

## Research Article

# Predicting Speech Recognition Using the Speech Intelligibility Index and Other Variables for Cochlear Implant Users

Sungmin Lee,<sup>a</sup> Lisa Lucks Mendel,<sup>b</sup> and Gavin M. Bidelman<sup>b,c,d</sup>

**Purpose:** Although the speech intelligibility index (SII) has been widely applied in the field of audiology and other related areas, application of this metric to cochlear implants (CIs) has yet to be investigated. In this study, SIIs for CI users were calculated to investigate whether the SII could be an effective tool for predicting speech perception performance in a population with CI.

**Method:** Fifteen pre- and postlingually deafened adults with CI participated. Speech recognition scores were measured using the AzBio sentence lists. CI users also completed questionnaires and performed psychoacoustic (spectral and temporal resolution) and cognitive function (digit span) tests. Obtained SIIs were compared with predicted SIIs using a transfer function curve. Correlation and regression analyses were conducted on perceptual and demographic predictor variables to investigate the association between these factors and speech perception performance.

**Result:** Because of the considerably poor hearing and large individual variability in performance, the SII did not predict speech performance for this CI group using the traditional calculation. However, new SII models were developed incorporating predictive factors, which improved the accuracy of SII predictions in listeners with CI.

**Conclusion:** Conventional SII models are not appropriate for predicting speech perception scores for CI users. Demographic variables (aided audibility and duration of deafness) and perceptual–cognitive skills (gap detection and auditory digit span outcomes) are needed to improve the use of the SII for listeners with CI. Future studies are needed to improve our CI-corrected SII model by considering additional predictive factors.

**Supplemental Material:** <https://doi.org/10.23641/asha.8057003>

A cochlear implant (CI) is a prosthetic device that converts acoustic signals into electrical stimuli to excite surviving auditory nerve fibers. Previous studies have shown that better aided thresholds for CI users were correlated with higher speech recognition performance, emphasizing the importance of a wider dynamic range that increases audibility in the CI map (Firszt et al., 2004; Holden et al., 2013). However, clinicians frequently fit CIs based on the patient's loudness comfort rather

than the consideration of audibility. This causes variability in aided thresholds across frequency among CI users with similar degrees of hearing loss. Given that aided audibility is a contributing factor, as well as other demographic/perceptual factors, to speech perception outcomes, investigating the traditional speech intelligibility model that predicts speech perception outcomes based majorly on the audibility is worth consideration. This study aimed to see how well the traditional speech intelligibility model works for CI users and attempted to improve the model using other contributing factors.

Over 60 years ago, efforts to quantify listeners' speech intelligibility led to the development of predictive models and articulation theory (French & Steinberg, 1947). The computational procedure has been enhanced and supplemented, such that the *articulation index* has been renamed the *speech intelligibility index* (SII; American National Standards Institute [ANSI], 1997). This model considers audibility ( $A_i$ ) and frequency importance functions ( $I_i$ ) as key components to predict speech intelligibility (Equation 1).

<sup>a</sup>Erik Jonsson School of Engineering and Computer Science, University of Texas at Dallas

<sup>b</sup>School of Communication Sciences and Disorders, University of Memphis, TN

<sup>c</sup>Institute for Intelligent Systems, University of Memphis, TN

<sup>d</sup>Department of Anatomy and Neurobiology, University of Tennessee Health Sciences Center, Memphis

Correspondence to Sungmin Lee: [Sung.Lee@utdallas.edu](mailto:Sung.Lee@utdallas.edu)

Editor-in-Chief: Frederick (Erick) Gallun

Editor: Jennifer Lentz

Received July 25, 2018

Revision received November 2, 2018

Accepted December 18, 2018

[https://doi.org/10.1044/2018\\_JSLHR-H-18-0303](https://doi.org/10.1044/2018_JSLHR-H-18-0303)

**Disclosure:** The authors have declared that no competing interests existed at the time of publication.

$$\text{SII} = \sum_{i=1}^n I_i A_i. \quad (1)$$

The amount of speech energy available to listeners and the relative importance weights for each frequency band, respectively, are multiplied, and all outcome values are summed to calculate SIIs ranging from 0 to 1. The SII unit, however, does not solely account for speech recognition outcomes. To predict speech perception scores using SIIs, a transfer function that establishes the relationship between the SII and speech perception is required.

The SII model is highly accurate at predicting speech perception scores in individuals with normal hearing (Pavlovic, Studebaker, & Sherbecoe, 1986; Sherbecoe & Studebaker, 2003), whereas incorporation of correction values is recommended when calculating the SII for individuals with hearing impairment. In other words, there are additional factors that affect speech recognition beyond audibility for those with hearing loss. Several correction factors, known as hearing loss desensitization (HLD) factors, that compensate for such suprathreshold deficits have been developed and proposed (Ching, Dillon, & Byrne, 1998; Fletcher & Galt, 1950; Pavlovic et al., 1986; Studebaker, Gray, & Branch, 1999; Studebaker, Sherbecoe, McDaniel, & Gray, 1997). Such correction factors—associated with hearing thresholds of individuals with mild-to-moderate hearing loss—have improved the accuracy of SII prediction to some extent. However, these modified SIIs have not improved predictive accuracy for people with hearing loss greater than moderate impairment (Ching et al., 1998; Ludvigsen, 1987; Pavlovic et al., 1986).

Given the limited success of the SII in cases of severe hearing loss, it is reasonable to assume that applying it to CI users (who typically have severe-to-profound losses) may also prove difficult. Furthermore, deteriorations in auditory processing and large individual variability in speech recognition outcomes among CI users could make it more challenging to use the SII as a predictable tool for speech perception in implant patients. Despite these concerns and in light of technological advancements in modern CIs that improve perceptual outcomes of speech, it is worth examining the feasibility of using the SII to predict speech perception performance in CI users. This preliminary study examined the application of the SII for predicting speech perception outcomes for CI users and scrutinized new ways to improve the SII model by considering important factors shown to be predictive of CI users' perceptual performance.

### ***Spectral/Temporal Resolution in CIs***

Spectral resolution refers to one's sensitivity in detecting fine acoustic changes in the frequency domain. CI users are known to have very poor spectral resolution for several reasons. Physiologically, neural excitation patterns of electrical hearing are broader than those of acoustic hearing (Macherey & Carlyon, 2014), resulting in poor frequency sensitivity caused by overlapping auditory filters for CI users. Functionally, a CI system primarily extracts and

transfers temporal envelope cues, and temporal fine structure cues in the speech signal are typically lost. Although it has been established that envelope cues alone can transfer sufficient information for speech perception in quiet, the role of temporal fine structure cannot be disregarded considering its significant contribution to pitch perception (Oxenham, Bernstein, & Penagos, 2004; Smith, Delgutte, & Oxenham, 2002) and speech perception in noise (Lorenzi, Gilbert, Carn, Garnier, & Moore, 2006). In addition, the approximately 22 electrical channels in CI systems that substitute for thousands of inner and outer hair cells may not be enough to deliver fine frequency information required for robust perception. Increasing the number of electrodes could result in adverse effects of channel interactions in certain circumstances such as with monopolar current. These physiological challenges, combined with the technical impossibility of designing auditory filter bands that function like the normal acoustic mechanism, cause poor spectral resolution for listeners with CI.

The ability to resolve or segregate temporal events in a stream of sound is called *temporal resolution*. CIs use a train of biphasic pulses as the carrier of envelope cues to transmit acoustic information. Theoretically, higher stimulation rates of electrical pulses are beneficial, as fine temporal modulation information can be delivered to listeners. However, many cases have been reported where CI users do not exploit these higher stimulation rates to improve speech perception (Lee & Mendel, 2016; Vandali, Whitford, Plant, & Clark, 2000), an effect thought to result from poorer temporal resolution. The reasons behind this are assumed to be due to the characteristics of auditory nerve firing patterns in response to electrical pulse trains used in CIs: (a) absolute refractory periods and resting potentials of neural firing patterns do not fit the fast rates of electrical stimulation and (b) a train of biphasic pulses is not appropriate to provide exact timing information because it consists of two opposite polarities that cause action potentials with different latencies (Macherey & Carlyon, 2014).

These spectral and temporal aspects of auditory processing are highly associated with speech perception performance. For this reason, many CI studies have examined these cues in relation to speech recognition (Fu & Shannon, 2000; Nie, Barco, & Zeng, 2006; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995; Xu & Zheng, 2007). Poorer auditory spectral and temporal processing in CI users is thought to contribute to the high variability in their speech perception outcomes. Here, we examined the degree to which individual differences in CI users' spectral and temporal resolution might be predictive of their speech perception.

### ***Working Memory as a Contributing Factor***

Over the past few years, the effect of central cognitive function on CI users' speech perception has received much attention (Burkholder & Pisoni, 2006; Collison, Munson, & Carney, 2004; Pisoni & Cleary, 2003; Pisoni & Geers,

2000). When assessing working memory associated with people with hearing loss, the concept of phonological loop is frequently applied. The phonological loop is a part of working memory model (Baddeley, 1993; Baddeley & Hitch, 1974) that stores phonological information and functions as an articulatory rehearsal process. In everyday life, linguistic information is encoded by sensory systems, and then, the phonological information is rehearsed and stored in one's memory. People with hearing loss, whose auditory sensory functions are diminished, may have problems with utilizing phonological representations of input information. To compensate for such diminished perceptual ability, patients with CI may depend more heavily on the top-down processes that make use of phonological/lexical access and long-term memory storage (Moberly, Harris, Boyce, & Nitttrouer, 2017). This top-down process is governed by cognitive abilities. Thus, associations between speech recognition and working memory for CI users may be quite important.

