Cross-frequency Integration for Consonant and Vowel Identification in Bimodal Hearing

Ying-Yee Kong[a),b),c)] and **Louis D. Braida**[c)]

[b)] Department of Speech Language Pathology & Audiology, Northeastern University, Boston, MA 02115

[c)] Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139

Abstract

Purpose—Improved speech recognition in binaurally combined acoustic-electric stimulation ("bimodal hearing") could arise when listeners integrate speech cues from the acoustic and electric hearing. The aims of this study are (1) to identify speech cues extracted in electric hearing and residual acoustic hearing in the low-frequency region and (2) to investigate cochlear-implant (CI) users' ability to integrate speech cues across frequencies.

Method—Normal-hearing (NH) and CI subjects participated in consonant and vowel identification tasks. Each subject was tested in three listening conditions: CI alone (vocoder speech for NH), hearing aid (HA) alone (low-passed speech for NH), and both. Integration ability for each subject was evaluated using a model of optimal integration – PreLabeling Integration model [L.D. Braida, Q. J. Exp. Psychol. A **43**,647-677 (1991)].

Results—Only a few CI listeners demonstrated bimodal benefit for phoneme identification in quiet. Speech cues extracted from the CI and the HA are highly redundant for consonants, but complementary for vowels. CI listeners also exhibited reduced integration ability for both consonant and vowel identification compared to their NH counterparts.

Conclusion—These findings suggest that reduced bimodal benefits in CI listeners are due to insufficient complementary speech cues across ears, a decrease in integration ability, or both.

I. Introduction

Cochlear implants (CIs) have evolved from being a supplemental aid to speechreading with a single-channel system to an auditory aid that provides sufficient speech cues for users to enjoy high levels of speech recognition without visual cues. As a result, audiological requirements for implant candidacy have been relaxed from profound to moderately-severe hearing loss. Many recently implanted users have some degree of low-frequency residual hearing. Patients with greater residual hearing in the low frequencies could benefit from a short-electrode CI array to preserve the residual hearing in the implanted ear (hybrid

[a)] Corresponding author: Address: Department of Speech Language Pathology & Audiology, 106A Forsyth Building, Northeastern University, Boston, MA 02115, USA; Tel: +1 (617) 373-3704; yykong@neu.edu.

hearing). More commonly, patients with more severe hearing loss are implanted with a long-electrode array in one ear and use a hearing aid (HA) in the opposite ear (bimodal hearing). While this supplement has been found to be beneficial to speech recognition when combined with electrical stimulation, the types of speech cues in the low-frequency residual acoustic hearing and the abilities of CI listeners to extract and integrate these cues with the electrical signals are not well understood.

Some studies demonstrated significantly better speech recognition performance in quiet with bimodal hearing compared to CI alone (Shallop, Arndt, & Turnacliff, 1992; Dooley et al., 1993; Armstrong, Pegg, James, & Blamey, 1997; Ching, Psarros, Hill, Dillon, & Incerti, 2001; Gifford, Dorman, McKarns, & Spahr, 2007a; Zhang, Dorman, & Spahr 2010), but others have reported no significant combined benefit in the majority of their subjects (Hamzavi, Pok, Gstoettner, & Baumagartner, 2004; Dunn, Tyler, & Witt, 2005; Mok, Grayden, Dowell, & Lawrence 2006), and even incompatibility between the two devices in some patients (Tyler et al., 2002). Results for speech recognition in noise, on the other hand, are more consistent across studies. Many studies reported combined acoustic-electric benefit for both bimodal and hybrid hearing when both speech and noise were presented from the front (e.g., Ching, Incerti, & Hill, 2004; Turner, Gantz, Vidal, Behrens, & Henry, 2004; Kong, Stickney, & Zeng, 2005; Mok et al., 2006; Gifford et al., 2007a; Dorman, Gifford, Spahr, & McKarns, 2008; Mok, Galvin, Dowell, & McKay, 2009; Zhang et al., 2010). However, the amount of benefit still varied among listeners. The sources of inter-subject variability in bimodal/hybrid outcomes are unclear. It is often assumed that the greater the amount of residual hearing, the greater the bimodal benefit. However, Ching et al. (2004) and Gifford et al. (2007a) failed to find significant correlation between the unaided threshold below 1000 Hz and the amount of bimodal benefit. Some researchers measured other aspects of auditory function, such as frequency selectivity and modulation detection in patients who are considered candidates for a CI, but found no correlation between residual functional abilities and speech recognition (Gifford, Dorman, Spahr, & Bacon, 2007b).

The possible underlying mechanisms for improved speech recognition performance in bimodal hearing include: (1) better detection of the target speech for sentence recognition in noise by glimpsing the target during the spectral and/or temporal dips of the masker (Kong and Carlyon, 2007; Li and Loizou, 2008; Brown and Bacon, 2009a); and (2) integration of speech cues between electric stimulation and acoustic stimulation at low frequencies (e.g., Ching et al., 2001; Ching, van Wanrooy, & Dillon, 2007; Kong and Carlyon, 2007). While the glimpsing mechanism has been studied in recent years, speech cue integration in bimodal hearing has not received much attention. In the present study, we investigated the ability of bimodal subjects to integrate speech cues across ears and across devices. The variability in the amount of combined benefit (Hamzavi et al., 2004; Dunn et al., 2005, Mok et al., 2006) could be attributed to the variability in the integration ability among CI users.

Integration of speech cues from multiple modalities or sources (Massaro, 1987, 1998; Braida, 1991; Grant and Seitz, 1998; Massaro and Cohen, 2000; Grant, 2002. Ronan, Dix, Shah, & Braida, 2004) has been researched and discussed extensively. Several quantitative models have been developed to characterize the processes of multimodal integration of speech segments and to predict integration performance based on observed performance for each separate modality. These models make predictions based on the observed confusion matrix (Miller and Nicely, 1955) for each separate source. The models assume that cues in each source are statistically independent from each other, combined without interference, and no new cues arise from inter-source comparisons. Among them, the PreLabeling (PreL) Model of Integration developed by Braida (1991) predicts combined-source scores using an "optimal" decision rule that assumes perfect integration of the cues derived from each separate source. The difference between observed combined scores and predicted combined

scores obtained by this ideal observer model allows us to evaluate a listener's ability to integrate cues from various sources independent of their ability to extract cues from each separate source. Detailed description of the PreL model is provided in Section II.B. Although the integration models were developed to describe perceptual integration based on audiovisual speech reception research, they have been formulated abstractly and have been shown to be capable of describing the integration of auditory cues across spectral bands (Ronan et al., 2004; Grant, Tufts, & Greenberg, 2007). In Ronan et al. (2004), the evaluation of predictions by a number of models for across frequency integration was performed with speech cues combined from different frequency bands. They first divided broadband speech into five discrete frequency bands (0-700 Hz; 700-1400 Hz; 1400-2100 Hz; 2100-2800 Hz; and 2800-4500 Hz) and then combined the bands that are adjacent to each other or remotely apart. The number of frequency bands was one of the testing variables. They concluded that both the PreL integration model and the Fuzzy Logical Model of Perception were able to predict performance for various combinations of frequency bands tested, with only a few exceptions.

