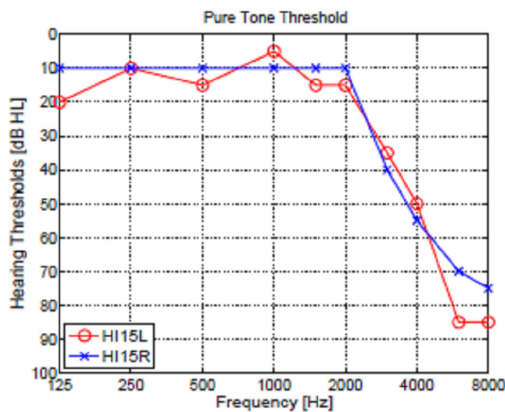### 3.1 Comparisons between the PTA and CLP

Both Fig. 1 and Fig. 2 show two PTAs (left panels) along with their CLP (right panels). In Fig.1, two HI subjects show a symmetrical hearing-loss in the left and right ears: (a) high-frequency and (c) high-frequency ski-slope hearing loss. In the Fig. 2, panel (a) shows two
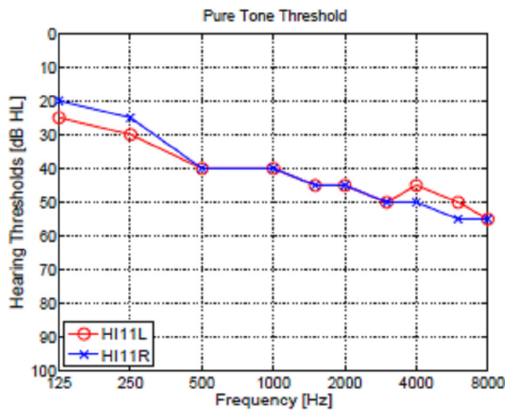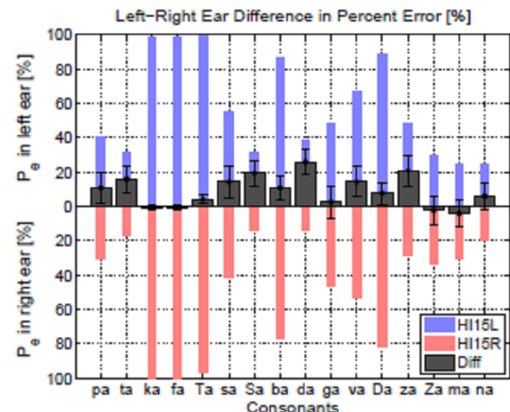

(a) Gradual slope: HI11L / 11R

(b) Consonant-loss of HI11L / 11R

(c) Ski slope: HI15L / 15R

(d) Consonant-loss of HI15L / 15R

Fig. 1. The two left panels show PTA results in the HI subjects and the right panels show their consonant loss profiles in left vs. right ears across the 16 consonants. On the right panels, bar graphs present percent error (%) of each consonant in blue for left ear and red for right ear. The gray bars show left ear vs. right ear advantage: above zero shows a right-ear advantage and below shows a left-ear advantage. To see the colors, check the website uploaded version. Error barsindicate 1 standard error (SE). Even though these subjects have symmetrical hearing loss (a,c), their consonant perception is asymmetrical and is inhomogeneous across consonants (b,d). PTA cannot predict individual HI ears' consonant-loss. ∗Due to limitation of creating IPA symbols in MATLAB, the consonants, /θa/, /ʃa/, /ɚa/, and /ʒa/ are displayed as Ta, Sa, Da, and Za, respectively.