Working memory for CI subjects are generally poorer than those of their normal hearing counterparts (Geers, Pisoni, & Brenner, 2013; Pisoni, Kronenberger, Roman, & Geers, 2011). A Digit Span Test (DST; Pisoni & Cleary, 2003; Pisoni & Geers, 2000) that measures working memory is known to be highly correlated with speech recognition in children with CI. We included DSTs to examine working memory capacity for CI users in relation to their speech perception performance. The tests presented stimuli auditorily and visually to investigate working memory load with and without the detrimental effect of hearing loss.

### ***Other Source of Variance in Speech Perception Performance for CI Users***

CIs do not provide equal benefit to all users, and there is enormous variability in speech perception performance among implant patients. Some CI recipients show near-normal performance exceeding expectations, whereas the performance of others is so poor they discontinue wearing their CI. Thus, determining factors that predict perceptual benefits from CI surgery is crucial in establishing realistic clinical expectations and rehabilitation strategies for CI recipients. In fact, a large number of studies have been conducted to address this issue by looking at the correlation between speech perception and surgical, demographic, psychophysical, and cognitive variables. Overall, studies agree that duration of deafness is one of the most critical factors that determine performance after implantation (Blamey et al., 1996; Daya et al., 1999; Gordon, Daya, Harrison, & Papsin, 2000; Green et al., 2007; Holden et al., 2013; van Dijk et al., 1999). Preimplant factors, however, cannot fully account for the limited speech perception outcomes seen in individuals with CI. Other factors, such as communication mode, audibility, etiology of deafness, habilitation, and cognitive function also have been found to contribute to variance in speech recognition among patients with CI (Collison et al., 2004; Holden et al., 2013; Pisoni,

Cleary, Geers, & Tobey, 1999; Schafer & Utrup, 2016). In this study, CI users' audiologic/demographic variables were investigated to examine their effects on speech perception variability.

### ***Aim of the Study***

SIIIs are often used with hearing aids, but little attention has been paid to the application of SIIIs for patients with CI. The lack of SII studies with patients with CI may be attributed to the significant hearing loss that typical individuals with CI have and the distorted electrical signals that CIs provide. Moreover, individual variability and heterogeneity typically observed in a population with CI may also be a primary reason for few studies using the SII with this population. This variability has caused uncertainty of how much clinicians and patients can expect from CIs during aural rehabilitation. To address this, we used the SII model that has never been examined with CI users. The major purpose of this preliminary study was to gauge the general applicability of the SII to the population with CI and improve the effectiveness of the SII prediction by including other perceptual-cognitive measures in its calculation.

First, a transfer function curve that established the relationship between SII and speech perception scores was used to determine if the SII could serve as an effective tool for predicting speech perception performance for this population. We then examined the role of other predictive factors in speech perception performance. Adult CI users' demographics, auditory processing ability, and working memory load were explored using a correlation analysis. Psychoacoustic and perceptual measures (e.g., masking, level distortion, HLD, and age) were also considered to investigate the predictive power of these factors with speech perception performance. A new SII<sub>CI</sub> model is proposed—incorporating these predictive factors as weights to the original model—that improves the accuracy of the SII for predicting speech perception performance in CI users. Our prediction model can be used when planning aural rehabilitation for CI recipients. In this case, if clinics have access to patient information such as that used in this study, audiologists would be able to set a target goal for maximizing speech perception for CI users based on the predicted model. This study lays the initial foundation to expand this idea for future SII models that may need less patient information but have higher accuracy.

## **Method**

### ***Participants***

Fifteen adults with CI ranging in age from 22 to 73 years ( $M = 53.13$ ,  $SD = 17.27$ ) participated in the current study. The inclusion criteria consisted of participants who were younger than 80 years of age, had experience with CI device(s) for at least 6 months, and whose first language was American English. All participants had severe-to-profound sensorineural hearing loss with bilateral pure-tone

averages (PTA) of greater than 70 dB HL. The group mean PTA for the left ears was 95 dB HL ( $SD = 7.2$  dB HL), and that for the right ears was 97 dB HL ( $SD = 4.8$  dB HL). The listeners with CIs signed informed consent forms, and all were paid for their participation after completing all procedures. The protocol employed in this research was approved by the University of Memphis Institutional Review Board. Participants completed a questionnaire addressing patient demographics and hearing history. Some of these demographics, such as duration of hearing loss, were used later for the regression analysis. Table 1 represents the demographic details of the participants with CI.

### Audiometric Testing

Aided and unaided audiometric tests were conducted to verify hearing thresholds and audibility with and without their CIs. Hearing thresholds were obtained at octave frequencies from 250 to 8000 Hz, and interoctave frequencies (125, 750, 1500, 3000, and 6000 Hz) were also confirmed. Aided audiometry was carried out in a free field condition with participants seated in the center of a

double-walled sound booth meeting ANSI standard S3.1-1999 (ANSI, R2013), facing the front speaker at a 1-m distance. In the case of bimodal participants who wore a CI on one ear and a hearing aid on the other ear, the hearing aid was removed during the test. This rule was applied to other experiments in the study as well. Unaided audiometry was conducted using pure tones presented through TDH-39 headphones.

### Speech Recognition Testing

Listeners with CIs' speech recognition was measured using the AzBio Sentence Test (Spahr & Dorman, 2005). The AzBio Sentence Test is one of the standardized speech perception tests in the Revised Minimum Speech Test Battery that was designed to be used with patients with CI. The AzBio stimuli are produced by two male and two female speakers that can be presented in quiet or in noise (10-talker babble). Each participant listened to three AzBio sentence lists in three test conditions presented in the sound field: (a) quiet, (b) +5 dB signal-to-noise ratio (SNR), and (c) +10 dB SNR. The level of the speech was fixed at 65 dB SPL, whereas noise levels were varied depending

**Table 1.** Demographic details of participants with cochlear implants (CIs).

No.	Gender	Age (years)	Onset of hearing loss (years)	Duration of deafness (years)	Etiology	CI manufacturer and model	Level of education	Uni/bilateral CI	Communication mode
1	Female	41	13	0	Unknown	Cochlear Nucleus	Master's degree	Uni	Oral/speaking, lipreading
2	Male	73	40	0	Noise exposure	Cochlear Nucleus	Master's degree	Uni	Oral/speaking, lipreading
3	Male	63	46	13	Unknown	Cochlear Nucleus	High school	Uni	Oral/speaking, lipreading, writing
4	Male	30	7	3	Unknown	Cochlear Nucleus	Doctoral degree	Uni	Oral/speaking, lipreading
5	Female	56	41	7	Ménière's	Cochlear Nucleus	Master's degree	Uni	Oral/speaking, lipreading
6	Female	65	51	10	Unknown	Cochlear Nucleus	Master's degree	Uni	Oral/speaking
7	Female	56	5	50	Nerve damage	Cochlear Kanso	Bachelor's degree	Uni	Oral/speaking, lipreading
8	Male	73	41	31	Noise exposure	Cochlear Kanso	Master's degree	Uni	Oral/speaking, lipreading
9	Female	40	15	14	Brain tumor	Cochlear Nucleus	Master's degree	Uni	Oral/speaking
10	Female	59	0	55	Rubella	Medel Opus	Bachelor's degree	Uni	Oral/speaking, lipreading, sign language
11	Female	65	56	2	Unknown	Cochlear Nucleus	Doctoral degree	Bi	Oral/speaking, lipreading
12	Female	64	42	21	Ménière's	Cochlear Nucleus	High school	Uni	Oral/speaking, lipreading
13	Male	24	2	2	Illness	Advance Bionics Harmony	Master's degree	Uni	Oral/speaking, lipreading
14	Female	22	2	0	Meningitis	Advance Bionics Harmony	Bachelor's degree	Uni	Oral/speaking
15	Male	66	50	0	Ménière's	Cochlear Nucleus	Bachelor's degree	Uni	Oral/speaking

on the desired SNRs. Each listener with CI was seated in the middle of the double-walled sound booth meeting ANSI standard S3.1-1999 (ANSI, R2013), 1 m away from the speaker, wearing his or her CI device. The listener's task was to repeat the sentences or words they heard. Among the 15 lists available in the AzBio test, Lists 8, 9, 10, 11, 12, and 13 were chosen as they were equally difficult based on results from a previous study (Bush, 2016). Among these lists, three were randomly selected and presented in the different conditions. Each list consists of 20 sentences that contain a different number of target words per sentence. The responses were scored in percentage based on the number of target words correctly repeated by the listeners.

### ***SII Calculation***

The SII was computed in the following way. For each patient with CI's aided thresholds to be used in the SII calculation, equivalent hearing threshold levels needed to be established. The aided audiometric thresholds measured in dB HL were converted to dB SPL by adding the transformation values proposed by Bentler and Pavlovic (1989). Critical ratios (Pavlovic, 1987) and bandwidth adjustments (ANSI, 1997) were further used to transform the obtained thresholds into equivalent hearing threshold levels. The equivalent hearing threshold levels were eventually used in the SII equation. To yield one-third octave pure-tone thresholds that could not be obtained from the audiometric procedures, interpolation or extrapolation was used.