To receive maximal benefit from multi-sensory or multi-source stimulation, listeners need to be able to extract *complementary* cues from each source and to integrate the cues from all sources (Braida, 1993). The PreL model conceptualizes cue extraction and cue integration as two *independent* processes. Therefore, predictions from the model will allow us to examine each factor separately:

> (1) Information extraction: The type and amount of speech cues provided by each source of stimulation can affect the degree of combined-source benefit. When comparing two different stimulation conditions, Grant, Walden, & Seitz (1998) demonstrated that the condition resulting in the highest recognition scores does not necessarily result in the highest combined score or the greatest improvement when combining the sources. Braida (1993) described two extreme cases: a) the sources provide cues that complement one another; and b) the cues provided by each source are completely redundant with one another. If speech cues in each stimulation condition are redundant, combined-source benefits will be less than if the cues are complementary.

> (2) Information integration: Cue extraction and integration are considered independent processes in integration models described above. According to these models, it is possible to measure multi-sensory integration that is independent of abilities to extract cues from each source. In other words, listeners' ability to integrate cues in a multi-source condition is not affected by the type and amount of cues extracted from each source. All other things being equal, it is assumed that greater ability to integrate speech cues from different sources will yield better performance in the combined-source condition.

Within this framework, the variability in bimodal outcomes (particularly in quiet) among CI users could be attributed to a decreased ability to extract speech cues, reduced ability to integrate speech cues, or both. This has been shown in cross-frequency integration in hearing-impaired (HI) and elderly listeners (Palva and Jokinen, 1975; Turner, Chi, & Flock, 1999; Healy and Bacon, 2002; Grant et al., 2007). Using the PreL model, Grant and colleagues (2007) examined cross-frequency cue extraction and cue integration abilities in a group of HI listeners. They reported that HI listeners showed proportionally less benefit than normal-hearing (NH) listeners when additional high-frequency cues were added to the low-frequency speech. They attributed the reduced benefit to HI listeners' difficulty extracting cues from the highest frequency band (4762-6000 Hz) when presented concurrently with the lower-frequency band (1890-2381 Hz), as well as their inefficiency in integrating speech cues across frequency regions.

The present study is the first attempt at applying an integration model to understand the process by which speech cues are integrated in combined acoustic-electric stimulation. We first measured phoneme (both consonants and vowels) recognition performance in bimodal CI listeners to examine the redundancy of speech cues extracted from electric stimulation and acoustic stimulation in the low-frequency region, and then used the model-based approach to investigate and compare NH and CI listeners' ability to integrate cross frequency speech cues.

II. Experiment I: Cross-frequency integration in NH listeners

The PreL integration model has shown the ability to predict performance for auditory-visual integration (Braidia, 1991; Braida, 1993; Grant et al., 1998) and auditory integration across frequencies for consonant identification (Ronan et al., 2004; Grant et al., 2007). In this experiment, we set out to further evaluate the application of the PreL model to predict a combined-band performance that was created to simulate bimodal hearing, in which low-frequency (<1000 Hz) speech was delivered to one ear and wideband (200-6000 Hz) vocoder-processed speech to the opposite ear.

Preliminary study: Before the start of this experiment, we conducted a preliminary study which assessed the model fit to a unique combined-band performance that was not examined in Ronan et al. (2004). In Ronan et al.'s study, multi-source speech cues came from non-overlapping frequency bands. The purpose of this preliminary study was to demonstrate that the PreL model is capable of predicting combined-band performance when the speech signals from different frequency regions undergo different signal processing (low-pass filtering with cutoff frequency at 500 Hz and channel vocoding above 900 Hz). Since this preliminary study is not the focus of this paper, a brief description of the methodology and results from our model predictions are included in the Appendix.

A. Methods

1. Subjects—A total of eight NH subjects (7 females), aged 19 to 31 years participated in the study. Six of them participated in the consonant identification task. Six subjects participated in the vowel identification task. Four of these six subjects also participated in the consonant identification task.

2. Stimuli—Two sets of speech stimuli were used. The first set consisted of 16 consonants / p,t,k,b,d,g,f,θ,s,ʃ,v,ð,z,ʒ,m,n/ used by Miller and Nicely (1955) in the /aCa/ context. These stimuli were recorded by five male and five female talkers by Shannon, Jensvold, Padilla, Robert, and Wang (1999). The second set consisted of nine monophthongs / i,ɪ,ɛ,æ,ɝ,ʌ,u,ʊ,ɔ/ in the /hVd/ context recorded from five male and five female talkers in our laboratory. All stimuli were scaled to the same overall root-mean-squared (RMS) level. For each stimulus set, recordings from two male and two female talkers were used in the practice sessions, and recordings from the remaining six talkers were used in the test sessions. Three utterances of each consonant and vowel from each talker were used.

Recorded stimuli were subjected to two types of processing (1) low-pass (LP) filtering and (2) channel-vocoding. LP was performed using Butterworth filters with a roll-off of 60 dB/ octave and a cutoff frequency of 500 Hz. These LP parameters mimicked a sloping hearing loss above 500 Hz, an audiometric configuration commonly seen in real bimodal CI users. Channel-vocoding processing preserved speech cues from 200 to 6000 Hz. The vocoder system simulated listening with a long-electrode array CI device. In this system, broadband speech (200 – 6000 Hz) was first processed through a pre-emphasis filer and then band-pass filtered into four logarithmic frequency bands. The lower, center, and upper frequencies of

the band-pass filters for this 4ch vocoder condition are listed in Table I. The amplitude envelope of the signal was extracted from each band by full-wave rectification and LP with a 400 Hz cutoff frequency. Sinusoids were generated with amplitudes equal to the RMS level of the envelope and frequencies equal to the center frequencies of the band-pass filters. The sinusoids were then summed and presented to the listeners.