(a) Identical PTA: HI36R / 40R

(b) Consonant-loss of HI36R / 40R

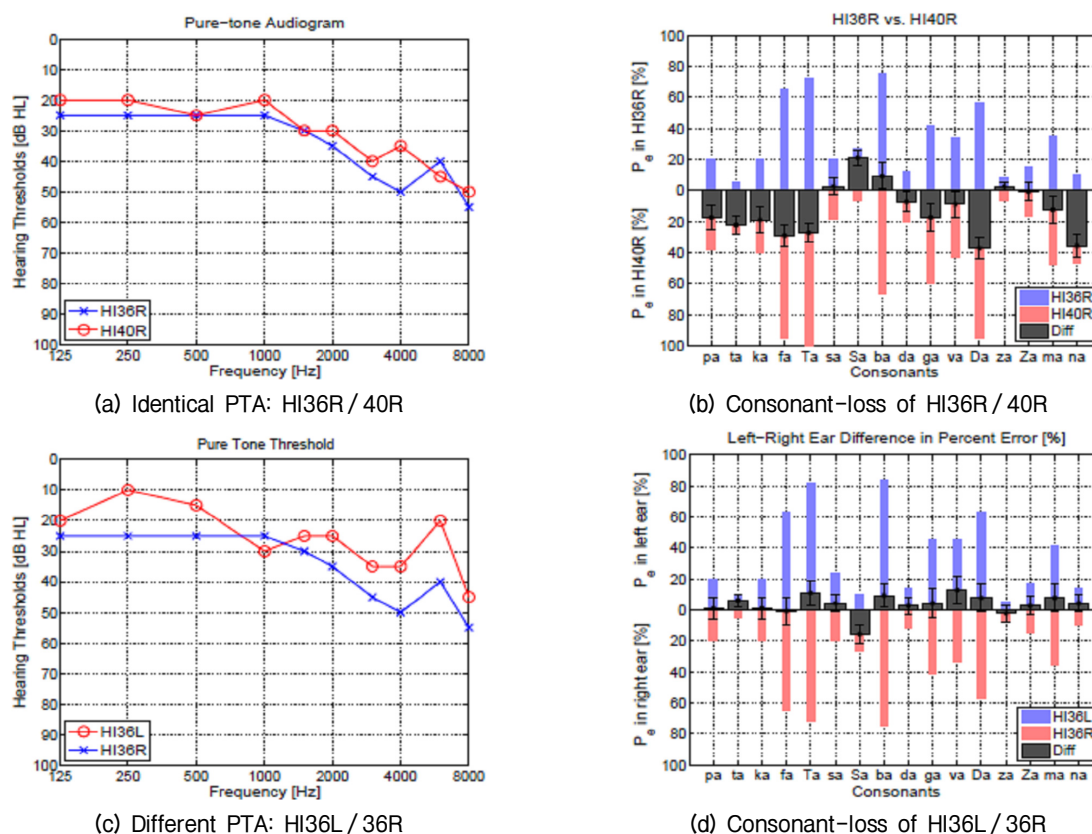(c) Different PTA: HI36L / 36R

(d) Consonant-loss of HI36L / 36R

Fig. 2 The two left panels show PTA results in the HI subjects and the right panels show their consonant loss profiles across the 16 consonants. There is a difference in CLP between two different HI subjects having identical PTA (a). The subject with the asymmetrical pure-tone loss (c) does not have an asymmetrical consonant loss profile (d).

different HI listeners with nearly identical PTAs, while the HI subject of panel (c) has an asymmetrical PTA.

Each of the right panels shows percent error for each consonant in both left and right ears as blue and red bars from the baseline, respectively. The difference in the percent error of consonant identification between the left and right ears across 16 consonants is presented as block wide bar graphs. The gray bar located above the horizontal axis indicates a right-ear advantage, while the bar below the horizontal axis indicates a left-ear advantage for that consonant. (To see the colors, check the website uploaded version.)

3.1.1 Gradual Sloping High Frequency Hearing Loss

Subject HI11 in Fig.1 (a,b) had high error rate in /fa, θa, ða/ for both ears. The /θa/ syllable had 100 % error in both

ears. She could not perceive /ða/ with her left ear, but correctly perceive it at 50 % in her right ear. The 4 consonants /ta/, /ka/, /ga/, and /ma/ resulted in low error rate and also elicited no significant difference between ears. HI11 has a left-ear advantage of about 18 % for /na/, a 46 % right-ear advantage in /ða/ and a small 10～15 % right-ear advantage for the /fa/, /sa/, and /za/ syllables.