As noted earlier, two key components, audibility and frequency importance functions, need to be established for calculation of the SII. For computation of audibility function in the three speech recognition task conditions, the Long-Term Average Speech Spectrum of the AzBio lists and its noise were measured. To this end, the overall root-mean-square level of 65 dB SPL and the levels of the concatenated speech and noise were measured separately using a Bruel and Kjaer Type 2250 sound level meter. Figure 1 shows the band-specific levels in LAeq across the one-third octave band frequencies. The shape of the speech and noise spectra reflects nearly identical patterns across the frequencies.

The audibility function was calculated by subtracting the larger of either the long-term noise levels or the thresholds from the spectral peaks in each band and dividing the difference by 30 dB. Table 2 shows the frequency importance function for the AzBio sentences that was derived in our prior study (Lee & Mendel, 2017). This band weighted information of the AzBio sentences was applied in the SII calculation using Equation 1. The entire procedure of computing SII values followed the ANSI standard (ANSI, 1997), which takes into consideration masking effects and a level of distortion factors. After the SII calculation, an age correction factor proposed by Studebaker et al. (1997) was multiplied for those who were older than 70 years. This correction was applied because of the tendency of speech processing to decline with age (Bidelman, Lowther, Tak, &

Alain, 2017; Bidelman, Villafuerte, Moreno, & Alain, 2014). SII calculations were generated using custom routines coded in Excel.

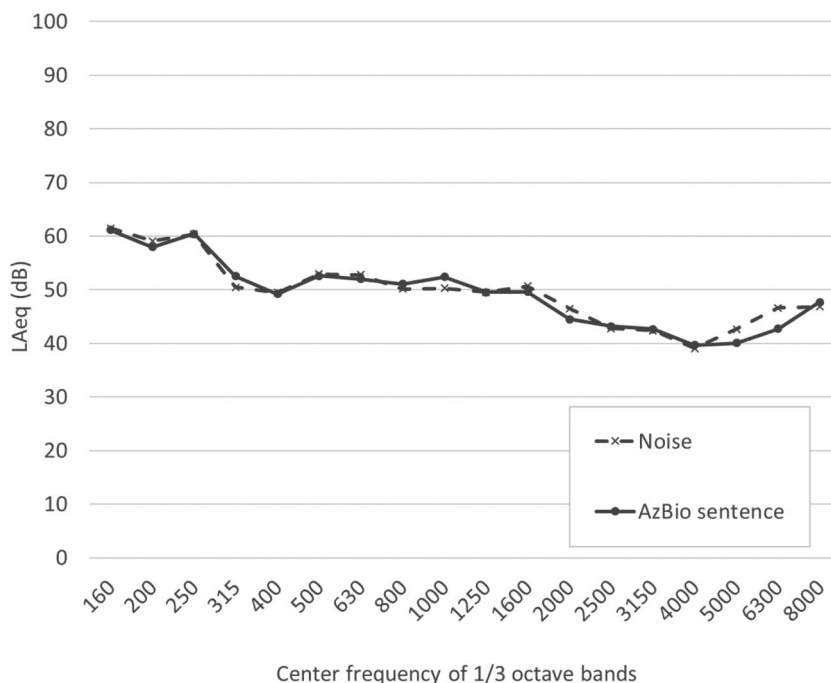
### ***Auditory Processing Tests***

It is reasonable to assume that suprathreshold deficits associated with poor speech recognition in individuals with hearing impairment are attributed to abnormal spectral and temporal resolution (Pavlovic et al., 1986). For measuring auditory processing, the Gap Detection Test (GDT; Florentine & Buus, 1984) and the Spectral-Temporally Modulated in Ripple Test (SMRT; Aronoff & Landsberger, 2013) were administered to determine each listener's temporal and spectral resolution, respectively. The auditory processing tests were administered twice for each participant, and the average of the two runs was used for subsequent statistical analysis.

We used PsyAcoustX (Bidelman, Jennings, & Strickland, 2015), a MATLAB-based software platform that implements three-alternative forced-choice psychophysical tasks, to measure perceptual GDTs. Three successive broadband noises, 500 ms each, were presented at 65 dB SPL through a loud speaker located 1 m and at 0° azimuth from the listener. One of the three stimuli was designed to have a short silent gap, whereas the other two were continuous broadband noise. The durations of the short gap were varied depending on the listener's response following a two-down/one-up adaptive tracking rule to determine one's GDT threshold with 71% criterion performance level (Levitt, 1971). The starting gap duration was set to 10 ms. On each trial, the listener was asked to click a button on the screen to reflect the detection of a brief gap that divided two successive stimuli. Threshold was measured as the geometric mean of the last eight of 12 reversals.

The Spectral Ripple Test that uses stimuli containing a different number of spectral peaks at a particular modulation depth is one of the most commonly used approaches to evaluate spectral resolution in modern CI studies (Henry & Turner, 2003; Henry, Turner, & Behrens, 2005). Won, Drennan, and Rubinstein (2007) found that better spectral ripple discrimination was significantly correlated with better speech perception in noise and quiet for CI users. In our study, spectral resolution of patients with CI was estimated using the SMRT software Version 1.1 (Aronoff & Landsberger, 2013). The SMRT measured the largest number of ripples per octave (RPO) that could be reliably detected by listeners. The test uses an adaptive procedure (one-up/one-down). Like the GDT, the SMRT was administered with three-alternative forced choice presenting stimuli at 65 dB SPL. The ripple density of reference stimuli was 20 RPO, and the target stimuli were adjusted starting from 0.5 RPO with a step size of 0.2 RPO. The trial ended when 10 reversals were found, and the mean of the last six reversals was reported as the RPO threshold. Listeners were instructed to click a button on the screen that reflected the stimulus that sounded different

**Figure 1.** Long-Term Average Speech Spectrum of AzBio sentences and noise across one-third octave band frequencies.



from the other two stimuli, discriminating the spectrally different sound.

### Cognitive Function Tests

We used the DST to evaluate auditory and visual working memory function. The DST is a subset of the Wechsler Adult Intelligence Scale–Revised (Wechsler & De Lemos, 1981), which assesses comprehensive cognitive ability for adults and consists of six verbal subtests (Information, Comprehension, Arithmetic, Digit Span, Similarities, and Vocabulary) and five performance subtests (Picture Arrangement, Picture Completion, Block Design, Object Assembly, and Digit Symbol). In the DST, a participant is required to memorize a series of numbers presented either visually or auditorily and repeat the correct numbers in the correct order. Outcomes from auditory DSTs may not reflect pure attention and memory deficits, as it is significantly influenced by hearing deficits for individuals with hearing loss. The DST can further be divided into two tasks, depending on the answering method. The forward task asks participants to answer in the presented order, whereas the backward task requires listeners to answer in reverse order. We used both visual and auditory

modalities and both forward and backward responses to compare the functional difference in working memory. The DST was run using the Inquisit computer software (Draine, 1998). For the visual DST, a sequence of numbers was shown on a computer screen and then disappeared. For the forward DST, participants were instructed to click the correct digits on the monitor in the correct order. For the backward DST, they were instructed to click the correct digits in the reverse order. The auditory DST was administered in the same way, but the sequence of numbers was presented from the front speaker at 65 dB SPL, instead of the monitor screen. The digit string was increased in length with each trial until the participant was unable to remember the correct numbers in the correct sequences. The final (span) score was the maximum length of numbers that were correctly recalled in order (or reverse order for backward DST).

### Results

#### Prediction of Speech Perception Scores Using SII<sub>s</sub>

Mean speech perception scores for three different SNR conditions (Quiet, SNR 5, and SNR 10) are shown

**Table 2.** Frequency importance function (FIF) across the one-third octave center frequencies (CFs).

CF	160	200	250	315	400	500	630	800	1000	1250	1600	2000	2500	3150	4000	5000	6300	8000
FIF	2.20	2.41	1.74	1.35	3.08	5.42	8.71	7.43	7.35	9.72	10.21	9.13	9.65	7.91	6.65	3.65	2.56	0.84

in Figure 2 (speech perception scores of individuals with CI are also provided along with hearing thresholds on Supplemental Material S1). Listeners with CI had particular difficulty under the noise compared to the quiet conditions. Scores decreased when background noise was presented along with the AzBio sentences. A one-way repeated-measures analysis of variance determined that mean speech perception scores differed between the presentation conditions of the AzBio sentences,  $F(2, 25.717) = 112.893, p < .001$ . Post hoc tests using Bonferroni corrections revealed that all pairs of conditions differed from one another, indicating that an increase in the amount of background noise resulted in a decrease in speech perception scores ( $p < .001$ ). In addition, the extended boxes and whiskers in Figure 2 suggested that individual differences in speech perception performance were large.