3. Procedures—Subjects were tested in a double-walled soundproof booth. Stimuli were presented from a sound card using 16-bit resolution at 44.1k Hz sampling rate to Sennheiser HD 265 headphones. Each subject was presented with two different speech signals: low-frequency (LP speech) to one ear and vocoder speech to the opposite ear. Three types of listening conditions were tested: low-frequency speech alone (LP-alone), vocoder speech alone (vocoder-alone), or LP and vocoder speech combined (LP + vocoder). For both consonant and vowel identification tasks, half of the subjects were presented with vocoder stimuli to the left ear and the LP stimuli to the right ear, the remaining half received the stimuli on the opposite sides. All stimuli were presented at an RMS level of 70 dBA.

For each condition, listeners first received practice trials identifying the consonant and vowel with visual correct-response feedback provided. Performance usually reached a plateau (i.e., within 3 percentage points difference) within three blocks of practice. If not, additional practice was given until the criterion was met. Each subject was then tested in blocks of 96 trials (16 consonants × 6 talkers) for consonant identification, and 54 trials (9 vowels × 6 talkers) for vowel identification. Each utterance from each talker was used for three blocks of testing. A total of nine blocks of testing were presented in each test condition, yielding a total of 54 trials (6 talkers × 9 blocks) per stimulus per condition per subject. No feedback was provided during test sessions. The order of presentation of the listening conditions was counterbalanced across subjects. Stimuli within each block were presented in random order. A list of 16 /aCa/ or nine /hVd/ syllables was displayed on a computer screen and subjects responded by clicking a button corresponding to the syllable they heard. Consonant and vowel confusion matrices were constructed from each subject's responses.

B. Data analysis and model fits

The mean overall percent (%) correct scores were calculated for consonant and vowel identification tasks. In addition, based on the confusion matrices of the group data, information transmission (Miller and Nicely, 1955) was computed for the features of voicing, manner of articulation, and place of articulation for consonant identification for each test condition. This was also done for the features of height, back, and tense for vowel identification. The different consonant and vowel features are listed in Table II.

Model predictions for combined scores were made for each subject using the PreL model of integration (Braida, 1991). The model first analyzes observed confusion matrices for each separate source of information (i.e., LP speech vs. vocoder speech), and then makes predictions for the confusion matrix when these sources are presented simultaneously, using an "optimal" decision rule that assumes perfect integration of available cues in an ideal observer without any bias. The PreL model is a special form of multidimensional scaling (MDS) and a multidimensional extension of the signal detection theory. Unlike traditional MDS, the scaled distances between stimuli in separate source spaces are converted into a common metric d′. It is assumed that there is a D-dimensional vector of cues $\vec{X} = \langle x_1, x_2,\ldots, x_D \rangle$ associated with each presentation of one of the N possible consonants S_i. The cue vector X is described by the conditional probability density

$$p\left(\vec{X}|S_i\right)=\prod_{k=1}^{D}p\left(x_k|S_i\right)=\prod_{k=1}^{D}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(x_k-s_{ik})^2}{2}\right)$$

(1)

and may be thought of as displaced from the stimulus center $\vec{S}=\langle s_{i1}, s_{i2}, ..., s_{iD}\rangle$. Corresponding to each response there is a response center or prototype $R=\langle r_{i1}, r_{i2}, ..., r_{iD}\rangle$. The decision processes assumes a comparison between the stimulus attributes (i.e., the observed vector of cues $\vec{X}$) and the response center $\vec{R}_i$ in memory. The listener is assumed to respond $\vec{R}_k$ if and only if the distance from the observed vector of cues $\vec{X}$ to $\vec{R}_k$ is smaller than the distance to any other prototype, i.e., $|\vec{X}-\vec{R}_k| < |\vec{X}-\vec{R}_j|$. A listener's sensitivity $d'(i,j)$ in distinguishing stimulus S_i from stimulus S_j is given by

$$d'(i,j)=\left\|\vec{S}_i-\vec{S}_j\right\|=\sqrt{\Sigma_{k=1}^{D}\left(s_{ik}-s_{jk}\right)^2}$$

(2)

where $\|\vec{S}_i-\vec{S}_j\|$ is the distance between the D-dimensional vector of cues generated by stimuli S_i and S_j. The predictions for the multi-source condition (i.e., dichotic presentation of LP and vocoder speech) are made based on the performance in the single-source conditions (i.e., LP-alone and vocoder-alone). In the multi-source condition, the model assumes that cues in each source are statistically independent from each other, combined without interference, and no new cues arise from intersource comparisons. Integration of cues is modeled by assuming that the cue densities are the "Cartesian products" of the densities corresponding to the separate sources. In the multi-source condition, the cue space has dimension $D_{AB} = D_A + D_B$ and each stimulus center $\vec{S}_i$ has the coordinates

$$\vec{S}_i=\langle s_{i1}, s_{i2}, ..., s_{iD_{AB}}\rangle=\langle a_{i1}, a_{i2}, ..., a_{iD_A}, b_{i1}, b_{i2}, ..., b_{iD_B}\rangle.$$

(3)

This model for the multi-source condition predicts that there is a simple Pythagorean relationship between a subject's sensitivity in the multi-source condition, $d'_{AB}(i,j)$, and the corresponding source sensitivities $d'_A(i,j)$, and $d'_B(i,j)$:

$$d'_{AB}(i,j)=\sqrt{d'_A(i,j)^2+d'_B(i,j)^2}.$$

(4)

In this model, the configurations derived from the single-source confusion matrices determine the predicted stimulus centers for the multi-source configurations. Prediction accuracy in the multi-source condition also requires a specification of the response centers (i.e., the prototypes). In the PreL model described by Braida (1991), the response centers of the multi-source condition were assumed to coincide with the multi-source stimulus centers. Recent work by Ronan et al. (2004) has evaluated the locations of the response centers in the multi-source case for across-frequency consonant identification in NH listeners. In that study, cross-frequency consonant identification performance in the multi-band case was made with three different assumptions of the location of the multiband response centers:

(A) PreLI0: the response centers in the multi-source case are the Cartesian products of the response centers in the single-source case. As mentioned in Ronan et al. this assumption represents a case of minimal adjustment(s) to the multi-source stimulation condition, in which the response centers are located at the same locations they had in the single-source conditions. This implies that when provided with multi-source

stimulation, listeners elicit a response that concurs to the response from a single source, i.e. HA-alone or CI-alone conditions in the present study.