3.1.2 Ski-slope High Frequency Hearing Loss

Subject HI15 Fig.1 (c,d) showed 100 % error rate for /ka/, /fa/, and / θa/ syllables and about 80 % error rate for /ba/ and /ða/ in both ears. Compared to subject HI11, this subject has higher error rates in many consonants although she has a better pure-tone threshold below 4 kHz. In spite of her symmetrical PTA, the subject HI15 showed a right-ear advantage for 12 out of 14 consonants (about 2～ 2 5 %). Even though the PTA threshold was 10～15 dB

HL higher (worse) at 6 - 8 kHz, her HL could not explain better performance in the right ear even for syllables containing low frequency consonants, /pa/ and /ba/.

3.1.3 Identical Audiogram and Different Consonant-loss

Two subjects with identical pure-tone threshold, HI36R and HI40R Fig. 2 (a,b), show dissimilar error rates and patterns in their consonant perception. HI36R has a lower consonant error rate overall (excluding /ʃa/), compared to HI40R who has almost 100 % error rate for /fa/, /θa/, and /ða/ syllables. The largest difference in consonant error rate between the two subjects was for the /ða/ and /na/ syllables, about 38 %. Again, their obviously different CLPs are not predicted by their nearly identical audiograms.

3.1.4 Dissimilar Audiogram and Same Consonant-loss

Subject HI36 of Fig. 2 (c,d) has an asymmetrical pure-tone hearing loss and about 20 dB HL better audibility in the left ear. However, his consonant-loss profile is not consistent with this difference. Overall, he poorly perceives the /fa/, /θa/, /ba/, and /ða/ syllables, with less than a 20 % difference between the two ears. The better audiogram in the left ear does not lead to a left-ear advantage in consonant perception; instead, there is a small right-ear advantage for a number of consonants.

3.2 Comparisons between the CRT and CLP

Fig. 3 shows that consonant-loss and the consonant recognition threshold (CRT) can be poorly correlated. CRT is defined as the SNR at which the listener achieves 50 % accuracy for recognizing consonants, which is same criteria as SRT and might show a limitation of average measurement. In Fig. 3 (a), six HI ears are paired in terms of their CRTs (-3, 0, and 4.5 dB SNR), shown by black dashed lines. Their consonant-loss is shown in sub-figures (b), (c), and (d). Note that the paired ears do not have the same CLP, even though they have the same average consonant scores. In Fig. 3 (b), although both ears have a

CRT of -3 dB SNR, HI29L heard /ba/ 40 % better than HI36R. The difference in /ba/ perception was up to 60 % at 0, 6, and 12 dB SNR (not shown). The ear also performed 20 % better for /ʃa/. However, the same ear (HI29L) showed 20 ~ 38 % poorer performance for /ka/, /θa/, /ða/, and /na/, when compared to HI36R. In Fig. 3 (c), HI26R was better than HI40L in most of the CVs. Interestingly, however, HI26R could not correctly perceive /ʒa/ at all, while HI40L could (a 70 % difference). Of the two HI ears having a 4.5 dB CRT (Fig. 3 [d]), HI15L was much better with /ʒa/, while the other ear was better with /ka/.

While the CRTs in this example are consistent with the extent of consonant-loss, it cannot explain the random nature of the CLP. The audiogram configurations were mild flat, mild-to-moderate gradual high frequency, and mild-to-moderate ski-slope high-frequency hearing loss in (b), (c), and (d), respectively. While there was no difference in the average scores and PTAs for the paired ears, their consonant losses differ dramatically as shown by CLP measures. In summary, the ears' consonant perception abilities seem to differ randomly when compared to their PTA and SRT.

IV. Discussion and Conclusion

The goal of the present study was to measure CLPs for HI listeners and to compare them to existing clinical measurements, thereby underscoring well-known deficiencies in those measurements.