To investigate the predictive role of the SII for speech perception performance, a transfer function curve that established the relationship between predicted scores using SII and observed scores was determined. We used the transfer function for the AzBio test derived by Lee and Mendel (2017) using Equation 2. The transfer function equation was obtained from listeners with normal hearing who were administered the AzBio test in a variety of filtering/SNR conditions. The fitting constants,  $Q$  (0.287) and  $N$  (5.206) in Equation (2) resulted in a good fit between the observed and predicted scores (root-mean-square error [RMSE] = 0.069 and  $R^2 = .923$ ). With the appropriate reference transfer function for listeners with normal hearing as a normative point, our speech perception data and corresponding SII for listeners with CI were examined.

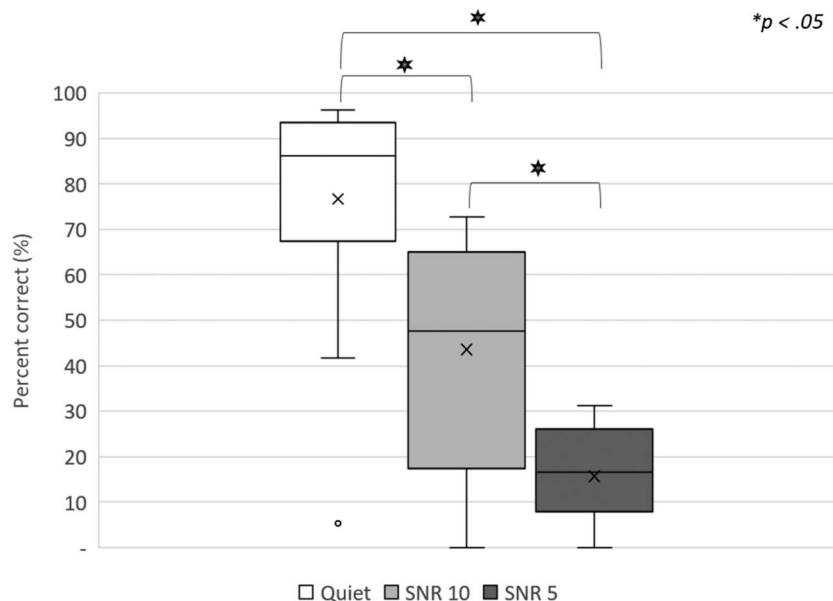
$$\text{Speech perception score} = \left(1 - 10^{-(\text{SII}/0.287)}\right)^{5.206}. \quad (2)$$

Figure 3a provides the scores-versus-SII transfer function for the AzBio test derived from listeners with normal hearing (solid line). In addition, the SII values and the corresponding speech perception scores obtained from this study are represented by filled gray circles for the +10-dB SNR condition, filled black circles for the +5-dB SNR condition, and white for quiet. Regardless of the test condition, scores fell considerably below the predicted scores using the transfer function curve for listeners with normal hearing, suggesting it is not capable of predicting speech perception scores for listeners with CI using the SII model. In an attempt to address the issue of overestimation by the conventional SII calculation, we applied an HLD factor to the SII calculation. Among several HLD models, we adopted an equation similar to the one developed by Sherbecoe and Studebaker (2003; Equation 3), where PTA is the average of unaided hearing threshold of the better ear at 1, 2, and 4 kHz. This correction factor was then applied by multiplying the SII values with the HLD for each CI user.

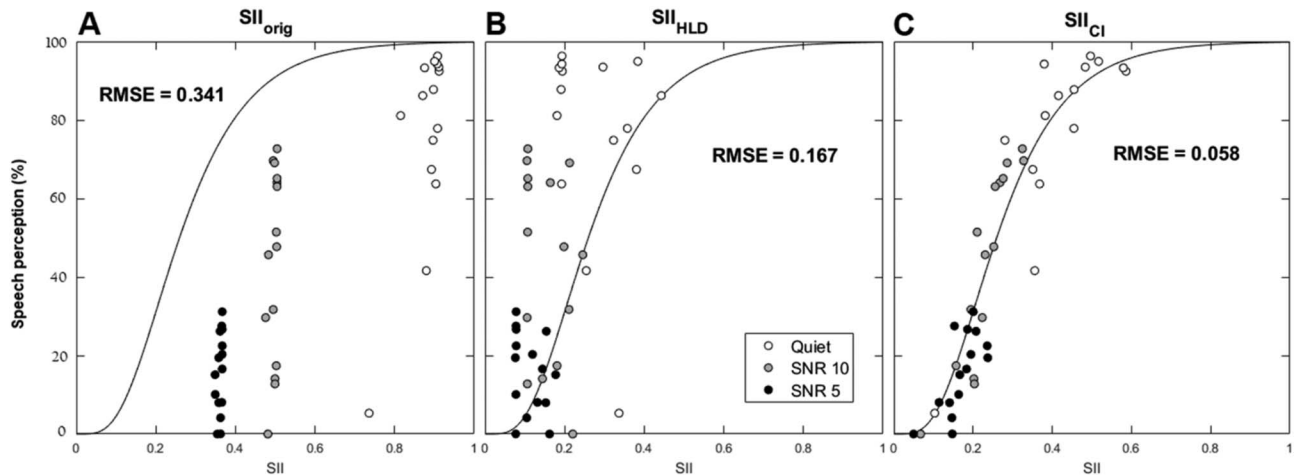
$$\text{Hearing loss desensitization} = 1 - \left(\frac{\text{PTA}}{108.3072}\right)^3. \quad (3)$$

Paired-sample  $t$  tests were conducted to compare conventional SII values and SII values with HLD correction (HLD SII) for the three conditions. There were statistical

**Figure 2.** Mean speech perception scores ( $\chi$ ) for the AzBio sentences in the three different signal-to-noise ratio (SNR) conditions.



**Figure 3.** (A) Score versus SII distribution scatter plots with the reference transfer function curve. (B) Score versus HLD SII distribution scatter plots with the reference transfer function curve. (C) Score versus CI-corrected scatter plots with the reference transfer function curve. RMSE denotes the root-mean-square error between the empirical data points and transfer function curve. CI-corrected SII values ( $SII_{CI}$ ) were computed by factoring in listeners with CI's duration of deafness, working memory, aided audibility, and Gap Detection Test scores. See Table 6 and the text for details. SII = speech intelligibility index; HLD = hearing loss desensitization; CI = cochlear implant; SNR = signal-to-noise ratio.



differences in the SII values with and without HLD corrections for all conditions: Quiet:  $t(14) = -21.958, p < .001$ ; SNR 10:  $t(14) = -23.743, p < .001$ ; and SNR 5:  $t(14) = -23.760, p < .001$ . Multiple comparisons showed that the HLD corrections reduced the original SII values for all listening conditions.

In contrast to the predictions from the conventional SII that excluded the influence of the hearing loss, the SII calculations using the HLD corrections resulted in lower SII values compared to the normative transfer function predictions in most cases (see Figure 3b), suggesting the normative transfer function curve was not suitable for predicting performance for listeners with CI even when the influence of hearing loss was considered. As was the case for SIIs calculated without the HLD factor, the SIIs with the HLD factor in Figure 3b also displayed a large individual variation in speech perception performance.

### Cognitive Function Tests

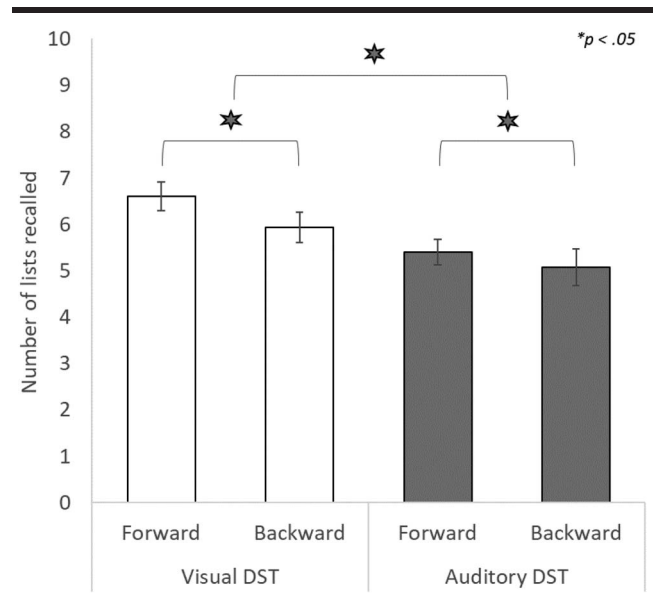
The DSTs administered in four different ways (visual/auditory, forward/backward) were further analyzed to compare differences in performance. Group mean maximum lengths of numbers correctly answered by 15 CI users are shown in Figure 4 (visual-forward DST,  $M = 6.6, SD = 1.18$ ; visual-backward DST,  $M = 5.93, SD = 1.27$ ; auditory-forward DST,  $M = 5.4, SD = 1.05$ ; auditory-backward DST,  $M = 5.06, SD = 1.53$ ). A two-way repeated-measures analysis of variance was conducted with the two presentation modalities (visual and auditory) and two reproduction orders (forward and backward) as the two within-subject variables and the maximum length of numbers as the dependent variable. Main effects were found for the presentation modalities,  $F(1, 14) = 14.252, p < .05$ , and reproduction orders,  $F(1, 14) = 6.563, p < .05$ , with no

interaction,  $F(1, 14) = 1.094, p = .072$ . CI users performed better on the forward and visual DSTs compared to backward and auditory DSTs, respectively.