(B) PreLI1: the response centers in the multi-source case coincide with the multi-source stimulus centers. This assumption predicts optimal integration of cues from different sources and predicts maximum overall performance. As Grant and colleagues (1998, 2007) pointed out, since the PreL model is an optimal integration model, predicted scores should always be higher than or equal to the observed scores in real listeners if no new cues arise from the simultaneous presentation of multiple sources.

(C) PreLIH: the response centers in the multi-source case are half-way between the response centers in the single-source case and multi-source stimulus centers. As pointed out by Ronan et al. (2004), this represents an intermediate case of adjustment to the multi-source condition in which the response centers are halfway between the response centers in case A and case B.

In this study, we made the predictions for the combined LP and vocoder performance using these three assumptions to examine where listeners place their multi-source response centers. For example, if the listener's multi-source response centers coincide with the multi-source stimulus centers, the PreLI0 method and the PreLIH method will provide predictions that underestimate the observed scores, while the PreLI1 model will provide predictions equal to or slightly higher than observed scores. If listeners' multi-source response centers are located at the same locations as in one of the single-source conditions, suggesting a potential reduced integration ability or response bias, the PreLI0 assumption will provide predictions close to observed scores while the PreLIH and the PreLI1 assumptions will overpredict by a greater extent. The evaluation of the model fit with these three assumptions will be based on (1) overprediction or underprediction of the observed data, and (2) the amount of error between predicted and observed scores. Given that the PreLI1 model is an ideal observer model, a model fit is expected to over-estimate real human observers' scores. Thus, the model that estimates the observed scores with smaller amount of error is considered the best fit compared to those that produce greater error between the predicted scores and the observed scores.

C. Results

1. Phoneme Identification and Information Transmission

1.1 Consonant Identification: Figure 1 shows the mean overall % correct consonant identification (left) for three listening conditions. Paired-t tests on the arcsine transformed data revealed a significant combined benefit of 5.6 percentage points compared to vocoder alone in NH listeners [paired-t(5)=7.37, p<0.001], suggesting that NH listeners are able to integrate LP and vocoder speech cues across ears.

Percent information transmission for the consonant features voicing, stop, nasal, and fricative, and place of articulation was calculated for the NH group data (Fig. 2 left panel). The group data were computed from confusion matrices combined across subjects. First, there was substantial cue redundancy between LP-alone and vocoder-alone stimuli. Unlike the complementary cues provided by the auditory (voicing and manner of articulation) and visual (place of articulation) stimuli in combined auditory-visual stimuli, the LP speech provided cues -- mostly voicing and manner of articulation, that were largely redundant to those delivered by the vocoder speech. However, all features, except for place of articulation, were transmitted above 80% by the vocoder. Second, the presence of combined benefit of LP+vocoder speech over vocoder speech alone is noticeable for voicing, fricative, and place of articulation features, but not for stop and nasality, perhaps due to a ceiling effect. The amount of improvement ranged from 5 to 8 percentage points. This pattern of results is similar to that found in our preliminary study (Appendix). Taken together, these

results suggest that NH listeners are able to combine consonant features across ears and across frequencies to improve their overall consonant identification performance as reported in Ronan et al. (2004).

1.2 Vowel Identification: Overall mean vowel identification for NH listeners was computed and compared between combined and vocoder-alone conditions (see Fig. 1). Paired-t tests on the arcsine transformed data revealed a significant combined benefit of 15.6 percentage points compared to vocoder-alone conditions for vowel identification [paired-t(5)=9.60, p<0.001;], suggesting that NH listeners are able to integrate LP and vocoder speech cues across ears for better identification of vowels. The amount of the combined benefit was greater for vowel identification than consonant identification.

Percent information transmission was computed for three vowel features (Chomsky and Halle, 1968): height (high, mid, low), back (front, central, back), and tense (tense, lax). Acoustically, the first formant (F1) is associated with the height of the vowel (F1 increases as the vowel height decreases). F1 frequencies for American vowels are generally below 1000 Hz for adult talkers (Hillenbrand, Getty, Clark, & Wheeler, 1995), within the range of the LP stimuli. Results from the acoustical analysis of our stimuli are consistent with this finding. The second formant (F2), on the other hand, is associated with the backness of the vowel (F2 decreases as the production of the vowel moves towards the back of the vocal tract), which normally at a higher frequency region (>900 Hz). (Hillenbrand et al., 1995). This was also confirmed by the acoustical analysis on our stimuli. The difference between tense and lax vowels is based on an articulatory criterion of muscular tenseness or laxness. Acoustically, tense and lax vowels generally differ in terms of their durations and F1 frequencies. Figure 2 (right panel) shows the average percent information transmission for these three vowel features. It is not surprising that only a small amount of information about the feature back was transmitted by the LP speech, which only contained frequencies up to 1000 Hz, below the frequency range of F2 in most vowels. About 31% of information regarding vowel height was delivered to the listeners with the LP speech. A closer examination of the confusion matrix reveals that confusions occurred more frequently between high lax vowels (e.g., /ɪ,ʊ/) and mid vowels (/ɛ,ɝ,ʌ,ɔ/) and between mid vowels and low vowels /æ/. In these cases, the F1 frequencies of these vowels were generally closer to each other, suggesting that listeners were able to distinguish the vowel height when F1 differences are large. Information transmission for the feature of tense by the LP stimuli was 74%, significantly higher than the feature of vowel height. This may be because listeners can use both durational and F1 cues to distinguish tense vowels from lax vowels. Percent information transmission for vowel height for 4ch-vocoder speech was 39%, for the vowel feature back was 48%, and for vowel tense was 60%.

Like the consonant identification results, the combined benefit of LP + vocoder speech compared to vocoder-alone is seen for the three vowel features evaluated in this study. The combined benefit was 26, 13, 16 percentage points for vowel features height, back, and tense, respectively. It is noted that this pattern of results is also similar to that in our preliminary study. These findings suggest that NH listeners are able to combine vowel features across ears and across frequencies to improve their overall vowel identification performance.

2. Model predictions for NH listeners—Integration ability across frequencies was evaluated using the PreLI0, PreLI1, and PreLIH integration models that differ in the location of the response centers in the multi-source condition described in Ronan et al. (2004).