4.1 CLP is not correlated with PTA

HI individuals with symmetrical hearing loss can have asymmetrical consonant perception, whereas the individuals who have asymmetrical PTAs can show little differences in CLP between two ears. Earlier studies have supported these results. Killion (1997) states that pure-tone threshold is limited in its utility for predicting speech perception because the loss of audibility and loss of speech clarity (i.e., SNR-loss) are functionally separated [29]. In other words, there is a major difference between hearing

(a) Gradual slope: HI11L / 11R

(b) Consonant-loss of HI11L / 11R

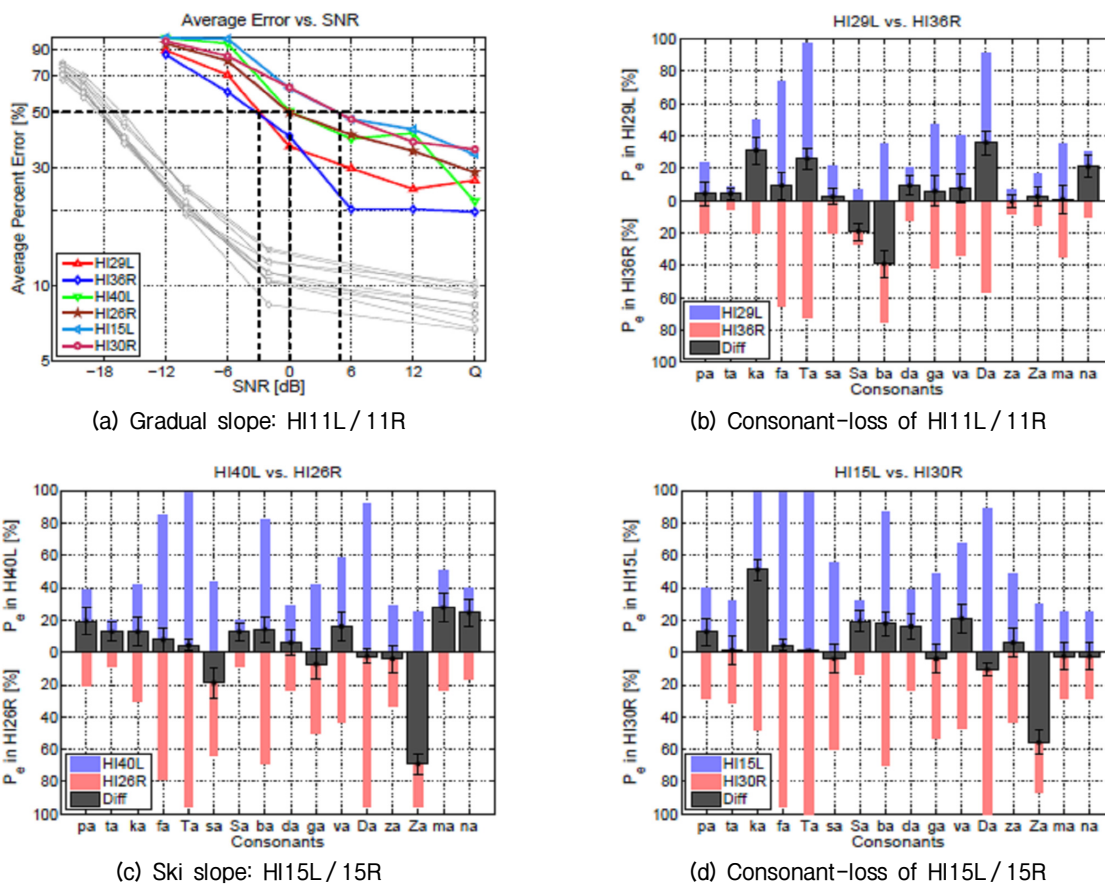(c) Ski slope: HI15L / 15R

(d) Consonant-loss of HI15L / 15R

Fig. 3. The CRT and CLP of HI ears are compared. The left top panel (a) shows the CRT threshold defined as the SNR at 50 % average error, for six pairs of ears showing the same CRT: −3, 0, and 4.5 dB SNR. The right top and two bottom panels show plots of consonant-loss difference between two ears as a function of consonants. Bar graphs present percent error of each consonant as blue for one ear and red for the other ear. The gray bars show left ear vs. right ear advantage: above the zero line one ear has a higher error (disadvantage), and below the line the right ear has the disadvantage. Note that one ear is much better than the other in some consonants although they have same CRT.