### Correlation of Speech Perception Scores With Other Factors

Correlation analysis was conducted to investigate a relationship between predictive variables and speech perception performance. To perform the correlation analysis,

**Figure 4.** Mean number of lists correctly recalled for forward and backward Digit Span Tests (DSTs) presented with two different modalities (visual and auditory). Error bars denote  $\pm 1 SEM$ .





speech perception scores obtained from the three conditions were averaged to represent the broad concept of speech perception ability, yielding one dependent variable. We then assessed correlations between independent variables that were likely to be associated with speech perception scores, including both demographic (age, onset of deafness, and duration of deafness) and perceptual (SMRT, GDT, DST) measures. Onset of deafness was defined as the age at which hearing loss occurred. The demographic variables were primarily defined based on the history of the poorer ear. In order to extract a single value representing cognitive function, visual and auditory DST values used for this analysis were the mean of forward and backward DSTs for each modality, respectively. Unaided and aided audibility values were determined by averaging auditory thresholds at 0.5, 1, 2, and 4 kHz in the better ear. If patients with CI could not detect pure tones at the highest level presented in the unaided condition, thresholds were plotted as 100 dB at those frequencies. Lastly, the HLD SII and conventional SII were also included in the correlations. Like the speech perception scores, HLD SII and conventional SII obtained in the three different conditions were averaged, and individual HLD SII and conventional SII values were respectively obtained for the 15 participants with CI. Variables used for correlational analysis are shown in Table 3.

Pearson correlations were conducted (IBM SPSS, Version .24) to investigate the relationship between CI users' speech perception scores and their demographic, auditory processing, and cognitive function variables. The intercorrelations of the variables are shown in Table 4. We found strong, negative correlations between speech perception scores and duration of deafness, aided audibility, and GDT ( $r > .7$ ,  $n = 15$ ,  $p < .05$ ). Increases in speech perception scores were correlated with decreases in duration of deafness, aided hearing threshold represented as unaided

audibility, and GDT thresholds. A robust positive correlation between auditory DST and speech perception score was also found ( $r = .771$ ,  $n = 15$ ,  $p < .05$ ).

### Improving SII for Use in Listeners With CI

Our initial analysis revealed that the normative transfer function was not suitable for predicting performance for listeners with CI even when the influence of hearing loss (i.e., HLD) was considered (cf. Figures 3a and 3b). It is likely that different demographic and perceptual-cognitive factors may have contributed to the large variability in listeners with CI's speech perception and the stark mismatch between the empirical and predicted responses (i.e., Figure 3a). Consequently, we aimed to determine appropriate correction factors that might adjust the SII metric to create a CI-corrected version of the model (SII<sub>CI</sub>) and improve the fit between predicted and empirical data. The conventional SII and its modifications for hearing loss were not designed to predict speech performance in electrical hearing. Thus, we reasoned that incorporating important demographic and/or perceptual-cognitive measures that are known to vary across listeners with CI might improve the predictive power of the SII.

We examined several correction factors applied to listeners' raw SII scores that incorporated the most prominent demographic and perceptual-cognitive variables. To identify these best predictors, we first ran a stepwise multiple regression to isolate the factors most relevant to predicting speech perceptions scores and determine, empirically, the proper scaling coefficients to weight each variable in our SII<sub>CI</sub> model. Model selection was based on the regression model that maximized  $R^2$ . This identified three variables and their relative contribution for predicting speech perception scores where the sign of the coefficients

**Table 3.** Nine predictive variables and one dependent variable (speech perception test scores) for the multiple regression analysis.

No.	Onset of hearing loss (years)	Duration of deafness (years)	SMRT (RPO)	GDT (ms)	Unaided audibility (dB HL)	Aided audibility (dB HL)	Visual DST	Auditory DST	Speech perception scores (%)	HLD SII	SII
1	13	0	1.60	2.22	100.00	22.50	7	7	62.51	0.13	0.59
2	40	0	3.63	5.25	87.50	26.25	4.5	4	48.98	0.29	0.59
3	46	13	1.92	15.10	93.75	27.50	8	4.5	19.94	0.17	0.58
4	7	3	1.57	4.19	100.00	30.00	8.5	6.5	60.77	0.12	0.58
5	41	7	1.98	3.48	92.50	21.25	7	6	59.26	0.19	0.59
6	51	10	1.28	9.62	97.50	26.25	5.5	4	25.49	0.13	0.59
7	5	50	1.48	5.19	75.00	22.50	7	5.5	33.42	0.21	0.59
8	41	31	3.15	9.76	88.75	32.50	7	5.5	40.28	0.12	0.57
9	15	14	4.13	5.39	67.50	22.50	6	6	47.37	0.23	0.59
10	0	55	0.51	34.88	86.25	43.75	4.5	2.5	1.77	0.24	0.52
11	56	2	1.63	8.30	100.00	20.00	6	4.5	57.73	0.13	0.59
12	42	21	6.30	12.09	90.00	25.00	5	5	35.69	0.25	0.58
13	2	2	1.13	2.86	95.00	22.50	6	5	59.15	0.13	0.59
14	2	0	1.70	1.71	100.00	18.75	6	6	64.18	0.13	0.59
15	50	0	2.93	10.64	88.33	23.75	6	6.5	63.38	0.25	0.59

Note. SMRT = Spectral-Temporally Modulated in Ripple Test; RPO = ripples per octave; GDT = Gap Detection Test; DST = Digit Span Test; HLD = hearing loss desensitization; SII = speech intelligibility index.

**Table 4.** Pearson correlations between dependent measures.

Variable	1	2	3	4	5	6	7	8	9	10	11
Speech perception score 1	1	-.02	-.767*	.306	-.713*	.08	-.815*	.265	.771*	-.272	.697*
Onset of hearing loss 2		1	-.29	.167	-.162	.367	.015	-.08	-.128	.118	.281
Duration of deafness 3			1	-.534*	.626*	-.088	.647*	-.144	-.455	.261	-.689*
Unaided audibility 4				1	-.104	-.362	-.149	.203	.057	-.624*	.103
Aided audibility 5					1	-.171	-.836*	-.169	-.577*	.175	-.921*
SMRT 6						1	-.126	-.27	.160	.516*	.22
GDT 7							1	-.377	-.723*	.341	-.907*
Visual DST 8								1	.595*	-.567*	.277
Auditory DST 9									1	-.279	.602*
HLD SII 10										1	-.189
SII 11											1

Note. SMRT = Spectral-Temporally Modulated in Ripple Test; GDT = Gap Detection Test; DST = Digit Span Test; HLD = hearing loss desensitization; SII = speech intelligibility index.

\* $p < .05$ .

reflect either a positive or negative impact on  $SII_{CI}$ : GDT (temporal resolution), auditory DST (working memory), and years of deafness prior to implantation (deaf; see Table 5). In addition, we included aided audibility (AA) since pilot modeling indicated this variable substantially improved the match between predicted and empirical data. Scaling coefficients for GDT, duration of deafness, and auditory DST were taken from the coefficients identified via regression analysis (see Table 5), whereas the scaling of AA was set to 1. To ensure  $SII_{CI}$  was in the same range as the original SII (0–1), we found it necessary to scale all variables by 100 prior to computing  $SII_{CI}$ . Various combinations of these measures were then used to rescale listeners' original SII scores to create the  $SII_{CI}$ . We then assessed the relative benefit of each term by evaluating the RMSE between empirical speech perception responses and

**Table 5.** Best fitting model predicting speech perception scores from stepwise regression analysis.

Model	B	SE	$\beta$	p
Constant	24.364	17.449		
GDT	-0.6	0.471	-0.262	.23
Duration of deafness	-0.437	0.167	-0.419	.024
Auditory DST	6.161	2.789	0.39	.049

Note. The dependent variable was speech perception score,  $R^2 = .836$ , adjusted  $R^2 = .791$ . B = unstandardized coefficient;  $\beta$  = standardized coefficient; GDT = Gap Detection Test; DST = Digit Span Test.

the transfer function curve. Lower RMSE denotes better fit to the data.

CI-corrected models and their corresponding RMSE are shown in Table 6. Of the different corrections factors

**Table 6.** Comparisons between listeners with CI's empirical speech intelligibility scores and speech intelligibility index (SII) predictions incorporating different demographic and perceptual–cognitive predictors.

Model	RMSE
1. SII	0.341
2. $SII_{HLD}$	0.167
CI-corrected <sup>a</sup>	
3. $SII_{CI} = SII \frac{(6.16DST)}{100}$	0.117
4. $= SII \frac{(6.16DST - 0.60GDT)}{100}$	0.137
5. $= SII \frac{(6.16DST - 0.60GDT - 0.44deaf)}{100}$	0.166
6. $= SII \frac{(6.16DST - 0.60GDT - 0.44deaf + AA)}{100}$	0.058 <sup>b</sup>

Note. RMSE = root-mean-square error; SII = speech intelligibility index (normal hearing);  $SII_{HLD}$  = SII corrected for hearing loss desensitization; CI = cochlear implant;  $SII_{CI}$  = CI-corrected SII; DST = auditory digit span score; GDT = gap detection threshold (ms); deaf = duration of deafness (years); AA = aided audibility (dB HL).