2.1 Consonant Identification: Predictions were made separately for each subject and condition. They were computed by first fitting the vocoder-alone and LP-alone matrices in

D = 3 dimensions and then predicting the scores for the combined condition from a 6-dimensional model. Figure 3 (upper left panel) shows the predicted versus observed combined-source consonant identification scores for each response center location (triangles: PreLI0; squares: PreLI1; circles: PreLIH). The unity-slope line represents a perfect match between predicted and observed scores. Points falling above this line indicate that the predicted scores are better than the observed scores and vice versa. Since PreLI1 made predictions using an "optimal" decision rule that assumes perfect integration of cues from each source, predicted scores are expected to be equal to or greater than observed scores. As expected, PreLI1 constantly overpredicted the observed combined performance (by an average of 9.1 percentage points). PreLI0 consistently underpredicted the combined performance by an average of 7.2 percentage points, while PreLIH overpredicted the combined performance by an average of 4.6 percentage points. The deviation between predicted and observed scores, calculated as root-mean-squared-error (RMSE), was greater for PreLI0 (8.7 percentage points) than for PreLIH (5.1 percentage points). In general, this pattern of results is similar to those reported in Ronan et al. (2004) for cross-frequency integration (across or within the same ear) and results from our preliminary study for consonant identification in NH listeners, which also showed that PreLI0 underestimated the multiband scores and that PreLIH provided better predictions compared to PreLI1 (i.e., overpredicted to a lesser extent) for cross-frequency consonant identification. The consistent underestimation of combined performance by PreLI0 and the more accurate predictions (i.e., smaller errors between predicted and observed scores) by PreLIH suggests that NH listeners indeed integrated cues from both LP and vocoder speech and that the locations of the response centers in the combined condition were different than the original LP-alone or vocoder-alone response centers.

2.2 Vowel Identification: Predictions for vowel identification were made in the same way as for consonants for each subject and combined condition. The vocoder-alone and LP-alone matrices were first fit with three dimensions and then the combined scores were fit with a 6-D model. Figure 3 (upper right panel) shows the predicted versus observed combined-source vowel identification scores for each response center location (triangles: PreLI0; squares: PreLI1; circles: PreLIH) (note that some data points overlap on this graph). Consistent with consonant identification results, PreLI1 constantly overpredicted the observed combined performance (by an average of 7.0 percentage points). PreLI0 consistently underpredicted the combined performance by an average of 6.9 percentage points, while PreLIH overpredicted the combined performance by an average of 3.6 percentage points. The deviation (RMSE) between predicted and observed scores was greater for PreLI0 (8.4 percentage points) than for PreLIH (4.5 percentage points). This pattern of results is, again, consistent with that found in our preliminary results for the vowel identification task (see Appendix).

II. Experiment II: Cross-frequency integration in bimodal CI users

There has been no systematic investigation that examined CI listeners' ability to *extract* and *integrate* speech cues across electric and acoustic stimulation. Ching et al. (2001) calculated information transmission for consonant recognition, but did not investigate vowel recognition. Mok et al. (2006) used CNC phoneme recognition scores to calculate information transmitted in individual ears and combined hearing, and reported that differences in scores between bimodal hearing and CI alone were greatest in phonemes containing relatively low-frequency cues compared to phonemes with high-frequency cues. In the present study, CI listeners were tested with the same stimuli and procedures as those used for NH listeners to facilitate comparisons of performance between the two groups. In addition, we acquired a large number of repeated measures (42-54 trials) for each bimodal CI listeners per stimulus per listening condition, which was rarely done in pervious studies.

This is necessary in order to minimize estimation bias for the analysis of information transmission (Sagi and Svirsky, 2008).

A. Methods

1. Subjects—Twelve CI subjects (C1-C12, 7 females, 5 males) aged 15 to 69 years (mean 35.75 years) participated in the consonant identification task, and half of them (C2, C5, C7, C8, C9, and C12) also participated in the vowel identification task. Table III shows detailed demographic information for each subject, including age, onset of hearing loss, etiology of hearing loss, duration of severe-to-profound hearing loss prior to implantation, and the CI processor used. Seven subjects are under the age of 30, close to the age range of the NH subjects in experiment 1. The remaining subjects are in the age range from 46 to 69 years old. Seven subjects were congenitally hard-of-hearing bilaterally (C4, C5, C7, C9, C11) or diagnosed with bilateral hearing loss at a very young age (C3, and C8) but had enough residual hearing to receive benefits from HA use before they received a CI. For these seven subjects, hearing loss was progressive from mild to severe-to-profound in three subjects (C3, C7, C8). The other four subjects had a severe hearing loss at birth (C4, C5, C9, C11). Two subjects were congenitally deaf or acquired hearing loss at a young age in only one ear (C10 and C12). They subsequently acquired hearing loss in the other ear later in life (C10 at age 7; C12 at age 43). All subjects use oral communication and have developed normal language skills. Speech production is highly intelligible in all subjects. All subjects wore their HA in the implanted ear before their implant surgery (except for C10) and they all continued to wear their HA in the non-implanted ear on a daily basis after implantation. Figure 4 shows the unaided (upper panel) and aided (lower panel) thresholds in the non-implanted ear for each individual, except for subject C6 due to lack of audiological data and the lack of audiometric equipment on the day of the test. The threshold data was either obtained from subjects' most recent audiologic examination or obtained on the same day of the testing in our laboratory. A large variability in the amount of residual hearing is evident in our subject group. While some subjects (e.g., C2) have mild-to-severe hearing loss at the low frequencies (<1000 Hz), some (e.g., C11) have profound loss even in the lowest frequencies. With amplification, all subjects have aided thresholds in the mild to moderate hearing loss range below 1000 Hz.

2. Stimuli and Procedure—The unprocessed consonant and vowel stimuli in experiment 1 were used for CI listeners. Each subject was evaluated under three listening conditions: HA-alone, CI-alone, and combined use of a CI and a HA (CI+HA). All stimuli were presented via a loudspeaker one meter directly in front of the subject at a fixed level of 65 dBA. Subjects used their own CI and HA settings during the entire test session, except for the volume setting in the HA. Subjects adjusted the volume of their HAs until the presented stimuli reached their comfortable listening level. Those who could not achieve the comfortable level with the HA alone were asked to adjust the volume of their HA to the maximum setting before any distortion occurred. With this setting, they reported that the presented speech stimuli were just slightly below the comfortable level. The same HA and CI settings were used for the CI+HA condition for each subject. Given that the purpose of this study was to investigate listeners' ability to integrate speech cues, we used subjects' everyday device settings to minimize novelty effects, which are likely to affect integration ability.

Subjects were asked to turn off their CI when tested on the HA-alone condition. They were instructed to turn-off their HA but leave their ear-mold in place in the non-implanted ear when tested on the CI-alone condition. A foam earplug was inserted in the implanted ear during testing regardless of the listening condition to prevent any potential acoustic stimulation in case residual hearing is preserved in that ear. To verify that the ear-mold and

the foam earplug provided sufficient attenuation to completely eliminate any acoustic stimulation, prior to the experiment, each subject was instructed to turn off both the HA and CI, and have both the ear-mold and the foam earplug inserted in their ear canals. Test speech stimuli (10 tokens) were then presented at 65 dBA via the loudspeaker. Under this condition, none of the CI listeners reported hearing any sound.