speech (i.e., audibility of speech) and understanding it (i.e., intelligibility of speech). Theoretically speaking, patients with outer hair cell (OHC) and/or inner hair cell (IHC) loss could show the same hearing threshold, yet have different symptoms. This is because damage to the OHCs reduces the active vibration of the basilar membrane at the frequency of the incoming signal, resulting in an elevated detection threshold. Damage to the IHCs reduces the efficiency of transduction [29]. Given identical detection thresholds, it might be that OHCs and IHCs impact speech perception differently. For example, some individuals have a much greater loss of intelligibility in noise than might be expected from their audiogram [29]. In order to

avoid the limitations of pure-tone audiometry, Killion suggests that the graphic Count-the-Dot Audiogram Method be used to estimate the AI [30]. This method provides an easy and practical way to clinically measure the degree of the HI patient's loss of speech clarity by computing the number of dots on the audiogram [31]. Yet, the method cannot give an estimate of the inhomogeneous extent of speech perception. The Count-the-Dot Audiogram, like the AI, does not provide information regarding an asymmetry in speech perception between two ears.

Our CV syllable test may explain HI individuals' ear preference when using the telephone. Ten subjects with symmetrical hearing loss were asked about their phone ear

preference by e-mail survey. Eight of them reported to have an ear preference while using a cell phone, which correlates with their CLP. The CLP may be a useful in deciding which ear to fit in cases of monaural HAs although the threshold is the main variable considered when fitting hearing aids (e.g., NAL-R). The CV test may also predict problems in listeners who have normal hearing but complain that speech is unclear under specific noisy circumstances. Our findings are similar to those of Danhauer (1979), who showed that there is no relationship between PTA and CLP [32].

Dubno and Schaefer (1992) found a correlation between frequency selectivity and consonant recognition using 66 CV and 63 VC syllables for both HI and masked normal hearing (NH) listeners [33]. Their results showed that frequency selectivity is poorer for HI listeners than for masked NH listeners. However, there is no difference in consonant recognition between two groups having equal speech-spectrum audibility. A major study completed by Zurek and Delhorne (1987) also revealed that the average consonant reception performance is arguably not significantly different from that of masked NH listeners. They conclude that audibility is the primary variable in speech scores [4]. Note that their argument is based entirely on average consonant scores. Here we argue that perception of individual consonants is not dependent on PTA, even when thresholds between the two groups are matched. Thus, our conclusion is the opposite to that of that of both Dubno & Schaefer [33] and Zurek & Delhorne [4]. We argue the use of the CLP rather than average scores. The large difference between ears implies a significant cochlear-specific deficiency. Such a difference could be due to specific cochlear lesions, such as a cochlear dead region.

PTA thus correlates poorly with consonant accuracy and is useless in predicting the frequency regions where the consonant is included. Our speech test precisely identifies the consonant errors, which, when compared to our knowledge of the key frequencies of each speech feature, should allow one to precisely pinpoint dysfunctional frequency regions in that ear.

4.2 CLP is not also correlated with CRT

Although Plomp (1986) proposed the SRT test to connect the detection of PTA and speech perception [11], the SRT is not actually a perception test, rather it is a speech audibility test. Turner *et al.* (1992) found that consonant detection of HI listeners in a suprathreshold-level masking noise was not different from that of NH listeners. In addition, they explained that HI listeners' poor speech perception might be due to their inability to efficiently utilize audible speech cues [13]. Our CRT measure is poorly correlated to consonant recognition in an HI listener and is supporting Turner's study, while we have not obtained data from spondee SRT measurements (as is typically used in the clinics). Fig. 3 shows that consonants may not have the same errors in 2 ears having the same average scores. It is apparent that the consonant errors are independent of the CRT and 3TA. Since HI ears show errors in only a few sounds, average scores or word/ sentence scores obscure these unique and relevant errors.

4.3 CV syllable scores present CLP

All SNHL listeners have a loss of both sensitivity and speech clarity [11,28]. The loss of sensitivity is represented by the PTA and can be easily evaluated. However, as Plomp's distortion function and later Killion's SNR-loss express, clarity is not revealed by either PTA or SRT measurements. Our results show poorer consonant perception for most HI listeners in quiet as well as for lower SNR thresholds than for NH listeners, with respect to the average scores. This defines an SNR-loss for HI listeners and is consistent with the results of Killion's 1997 study.