<sup>a</sup>Scaling factors for GDT, deaf, and DST were based on the coefficients determined via multiple regression analysis (see Table 5). The 100 divisor term is used to ensure predicted  $SII_{CI}$  scores are bounded in the range of the original SII (0–1). <sup>b</sup>Best fitting model based on RMSE minimization (cf. Figure 3C).

we examined, incorporating GDTs, auditory DSTs, duration of deafness, and aided audibility provided the most accurate fit to listeners with CI's data (see Figure 3c). Other subsets of these predictors produced poorer fits. However, it should be noted that inclusion of auditory DST (i.e., working memory) scores alone proved to be one of the best weighting variables to correct the original SII for listeners with CI ( $RMSE_{SII_{CI}} = 0.117$  vs.  $RMSE_{SII} = 0.341$ ). Although alternate models are possible, these results suggest that including simple perceptual-cognitive measures (e.g., DST, GDTs), as well as important demographic variables (e.g., duration of deafness), can dramatically improve the predictive power of the SII for CI users.

## Discussion

### *Prediction of Speech Perception Scores Using SII*

One of our primary goals was to investigate whether the SII is an appropriate model to predict speech perception scores for adult CI users. We examined CI users' SII with and without HLD compared to the transfer function curve derived for adults with normal hearing. Two trends were apparent from our results. First, the transfer function curve tended to overestimate speech perception scores for CI users when the HLD factor was not taken into account, whereas the curve tended to underestimate speech perception when the HLD factor was applied to conventional SII calculations. This finding confirms, perhaps expectedly, that CI users' speech perception cannot be predicted using the existing SII model based on the transfer function from listeners with normal hearing. It is likely that severe-to-profound sensorineural hearing loss coupled with the limitations posed by electrical hearing could adversely affect SII prediction beyond just audibility. When taking HLDs into account, a decrease in SII values with HLD corrections for the CI users appears to be the reason why the prediction failed.

In general, HLD correction factors for modification of SII models are applied based on the hearing thresholds of the listeners. The amount of the correction factor increased with the greater hearing loss that these individual listeners have. This approach sometimes results in negative weights on SII values for people having very poor hearing thresholds. For example, an HLD factor proposed by Pavlovic et al. (1986) provides a desensitization factor of zero when hearing thresholds are above 94 dB HL. This eventually leads SII to 0 as a result of multiplication. This extreme application of correction factors has been criticized because individuals with thresholds of 94 dB HL would still be able to extract intelligible information from speech above 94 dB HL (Ching et al., 1998). Given the fact that CI candidates have mostly severe-to-profound hearing loss and the corresponding correction factors are substantially high, it is not surprising to see the underestimation effect observed in Figure 3b. In fact, the HLD correction was originally developed for the SII application in unaided conditions. Use of the HLD correction in aided conditions in

this study likely limited the SII's ability to show better predictive outcomes.

The other especially important observation is the variability seen in our listeners with CI's speech perception scores. Given the wide variation in individual speech perception outcomes among listeners with CI (Kiefer et al., 1998; Pisoni et al., 1999), the limitation of the conventional SII model to accurately predict speech perception was an anticipated result to some extent. Here, we took advantage of the variables that underlie such perceptual differences to propose a new CI-specific SII model ( $SII_{CI}$ ). Our model was developed by considering other characteristics of listeners with CI, which yielded an improvement in the fitting between predicted and observed speech perception scores. Variables that were highly correlated with empirical speech perception scores improved the accuracy of the SII's predictions. These findings are encouraging and promising for the applicability of CI-specific SII models to predict speech perception performance in the electrical hearing population. Our study is the first to investigate and provide supporting evidence that SII might be an effective tool with the caveat that studies must control (or at least consider) CI-related variables. Further research is required to strengthen the  $SII_{CI}$  model introduced here.

### *Prediction of Speech Perception Scores Using Perceptual and Cognitive Factors*

Ten independent variables considered as possible predictors of speech perception scores were included in the correlation analyses. The correlation coefficient of the group data from 15 adults with CI showed that duration of deafness, aided audibility, GDT, auditory DST, and SII were correlated with speech perception scores. Unexpectedly, several variables that we thought would be associated with speech perception performance were excluded from the best regression model. The regression analysis showed that auditory DST, GDT, and duration of deafness were the strongest variables associated with speech perception scores.

Working memory capacity represented by auditory DST showed high correlation with speech perception scores and contributed largely to our  $SII_{CI}$  model. It is also notable to look at the visual DST that showed relatively lower relationships with speech perception scores. This implies that CI users may have difficulty in restoring phonological structure from auditory digits, causing weakened phonological storage due to their limited auditory sensory input. It is reasonable to assume that speech perception and working memory capacity recruit similar underlying mechanisms to compensate for the poor auditory perception ability in people with CIs. Taken together, working memory does not play a major role in predicting speech perception unless the auditory modality is employed in CI users.

The high contribution of the GDT indicates that auditory processing abilities, especially temporal resolution, are highly associated with speech perception performance

in adults with CI. In contrast, the SMRT outcomes, which represent spectral acuity, did not show robust correlations with speech perception scores. The lack of any correlations between the SMRT outcomes and speech perception differs from previous studies that showed high correlations between these two measures (Lawler, Yu, & Aronoff, 2017; Litvak, Spahr, Saoji, & Fridman, 2007). The basis of this inconsistency is unclear, but possible explanations for this discrepancy may lie in the considerable degree of variability seen across the participants in this study. Individual differences in performance in populations with CI are typically observed in psychoacoustic experiments and in speech intelligibility tasks (Goldsworthy, Delhorne, Braida, & Reed, 2013). Even though an attempt was made to control for these effects during selection of participants with CI for this study, the small sample size ( $N = 15$ ) probably could not yield asymptotic performance on the SMRT tasks.

It is often stated that temporal resolution is comparatively better than spectral resolution in CI users. Indeed, we found that GDT performance for CI users was comparable to that of listeners with normal hearing (Goldsworthy et al., 2013; Shannon, 1989). CI users' poorer spectral resolution likely made it difficult to produce meaningful scores on the SMRTs in relation to speech perception scores. Taken together, our auditory processing outcomes suggested that temporal resolution is more associated with speech perception than spectral resolution for CI users. This observation supports the notion that CI users who are exposed to only limited spectral information rely heavily on temporal cues for speech perception (Kirby & Middlebrooks, 2010; Winn, Won, & Moon, 2016).

The high effect of duration of deafness on speech perception scores agrees well with reports from other studies that emphasize the prompt restoration of auditory feedback for prelingually deafened children with CI (Svirsky, Teoh, & Neuburger, 2004; Tong, Busby, & Clark, 1998). Even though most of the CI subjects in our study were postlingually deafened who are less likely to be affected by auditory deficits, duration of deafness was still the important factor in terms of speech perception outcomes. A consistent outcome was reported from Blamey et al. (1996) that examined some demographic variables for postlingually deafened CI users. These findings suggest that consistent auditory input is critical for speech perception regardless of whether or not the individual is deafened prelingually. Auditory neural circuits probably require continuous linguistic inputs for maintaining functional plasticity to decode speech inputs.

### ***Working Memory Capacity for CI Users***

To determine whether working memory capacity for CI users was affected by the deficits in auditory/phonological processing components, we administered the DST in two different modalities (auditory and visual presentations). Previous cognitive literature has shown mixed outcomes in terms of the superiority between two modalities. Some studies on human memory have shown that visual

memory is superior to auditory memory (Cohen, Horowitz, & Wolfe, 2009; Hilton, 2001). Hilton (2001) posited that visual (but not auditory) stimuli are stored in two different forms, mental image and repetition, which make them easier to recall than auditory information. She also noted that auditory processing may cause more fatigue than visual processing. On the other hand, other studies have claimed an auditory superiority effect with the assumption of higher strength of association between successive auditory stimuli compared to successive visual stimuli (Kemtes & Allen, 2008; Penney, 1989). These researchers also argued that visual stimuli likely give rise to more attentional load relative to auditory stimuli. This controversy can be seen in DST studies using CI subjects who have significant hearing problems. AuBuchon, Pisoni, and Kronenberger (2015) reported slightly higher performance for auditory DST than visual DST in a forward paradigm, but slightly lower performance for visual DST than auditory DST in a backward paradigm. This contradicts the results of Kronenberger, Pisoni, Henning, and Colson (2013) where visually presented stimuli resulted in slightly higher reproduction rates over auditorily presented stimuli in forward DSTs. Taken together, although there appears to be no definitive answer on this issue, it is clear that presentation modality plays a role in task performance.

Our DST results showed that CI users' performance on working memory tasks was better when they perceived stimuli visually rather than auditorily. The heavy demand on working memory load for processing the auditory stimuli may make it difficult to store the stimuli into short-term memory, resulting in such variance in performance. The other possibility is that the poorer performance on auditory tasks might have been caused by an auditory perception issue, not auditory processing or memory demands. Some CI users having very poor speech perception might have misunderstood the auditory stimuli, substantially affecting the group mean performance on the auditory DST. This assumption is supported by our finding of higher correlations of speech perception ability with auditory DST, compared to visual DST. However, because a normal hearing control group was not measured in this study and previous studies showed mixed outcomes, it would be difficult to argue that auditory deficits in CI users led to such differences in working memory function. Indeed, studies that assessed listeners with normal hearing and those with CIs using both auditory and visual modalities indicated that the poorer working memory function for CI users compared to those with normal hearing are not solely accounted for by their auditory perception or speech production abilities (AuBuchon et al., 2015; Cleary, Pisoni, & Geers, 2001; Kronenberger et al., 2013). Therefore, further investigations are needed to examine the mechanism of these two modalities in relation to performance on working memory capacity.