As in Experiment 1, each CI listener first received practice identifying the consonants and vowels with visual correct-response feedback provided for each test condition. Performance usually reached plateau (i.e., within 3 percentage points difference) within three blocks of practice; additional practice sessions were given until the criterion was met. After the practice sessions, each CI listener was then given nine blocks (3 blocks for each of the three utterances) of testing for each test condition yielding a total of 54 trials (6 talkers × 9 blocks) per stimulus per condition per subject, except for subjects C1 and C6 who only were tested with seven blocks (42 trials per stimulus per condition) due to time constraints. No feedback was provided during test sessions. The order of testing for the CI-alone and CI+HA conditions were counterbalanced across subjects (i.e., half of the subjects were tested with CI-alone condition first and vice versa). The HA-alone condition was tested last. This was done intentionally to minimize any anxiety arising if the subjects performed poorly with their HA alone.

B. Results

1. Phoneme Identification and Information Transmission

1.1 Consonant Identification: Overall consonant identification was calculated for each subject and listening condition (Fig.5, upper panel). There was a large inter-subject variability in scores, ranging from near-chance level 15-16% correct (C8 and C11) to a high level performance of 78% correct (C3) for the HA-alone condition, and from 32% (C12) to 87% correct (C5) for the CI-alone condition. Two subjects, C8 and C11 could not perform the task with HA alone: their scores were near chance levels of 16%. The majority of the subjects showed significantly better consonant identification with CI alone than with HA alone (p<0.05), except for C4 and C12. Subject C4 showed no significant difference in performance between the HA-alone and CI-alone conditions [t(16) = 1.94, p = 0.07]. Subject C12 performed significantly better with HA alone than with CI alone [t(16) = 14.18, p < 0.001]. Unlike the patterns of results obtained in NH listeners, the majority of CI listeners did not show a bimodal benefit (i.e., the CI+HA consonant identification scores were not significantly different from the scores for the better ear), except for C3. The bimodal benefit (CI+HA vs. CI-alone) noted in C3 was 3.8 percentage points. Subject C9, however, showed a significant decrease in performance (4.1 percentage points) with combined CI and HA use compared to CI alone [t(16) = 3.95; p<0.005]. Significant difference between the CI+HA score and the better ear score is marked with an asterisk (*) on the figure.

The group mean data is also shown in Fig. 5. We excluded subject C12 in the group analysis due to her atypical pattern of results. While consonant identification was better in the CI +HA condition than in the CI-alone condition in C12, her bimodal hearing performance was no better than her HA-alone performance. The inclusion of this subject in the group analysis could produce a false impression that there was a bimodal benefit compared to listening to a single device (HA alone or CI alone).

Percent information transmission for consonant features of voicing, stop, nasal, and fricative, and place of articulation was calculated for each subject and for the group data (Table IV). The group data was computed from confusion matrices combined across subjects, excluding subject C12. The HA provided information about voicing and nasality but very limited information about stops, fricative and place of articulation. On average

(excluding C12), CI alone provides 56% voicing information which is consistent with findings reported in the literature (e.g., Fishman, Shannon, & Slattery, 1997;Ching et al., 2001). Unlike the patterns of results from NH listeners which showed improvement in the combined conditions compared to vocoder-alone on some consonant features, there was no sizable bimodal benefit for all consonant features in the CI group, except for 8.9 points improvement in the nasality feature.

1.2 Vowel Identification: Overall vowel identification scores were calculated for individual and group mean data (excluding C12) for each listening condition (Fig. 5 lower panel). Only one subject, C12, showed better vowel identification with HA alone compared to CI alone [t(16) = 16.68, p<0.001]. The rest of the group showed a reverse pattern. Unlike the lack of bimodal benefits for consonant identification, half of the subjects tested (C2, C5, and C8) showed a bimodal benefit of 3.5 – 6.8 points for vowel identification compared to CI alone or HA alone (p<0.05). The bimodal benefit does not seem to correlate with the amount of residual hearing in the non-implanted ear. For example, subjects C8 and C9 had almost identical hearing loss from 250 to 1000 Hz, C8 showed bimodal benefit, but C9 did not.

Percent information transmission was calculated for the three vowel features: height, back, and tense for individual and group data (excluding C12) for each listening condition (Table V). The type and amount of information transmitted by the HA were very similar to those obtained in LP-alone condition in NH listeners. CI listeners also performed similarly to NH listeners listening to 4ch-vocoder stimuli. As a group (excluding C12), HA alone provided more information on vowel height (F1) (27%) but essentially no information on the feature back (F2) (4%). CI alone provided more information on the feature back (69%) than on vowel height (53%). This suggests that HA provided complementary information to a CI for vowel identification. Independent of listening condition, vowel tense was better identified than vowel height, suggesting CI listeners' ability to use multiple cues (duration and F1) to distinguish between tense and lax vowels. The three subjects (C2, C5, and C8) who demonstrated a significant bimodal benefit are the only ones who showed better CI+HA performance compared to CI-alone performance on all three vowel features. The lack of improvement between CI+HA and CI-alone or HA-alone on vowel features height and back was seen for subjects (C7, C9, and C12) who did not show overall bimodal benefit in vowel identification.

The information transmission analysis provides one test of whether performance of the NH subjects who listened to vocoder speech is a reasonable acoustic model to approximate performance in CI subjects. The patterns of results with vocoder speech alone obtained from NH listeners for both consonant and vowel identification are consistent with those obtained with the CI listeners in this study, as well as results reported in the CI literature (Ching et al., 2001; Mok et al., 2006). For consonants, both vocoder speech and CI alone provided more information on features voicing and manner of articulation than on place of articulation. For vowels, both vocoder speech and CI alone provided more information on the feature back than on vowel height.

2. Model predictions for CI listeners—The ability to integrate electric and acoustic speech cues in CI listeners was evaluated using the PreLI0, PreLIH, and PreLI1 models and results are compared to those obtained from NH listeners. Similar to the fitting procedures used for the NH data, CI data was first fit with a 3-D model on the HA-alone and CI-alone matrices and then fit with an 6-D model for the bimodal performance for both consonant and vowel identification.