4.4 New approach for considering Korean acoustic and perceptual characteristics

Although a test battery of Korean speech perception in audiology and hearing science field has been developed, not for long, to date, relatively little data exist based on naturally produced CV stimuli, which include the listener's perceptual information, such as specific acoustic/phonetic

features. We believe that approach for Korean acoustic characteristics and perceptual confusion using Korean CV syllables would give a better understand primarily regarding features within speech signals that yield a greater chance of extracting essential cues for human auditory/oral communication.

Acknowledgement

References

1. H. Dillon, *Hearing Aids*, Thieme, NewYork, 2001.
2. M.C. Killion, P. A. Niquette, G. I. Gudmundsen, L. J. Revit, and S. Banerjee, "Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.*, vol. 116, pp. 2395-2405, 2004.
3. R. Plomp, "Auditory handicap of hearing impairment and the limited benefit of hearing aids," *J. Acoust. Soc. Am.*, vol. 63, pp. 533-549, 1978.
4. P. Zurek and L. Delhorne, "Consonant reception in noise by listeners with mild and moderate sensorineural hearing impairment," *J. Acoust. Soc. Am.*, vol. 82, no. 5, pp. 1548-1599, 1987.
5. D. Byrne and H. Dillon, "The National Acoustic Laboratories' (NAL) New Procedure for Selecting the Gain and Frequency Response of a Hearing Aid," *Ear Hear.*, vol. 7, pp. 257-265, 1986.
6. W. T. Brandy, "Speech Audiometry," in *Handbook of Clinical Audiology* (5th Ed.), edited by J. Katz, pp. 96-110, Lippincott Williams and Wilkins, Baltimore, 2002.
7. G. F. Smoorenburg, "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.*, vol. 91, pp. 421-437, 1992.
8. R. Carhart, "Speech reception in relation to pattern of pure tone loss," *J. Speech Disorders*, vol. 11, pp. 97-108, 1946.
9. J. M. Festen and R. Plomp, "Speech-reception threshold in noise with one and two hearing aids," *J. Acoust. Soc. Am.*, vol. 79, no. 2, pp. 465-471, 1986.
10. H. Fletcher, "A method of calculating hearing loss for speech from an audiogram," *J. Acoust. Soc. Am.*, vol. 22, pp. 1-5, 1950.
11. R. Plomp, "A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired," *J. Speech Hear. Res.*, vol. 29, pp. 146-154, 1986.
12. R. A. Dobie and C. S. Sakai , "Noise induced hearing loss basic mechanisms, prevention, and control," in *Estimation of Hearing Loss Severity from the Audiogram*, edited by D. Henderson, D. Prasher, R. S. Kopke, R. Hamernik, pp.351-363, NRN publications, London, 2001.
13. C. W. Turner, D. A. Fabry, S. Barrett, and A. R. Horwitz, "Detection and recognition of stop consonants by normal-hearing and hearing-impaired listeners," *J. Speech Hear. Res.*, vol. 35, pp. 942-949, 1992.
14. J. Egan, "Articulation testing methods," *Laryngoscope*, vol. 58, pp. 955-991, 1948.
15. M. Nilsson, S. Soli, and J. Sullivan, "Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise," *J. Acoust. Soc. Am.*, vol. 95, pp. 1085-1099, 1994.
16. R. C. Bilger, J. M. Nuetzel, and W. M. Rabinowitz, "Standardization of a test of speech perception in noise," *J. Speech Hear. Res.*, vol. 27, pp. 32–48, 1984
17. D. N. Kalikow, K. N. Stevens, and L. L. Elliot, "Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability," *J. Acoust. Soc. Am.*, vol. 61, pp. 1337-1351, 1977.
18. G. A. Miller, G. A. Heise, and W. Lichten, "The intelligibility of speech as a function of the context of the test materials," *J. Exp. Psychol.*, vol. 41, no. 5, pp. 329-335, 1951.
19. S. Phatak, Y. Yoon, D. Gooler, and J. Allen, "Consonant recognition loss in hearing impaired listeners," *J. Acoust. Soc. Am.*, vol. 126, no. 5, pp. 2683-2694, 2009.
20. A. Boothroyd, "Speech perception by hearing-impaired listeners," *J. Acoust. Soc. Am.*, vol. 95, no. 5, p. 2998, 1994.
21. N. J. Versfeld, L. Daalder, J. M. Festen, and T. Houtgast, "Method for the selection of sentence materials for efficient measurement of the speech reception threshold," *J. Acoust. Soc. Am.*, vol. 107, no. 3, pp. 1671-1684, 2000.
22. C. M. Connine, J. Mullennix, E. Shernoff, and J. Yelen, "Word familiarity and frequency in visual and