### ***Limitations***

The current study began with the idea of applying the conventional SII model, which is typically used for

people with normal hearing and mild to moderate hearing loss to CI users. We acknowledge that adopting the original SII model necessarily disregards some important aspects of CI users' perception. For instance, the SII model is based on acoustic levels at the eardrum, but the levels were inevitably determined near the microphone of the CIs in this study. Other distinct mechanisms of CI processing, such as significantly reduced dynamic range or frequency allocation, may cause the effects of nonlinearity that interfere with SII assumptions. Previous studies have found that listeners with CI may not optimally weight spectral information (Moberly et al., 2014), and the importance of individual channels in a CI are highly variable across subjects (Bosen & Chatterjee, 2016). Entirely new approaches that deal with such uniqueness of CIs should be developed and used for predictions of CI outcomes.

The primary shortcoming of this study is the small number of participants with CI ( $N = 15$ ). It is well known that large sample sizes are necessary for examining factors associated with experimental performance for CI users (Schafer & Utrup, 2016). Speech perception performance in individuals with CI varies considerably from person to person, and a large number of variables are associated with such variability. In addition, the small sample size may not have represented a typical population with CI. Recruitment challenges aside, larger samples may provide more reliable predictions in future studies and allow for the development of a more robust  $SII_{CI}$  model by taking into account other contributing factors to speech perception performance in CI users.

## Conclusion

This study investigated whether the SII could be a reliable predictor for speech perception performance in adults who use CIs. The speech perception scores for CI recipients obtained in three different SNR conditions yielded observed SIIs, which were compared to the predicted SIIs based on the transfer function curve for the AzBio test. Predictions of speech perception performance using the conventional SII calculation alone overestimated CI users' abilities, whereas SII calculations using HLD corrections underestimated performance. Furthermore, the large variability in speech perception performance across the CI users was shown to be a significant barrier for the SII to prove to be a reliable predictor. Other demographic and experimental variables associated with speech perception were used to construct a new SII model ( $SII_{CI}$ ). Models incorporating CI-relevant factors (i.e., GDT, duration of deafness, aided audibility and auditory DST) improved the fit between SII-predicted and observed scores. The results obtained with this initial sample of subjects suggest that conventional SII models are not appropriate for predicting speech perception scores for CI users without these additional demographic and perceptual weighting terms. Thus, the application of the  $SII_{CI}$  is based on the promise that clinics have a broad range of information about their patients. We believe that obtaining such information is certainly worthwhile,

considering the benefits of  $SII_{CI}$ . This new model may contribute to establishing realistic expectations and customized aural rehabilitation strategies for individual patients with CI. The long-term goal of future studies would include improving the  $SII_{CI}$  so that it can quantify or control the enormous individual variability seen in populations with CI. In addition, clinical applicability of the new  $SII_{CI}$  will eventually be investigated by replicating this preliminary study with a large group of patients with CI.

## Acknowledgments

This study was supported by the American Academy of Audiology/American Academy of Audiology Foundation Research Grant in Hearing & Balance Program (S. L.) and National Institute on Deafness and Other Communication Disorders Award R01DC016267 (G. M. B.). This study is a part of the first author's PhD dissertation. We would like to thank all the dissertation committee members, Eugene Buder and George Relyea, who contributed to discussing some of the issues raised in this study.

## References

- American National Standards Institute. (1997). *Methods for calculation of the speech intelligibility index* (ANSI S3.5-1997). New York, NY: Acoustical Society of America.
- American National Standards Institute. (1999). *American national standard maximum permissible ambient noise levels for audiometric test rooms ANSI S3.1-1999 (R2013)*. New York, NY: Acoustic Society of America.
- Aronoff, J. M., & Landsberger, D. M. (2013). The development of a modified Spectral Ripple Test. *The Journal of the Acoustical Society of America*, *134*(2), EL217–EL222.
- AuBuchon, A. M., Pisoni, D. B., & Kronenberger, W. G. (2015). Short-term and working memory impairments in early-implanted, long-term cochlear implant users are independent of audibility and speech production. *Ear and Hearing*, *36*(6), 733–737.
- Baddeley, A. (1993). Working memory and conscious awareness. In A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp. 11–28). Hillsdale, NJ: Erlbaum.
- Baddeley, A., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47–89.
- Bentler, R. A., & Pavlovic, C. V. (1989). Transfer functions and correction factors used in hearing aid evaluation and research. *Ear and Hearing*, *10*(1), 58–63.
- Bidelman, G. M., Jennings, S. G., & Strickland, E. A. (2015). PsychoAcoustX: A flexible MATLAB® package for psychoacoustics research [Technology report]. *Frontiers in Psychology*, *6*, 1498.
- Bidelman, G. M., Lowther, J. E., Tak, S. H., & Alain, C. (2017). Mild cognitive impairment is characterized by deficient brainstem and cortical representations of speech. *Journal of Neuroscience*, *37*(13), 3610–3620.
- Bidelman, G. M., Villafuerte, J. W., Moreno, S., & Alain, C. (2014). Age-related changes in the subcortical-cortical encoding and categorical perception of speech. *Neurobiology of Aging*, *35*(11), 2526–2540.
- Blamey, P., Arndt, P., Bergeron, F., Bredberg, G., Brimacombe, J., Facer, G., . . . Whitford, L. (1996). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants. *Audiology and Neurootology*, *1*(5), 293–306.
- Bosen, A. K., & Chatterjee, M. (2016). Band importance functions of listeners with cochlear implants using clinical maps.

*The Journal of the Acoustical Society of America*, 140(5), 3718–3727.