2.1 Consonant Identification: Figure 3 (lower left panel) shows predictions for individual CI subjects for the bimodal hearing condition. Triangles represent predictions from PreLI0,

squares represent predictions from PreLI1, and the circles represent predictions from PreLIH. Note that we were unable to model the data for subject C2. In order to apply the model to the data, errors must be made in identifying the stimuli. This subject made no errors for one stimulus. Our modeling results show that the predictions of the PreLI0 model were either equal to or just slightly higher (1.3 – 4.6 points) than the combined scores for eight subjects. It only underpredicted the combined scores for three subjects – C7, C8, and C11 by 6.1, 6.3, and 9.7 percentage points, respectively. This is very different than the patterns of results seen in NH listeners in which PreLI0 almost always underpredicted the combined score (upper left panel). PreLI1 and PreLIH on the other hand, overpredicted the combined score by a greater amount, ranging from 8.5 points (C5) to 35.2 points (C1) for PreLI1, and 5.2 points (C11) to 21.8 points (C1) for PreLIH. In sum, the data for the CI listeners, unlike NH listeners, were better fit with response centers located somewhat nearer to the original response centers in the single-source condition. This suggests that the majority of the CI listeners (except for three subjects: C7, C8, and C11) made minimal adjustment to the bimodal hearing condition and their responses to the bimodal stimulation were similar to the responses made to the HA-alone or CI-alone condition; whichever produced higher overall performance.

2.2 Vowel Identification: Figure 3 (lower right panel) shows the predicted versus observed combined vowel scores for six CI listeners using PreLI0 (triangles), PreLI1 (squares), and PreLIH (circles) models. While PreLI0 consistently underpredicted the bimodal performance in NH listeners, it produced mixed results for the CI listeners – an underprediction for half of the six subjects tested (C2, C5, and C8). It should be noted that the three subjects (C2, C5, and C8) for whom PreLI0 underpredicted combined scores are the same subjects who achieved significantly higher overall vowel identification scores with bimodal hearing compared to CI alone. The rest of the subjects (C7, C9, C12) who did not receive a bimodal benefit for vowel identification were also best fit by PreLI0 in which the predictions were either equal to (C7) or just slightly higher (2.3 points for C9 and 4.6 points for C12) than the observed combined scores. This suggests that these three subjects may not have been able to integrate as well as the other subjects who showed significant bimodal benefit for vowel identification. Both PreLIH and PreLI1 overpredicted the combined scores for all subjects by an average of 6.6 and 9.0 percentage points, respectively.

III. Discussion

A. Information transmission by HA and CI

Information transmitted via the HA and the CI is largely redundant for consonants. Both HA and CI provided greater voicing and manner of articulation information, but less place of articulation information. Note that seven of the 12 CI subjects tested achieved voicing scores below 40%, considerably lower than the voicing score obtained from the NH group when only low frequencies were presented (LP-alone condition). One possibility for the reduced HA performance is that the HA fitting in our CI subjects may not be optimal. More likely, the difference in the voicing score is due to the presence of severe to profound hearing loss in the CI listeners while NH listeners have normal low-frequency hearing. At first, the low voicing scores in the impaired ears may seem surprising given that voicing cues are preserved at the low frequencies. However, similar results have been reported in the literature on HI individuals with substantial hearing loss at low frequencies (Boothroyd, 1984; Ching et al., 2001; McDermott, Dorkos, Dean, & Ching, 1999). McDermott et al. (1999) tested five adults with sensorineural hearing loss on phoneme, word, and sentence recognition tasks. All subjects had moderate to profound sensorineural hearing loss from 500 to 1000 Hz and a profound loss above 1000 Hz. Two subjects had a mild loss at 250 Hz and the rest had moderate to moderately-severe loss at that frequency, very similar to our CI

listeners in the non-implanted ear. With conventional amplification, three of their subjects showed voicing scores below 40%, consistent with our findings. Ching et al. (2001) investigated a group of 16 bimodal CI children (age ranged from 6 to 18 years) who had severe to profound hearing loss in the non-implanted ear. They reported a low voicing score in the HA-alone condition with an average voicing score of about 36% in quiet. Careful examination of our data revealed that voicing confusions occurred mostly for fricative consonants. For example, /f/ and /θ/ were frequently mis-identified as voiced stop /b/ or voiced fricatives /v,ð,z/; /s/ and /ʃ/ were mis-identified as voiced stop /d/ or voiced fricatives /ð,z,ʒ/. The reason for the poor voicing perception in our subjects is unclear.

Unlike consonants, information transmitted via the HA and the CI is somewhat complementary for vowels. The HA received mostly F1 information, but the CI received more F2 information than F1. This finding is consistent with results reported in the literature. Mok et al. (2006) tested a group of bimodal users on CNC words in quiet. The degrees of hearing loss in the non-implanted ears of their subjects are similar to our subjects. They reported about 14 percentage points better F1 transmission than F2 on the HA side, and about 8 percentage points better for F2 than F1 on the CI side.

B. Bimodal benefits

While all NH listeners showed improved consonant and vowel identification in the LP +vocoder condition compared to the vocoder-alone condition, only a few CI listeners showed a bimodal benefit. This pattern of results has been reported previously for bimodal hearing users (Ching et al., 2001; Dunn et al., 2005; Mok et al., 2006) as well as for hybrid users for consonant identification (Reiss, Gantz, & Turner 2008). Ching et al. (2001) reported only three out of 11 subjects showed significant bimodal benefit for speech recognition in quiet with HAs that were adjusted to provide the target insertion gain using the prescriptive NAL-RP method. Dunn et al. (2005) and Mok et al. (2006) reported that only four out of 12 subjects and two out of 14 subjects showed significant bimodal benefit for CNC word recognition and CNC phoneme recognition in quiet, respectively.

Reiss et al. (2008) tested consonant discrimination on 20 hybrid users. Their subjects had substantially better residual hearing than our bimodal CI subjects. They were implanted with short-electrode arrays: the most apical electrode (electrode 6) encoded frequencies down to 688 Hz, not covering the entire speech frequency range. Interestingly, their results showed that only about one quarter of the subjects demonstrated substantial improvement with combined acoustic-electric (A+E) conditions compared to A-only or E-only conditions (Fig. 2B and Fig. 3A in that paper) for consonant discrimination. This suggests that the reduced benefit observed in our bimodal users is not likely due to overlapping frequency ranges between electric and acoustic stimulation, or the severity of the hearing loss in the acoustic ear in our bimodal CI subjects.