auditory word recognition," *J. Exp. Psychol.*, vol. 16, no. 6, pp. 1084-1096, 1990.

23. A. Boothroyd and S. Nittrouer, "Mathematical treatment of context effects in phoneme and word recognition," *J. Acoust. Soc. Am.*, vol. 84, no. 1, pp. 101-114, 1988.

24. A. W. Bronkhorst, A. J. Bosman, and G. F. Smoorenburg, "A model for context effects in speech recognition," *J. Acoust. Soc. Am.*, vol. 93, no. 1, pp. 499-509, 1993.

25. M. K. Pichora-Fuller, B. A. Schneider, and M. Daneman, "How young and old adults listen to and remember speech in noise," *J. Acoust. Soc. Am.*, vol. 97, no. 1, pp. 593-608, 1995.

26. R. C. Bilger and M. D. Wang, "Consonant Confusions in Patients with Sensorineural Hearing Loss." *J. Speech Hear. Res.*, vol. 19, pp. 718-748, 1976.

27. P. Fousek, P. Svojanovsky, F. Grezl, and H. Hermansky, "New nonsense syllables database - Analyses and preliminary ASR experiments," in The *Eighth International Conference on Spoken Language Processing (ICSLP)*, May, 2004.

28. G. A. Miller and P. Nicely, "An Analysis of Perceptual Confusions among some English Consonants," *J. Acoust. Soc. Am.*, vol. 27, no. 2, pp. 338-352, 1955.

29. M. C. Killion, "SNR Loss: I can hear what people say, but I can't understand them," *Hear. Review*, vol. 4, no. 12, pp. 8-14, 1997.

30. S. Kujawa and M. Liberman, "Adding insult to injury: cochlear nerve degeneration after temporary noise-induced hearing loss," *J. Neurosci.*, vol. 29, no. 45, pp. 14077-14085, 2009.

31. H. G. Mueller and M. C. Killion, "An easy methold for calculating the articulation index," *Hear. J.* vol. 43, no. 9, pp. 14-17, 1990.

32. J. L. Danhauer and R. M. Lawarre, "Dissimilarity ratings of English consonants by normally-hearing and hearing-impaired indivisuals," *J. Speech Hear. Res.*, vol. 22, no. 2, pp. 236-246, 1979.

33. J. R. Dubno and A. B. Schaefer, "Comparison of frequency selectivity and consonant recognition among hearing-impaired and masked normal-hearing listeners," *J. Acoust. Soc. Am.* vol. 91, no. 4, pp. 2110-2121, 1992.

Profile

‣ Woojae Han

She received the PhD degree from Dept. of Speech and Hearing Science in Univ. of Illinois at Urbana–Champaign (UIUC), IL, USA, in Aug. 2011. Since Sept. 2011, she has served as an assistant professor in Hallym University, Div. of Speech Pathology and Audiology. Her interests are hearing pathology, tinnitus, as well as speech perception in hearing–impaired listeners.

‣ Jont B. Allen

He received his PhD degree from Electrical Engineering in Univ. of Pennsylvania, PA, USA in May 1970. During his 32 year AT&T Bell Labs career, he has specialized in auditory signal processing. Since Aug. of 2003, he has joined the ECE faculty in UIUC, and has concentrated on cochlear modeling, speech music coding, speech perception, and aspects of acoustics.