- Burkholder, R. A., & Pisoni, D. B.** (2006). Working memory capacity, verbal rehearsal speed, and scanning in deaf children with cochlear implants. In P. E. Spencer & M. Marschark (Eds.), *Advances in the spoken language development of deaf and hard-of-hearing children* (pp. 328–357). New York, NY: Oxford University Press.
- Bush, L. C.** (2016). List equivalency of the AzBio Sentence Test. *Honors Theses Paper*, 419.
- Ching, T. Y., Dillon, H., & Byrne, D.** (1998). Speech recognition of hearing-impaired listeners: Predictions from audibility and the limited role of high-frequency amplification. *The Journal of the Acoustical Society of America*, 103(2), 1128–1140.
- Cleary, M., Pisoni, D. B., & Geers, A. E.** (2001). Some measures of verbal and spatial working memory in eight- and nine-year-old hearing-impaired children with cochlear implants. *Ear and Hearing*, 22(5), 395–411.
- Cohen, M. A., Horowitz, T. S., & Wolfe, J. M.** (2009). Auditory recognition memory is inferior to visual recognition memory. *Proceedings of the National Academy of Sciences of the United States of America*, 106(14), 6008–6010.
- Collison, E. A., Munson, B., & Carney, A. E.** (2004). Relations among linguistic and cognitive skills and spoken word recognition in adults with cochlear implants. *Journal of Speech, Language, and Hearing Research*, 47(3), 496–508.
- Daya, H., Figueirido, J. C., Gordon, K. A., Twitchell, K., Gysin, C., & Papsin, B. C.** (1999). The role of a graded profile analysis in determining candidacy and outcome for cochlear implantation in children. *International Journal of Pediatric Otorhinolaryngology*, 49(2), 135–142.
- Draine, S.** (1998). Inquisit [Computer software]. Seattle, WA: Millisecond Software.
- Firszt, J. B., Holden, L. K., Skinner, M. W., Tobey, E. A., Peterson, A., & Gaggl, W.** (2004). Recognition of speech presented at soft to loud levels by adult cochlear implant recipients of three cochlear implant systems. *Ear and Hearing*, 25(4), 375–387.
- Fletcher, H., & Galt, R. H.** (1950). The perception of speech and its relation to telephony. *The Journal of the Acoustical Society of America*, 22(2), 89–151.
- Florentine, M., & Buus, S.** (1984). Temporal gap detection in sensorineural and simulated hearing impairments. *Journal of Speech and Hearing Research*, 27, 449–455.
- French, N. R., & Steinberg, J. C.** (1947). Factors governing the intelligibility of speech sounds. *The Journal of the Acoustical Society of America*, 19(1), 90–119.
- Fu, Q.-J., & Shannon, R. V.** (2000). Effect of stimulation rate on phoneme recognition by Nucleus-22 cochlear implant listeners. *The Journal of the Acoustical Society of America*, 107(1), 589–597.
- Geers, A. E., Pisoni, D. B., & Brenner, C.** (2013). Complex working memory span in cochlear implanted and normal hearing teenagers. *Otology & Neurotology*, 34(3), 396–401.
- Goldsworthy, R. L., Delhorne, L. A., Braid, L. D., & Reed, C. M.** (2013). Psychoacoustic and phoneme identification measures in cochlear-implant and normal-hearing listeners. *Trends in Amplification*, 17(1), 27–44.
- Gordon, K. A., Daya, H., Harrison, R. V., & Papsin, B. C.** (2000). Factors contributing to limited open-set speech perception in children who use a cochlear implant. *International Journal of Pediatric Otorhinolaryngology*, 56(2), 101–111.
- Green, K. M., Bhatt, Y., Mawman, D. J., O'Driscoll, M. P., Saeed, S. R., & Ramsden, R. T.** (2007). Predictors of audiological outcome following cochlear implantation in adults. *Cochlear Implants International*, 8(1), 1–11.
- Henry, B. A., & Turner, C. W.** (2003). The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. *The Journal of the Acoustical Society of America*, 113(5), 2861–2873.
- Henry, B. A., Turner, C. W., & Behrens, A.** (2005). Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners. *The Journal of the Acoustical Society of America*, 118(2), 1111–1121.
- Hilton, E.** (2001). Differences in visual and auditory short-term memory. *IU South Bend Undergraduate Research Journal*, 4, 47–50.
- Holden, L. K., Finley, C. C., Firszt, J. B., Holden, T. A., Brenner, C., & Potts, L. G.** (2013). Factors affecting open-set word recognition in adults with cochlear implants. *Ear and Hearing*, 34(3), 342–360.
- Kentes, K. A., & Allen, D. N.** (2008). Presentation modality influences WAIS Digit Span performance in younger and older adults. *Journal of Clinical and Experimental Neuropsychology*, 30(6), 661–665.
- Kiefer, J., von Ilberg, C., Reimer, B., Knecht, R., Gall, V., & Diller, G.** (1998). Results of cochlear implantation in patients with severe to profound hearing loss—Implications for patient selection. *International Journal of Audiology*, 37(6), 382–395.
- Kirby, A. E., & Middlebrooks, J. C.** (2010). Auditory temporal acuity probed with cochlear implant stimulation and cortical recording. *Journal of Neurophysiology*, 103(1), 531–542.
- Kronenberger, W. G., Pisoni, D. B., Henning, S. C., & Colson, B. G.** (2013). Executive functioning skills in long-term users of cochlear implants: A case control study. *Journal of Pediatric Psychology*, 38(8), 902–914.
- Lawler, M., Yu, J., & Aronoff, J. M.** (2017). Comparison of the spectral-temporally modulated ripple test with the Arizona Biomedical Institute Sentence Test in cochlear implant users. *Ear and Hearing*, 38(6), 760–766.
- Lee, S., & Mendel, L. L.** (2016). Effect of the number of maxima and stimulation rate on phoneme perception patterns using cochlear implant simulation. *Clinical Archives of Communication Disorders*, 1(1), 87–100.
- Lee, S., & Mendel, L. L.** (2017). Derivation of frequency importance functions for the AzBio sentences. *The Journal of the Acoustical Society of America*, 142(6), 3416–3427.
- Levitt, H.** (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 49(2B), 467–477.
- Litvak, L. M., Spahr, A. J., Saoji, A. A., & Fridman, G. Y.** (2007). Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. *The Journal of the Acoustical Society of America*, 122(2), 982–991.
- Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C.** (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences of the United States of America*, 103(49), 18866–18869.
- Ludvigsen, C.** (1987). Prediction of speech intelligibility for normal-hearing and cochlearly hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 82(4), 1162–1171.
- Macherey, O., & Carlyon, R. P.** (2014). Cochlear implants. *Current Biology*, 24(18), R878–R884.
- Moberly, A. C., Harris, M. S., Boyce, L., & Nittrouer, S.** (2017). Speech recognition in adults with cochlear implants: The effects of working memory, phonological sensitivity, and

- aging. *Journal of Speech, Language, and Hearing Research*, 60, 1046–1061.
- Moberly, A. C., Lowenstein, J. H., Tarr, E., Caldwell-Tarr, A., Welling, D. B., Shahin, A. J., & Nittrouer, S.** (2014). Do adults with cochlear implants rely on different acoustic cues for phoneme perception than adults with normal hearing? *Journal of Speech, Language, and Hearing Research*, 57, 566–582.
- Nie, K., Barco, A., & Zeng, F.-G.** (2006). Spectral and temporal cues in cochlear implant speech perception. *Ear and Hearing*, 27(2), 208–217.
- Oxenham, A. J., Bernstein, J. G., & Penagos, H.** (2004). Correct tonotopic representation is necessary for complex pitch perception. *Proceedings of the National Academy of Sciences of the United States of America*, 101(5), 1421–1425.
- Pavlovic, C. V.** (1987). Derivation of primary parameters and procedures for use in speech intelligibility predictions. *The Journal of the Acoustical Society of America*, 82(2), 413–422.
- Pavlovic, C. V., Studebaker, G. A., & Sherbecoe, R. L.** (1986). An articulation index based procedure for predicting the speech recognition performance of hearing-impaired individuals. *The Journal of the Acoustical Society of America*, 80(1), 50–57.
- Penney, C. G.** (1989). Modality effects and the structure of short-term verbal memory. *Memory and Cognition*, 17(4), 398–422.
- Pisoni, D. B., & Cleary, M.** (2003). Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation. *Ear and Hearing*, 24(1, Suppl.), 106S–120S.
- Pisoni, D. B., Cleary, M., Geers, A. E., & Tobey, E. A.** (1999). Individual differences in effectiveness of cochlear implants in children who are prelingually deaf: New process measures of performance. *The Volta Review*, 101(3), 111–164.
- Pisoni, D. B., & Geers, A. E.** (2000). Working memory in deaf children with cochlear implants: Correlations between digit span and measures of spoken language processing. *Annals of Otolaryngology, Rhinology & Laryngology*, 115, 92–93.
- Pisoni, D. B., Kronenberger, W. G., Roman, A. S., & Geers, A. E.** (2011). Measures of digit span and verbal rehearsal speed in deaf children after more than 10 years of cochlear implantation. *Ear and Hearing*, 32(1, Suppl.), 60S–74S.
- Schafer, E., & Ustrup, A.** (2016). The effect of age of cochlear implantation on speech intelligibility to others. *Journal of Educational, Pediatric & (Re)Habilitation Audiology*, 22, 1–11.
- Shannon, R. V.** (1989). Detection of gaps in sinusoids and pulse trains by patients with cochlear implants. *The Journal of the Acoustical Society of America*, 85(6), 2587–2592.
- Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., & Ekelid, M.** (1995). Speech recognition with primarily temporal cues. *Science*, 270(5234), 303–304.
- Sherbecoe, R. L., & Studebaker, G. A.** (2003). Audibility-index predictions of normal-hearing and hearing-impaired listeners' performance on the connected speech test. *Ear and Hearing*, 24(1), 71–88.
- Smith, Z. M., Delgutte, B., & Oxenham, A. J.** (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 416(6876), 87–90.
- Spahr, A. J., & Dorman, M. F.** (2005). Effects of minimum stimulation settings for the Med El Tempo+ speech processor on speech understanding. *Ear and Hearing*, 26(4), 2S–6S.
- Studebaker, G. A., Gray, G. A., & Branch, W. E.** (1999). Prediction and statistical evaluation of speech recognition test scores. *Journal of the American Academy of Audiology*, 10(7), 355–370.
- Studebaker, G. A., Sherbecoe, R. L., McDaniel, D. M., & Gray, G. A.** (1997). Age-related changes in monosyllabic word recognition performance when audibility is held constant. *Journal of the American Academy of Audiology*, 8(3), 150–162.
- Svirsky, M. A., Teoh, S. W., & Neuburger, H.** (2004). Development of language and speech perception in congenitally, profoundly deaf children as a function of age at cochlear implantation. *Audiology and Neurootology*, 9, 224–233.
- Tong, Y. C., Busby, P. A., & Clark, G. M.** (1998). Perceptual studies on cochlear implant patients with early onset of profound hearing impairment prior to normal development of auditory, speech, and language skills. *The Journal of the Acoustical Society of America*, 84, 951–962.
- Vandali, A. E., Whitford, L. A., Plant, K. L., & Clark, G. M.** (2000). Speech perception as a function of electrical stimulation rate: Using the Nucleus 24 cochlear implant system. *Ear and Hearing*, 21(6), 608–624.
- van Dijk, J. E., van Olphen, A. F., Langereis, M. C., Mens, L. H., Brokx, J. P., & Smoorenburg, G. F.** (1999). Predictors of cochlear implant performance. *International Journal of Audiology*, 38(2), 109–116.
- Wechsler, D., & De Lemos, M. M.** (1981). *Wechsler Adult Intelligence Scale-Revised*. San Diego, CA: Harcourt Brace Jovanovich.
- Winn, M. B., Won, J. H., & Moon, I. J.** (2016). Assessment of spectral and temporal resolution in cochlear implant users using psychoacoustic discrimination and speech cue categorization. *Ear and Hearing*, 37(6), e377–e390.
- Won, J. H., Drennan, W. R., & Rubinstein, J. T.** (2007). Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. *Journal of the Association for Research in Otolaryngology*, 8(3), 384–392.
- Xu, L., & Zheng, Y.** (2007). Spectral and temporal cues for phoneme recognition in noise. *The Journal of the Acoustical Society of America*, 122(3), 1758–1764.