Our modeling results and the estimates of information transmission for various consonant and vowel features provide an insight on the mechanisms that underlie this deficit. The lack of bimodal benefits for consonant identification in CI listeners can be attributed to a combination of factors, including: (1) insufficient, misrepresented, or redundant information provided by an individual ear, and (2) reduced ability to integrate speech cues across ears. The CI data were best fit with the PreLI0 model that assumes a minimal adjustment in response from the single-source conditions to the multi-source condition. This indicates that CI listeners performed sub-optimally compared to their NH counterparts. For consonant identification, all subjects (except for C3) showed an absence of bimodal benefit with bimodal scores essentially the same as the CI-alone or HA-alone scores; whichever was higher. For example, subjects C4 and C10 performed similarly with HA- and CI-alone to the

LP- and vocoder-alone observed in NH listeners, but they did not obtain the bimodal benefits that NH listeners did.

While the PreLI0 model produced a good fit for many of the CI subjects for consonant identification, it slightly underpredicted the combined scores for subjects C7, C8, and C11. However, none of these subjects showed overall bimodal benefits for consonant identification, which could be attributed to the insufficient or redundant cues provided by the HA ear to improve their performance in the bimodal listening condition. Subjects C8 and C11 achieved overall scores of 15-16% correct, close to chance level performance (6.25%). Subject C7 performed better than C8 and C11 with HA alone, but his performance, particularly the extraction of voicing, stop, and fricative features, was still significantly substantially lower than the LP-alone performance in NH listeners.

Subject C3 showed a bimodal benefit for consonant identification despite the fact that his bimodal performance was best predicted by the PreLI0 model. Unlike other CI listeners, C3 could identify consonants in the HA-alone condition with an overall score of 78% correct, 34 percentage points higher than the LP-alone score in NH listeners, while his CI alone performance was comparable to the NH 4ch-vocoder performance. Despite his high level of performance on the HA condition, he only received a bimodal benefit of 4.4 points. His reduced bimodal benefit may be attributed to his sub-optimal integration ability to combine speech cues from both ears.

The relationship between model fit and bimodal benefit is more apparent and direct in the vowel identification task. Five out of the six CI subjects tested, except for C12, had similar vowel identification performance with HA alone and with CI alone. However, only three subjects (C2, C5, and C8) showed improved overall vowel identification scores. Interestingly, these three subjects also exhibited an ability to integrate speech cues: their combined scores were underpredicted by the PreLI0 model.

C. Cross-frequency integration deficits in HI listeners

The present study shows a reduced benefit from additional low-frequency speech cues for phoneme identification in bimodal CI listeners. This finding is consistent with Grant et al. (2007) who reported that HI listeners received less benefit than NH listeners when additional high-frequency cues were added to low-frequency speech in the auditory-alone condition, but their performance significantly improved and achieved a similar level of performance as NH listeners when the high-frequency cues were provided visually in the auditory-visual condition. They attributed these patterns of performance to HI listeners' reduced efficiency in integrating auditory speech cues across spectral regions compared to NH listeners, but similar efficiencies in integrating auditory-visual speech cues. In addition to the interplay between extraction and integration of cues, Grant et al. (2007) discussed other explanations that may underlie the reduced benefits of additional high-frequency speech cues in HI listeners, which may also apply to the CI population:

> (1) *Peripheral interference – masking*. Cross-frequency integration in the same ear in HI listeners could have been adversely affected by excessive peripheral upward spread of masking due to the broadening of auditory filters. However, the masking effect could not explain the reduced integration ability in our bimodal CI subjects because the low-frequency and high-frequency speech cues were presented to separate ears.

> (2) *Perceptual saliency of additional cues*. Turner and Henry (2002) reported that HI listeners can benefit from high-frequency speech cues when consonant recognition performance is relatively poor at low signal-to-noise ratios (SNRs). At this poor performance level, the additional high-frequency speech cues may help to decipher the more "difficult" features of speech. In our study, all of our CI subjects (except for C1)

achieved greater than 50% consonant identification with CI alone (or with HA alone for subject C12). The additional low-frequency cues from the non-implanted ear in these subjects may be insufficient to decipher the more "difficult" features, such as voicing and place of articulation. To test this hypothesis, future studies could deliver noise in the CI ear to decrease the CI-alone performance to below 50% in order to examine if the addition of low-frequency speech cues could help improve performance. The exception from subject C1, who achieved only 41% correct with CI alone but did not receive bimodal benefit, poses a challenge to this explanation.

(3) *Perceptual bias*. Ross, Saint-Amour, Leavitt, Javitt, & Foxe (2006) reported that auditory-visual integration is most effective for intermediate SNRs. In high or low SNRs, listeners may have a strong bias towards the cues from the dominant modality and ignore cues from the other modality instead of integrating available cues from both modalities. This form of "interference", as Grant et al. (2007) pointed out, would not be accounted for by the PreL model. Our CI listeners may have had a strong perceptual bias toward their CI ear and may have ignored cues from the HA ear because the HA provided weaker cues than the CI. However, evidence from our CI listeners who did not show ear dominance argues against this explanation. Subjects C3 and C4 showed a relatively good consonant identification score in the HA-alone condition, not significantly different than the CI-alone performance. For these two subjects, their reduced bimodal benefit is not likely to be attributed to a perceptual bias toward one ear, although it is still possible that the cues from one ear are ignored because they are redundant, as in the consonant stimuli. For vowel identification, however, half of the CI subjects showed bimodal benefit, suggesting that they did not ignore cues from the less dominant ear when the ear provided complementary cues.

(4) *Age effects and internal noise*. Although we did not investigate age effects in this study, we made an effort to recruit CI subjects from different age groups, ranging from 15 to 69 years old. A study by Souza and Boike (2006) showed that age impaired listeners' ability to integrate temporal-envelope cues across frequency bands for speech recognition. Seven of 12 CI subjects tested in the consonant identification task were under the age of 30, within the age range of the NH listeners. The differences in the combined benefit between the two groups cannot be accounted for by age. However, all of our younger subjects and some of the older subjects (e.g., C8 and C12) had an onset of hearing loss in at least one ear at a very young age (< 5 years). Although all subjects communicated orally prior to implantation and have developed essentially normal language and reading skills, it is possible that the degraded speech they received since childhood may have affected their development of a normal internal representation of the phonemes compared to their NH counterparts, which in turn might impair their ability to integrate speech cues and to identify stimuli in the combined case. The imperfections of the internal representation of the stimulus due to the excess sensory noise and memory noise (Sagi, Meyer, Kaiser, Teoh, & Svirsky, 2010) in the single-source and/or combined-source case can also occur in post-lingually deafened adults, especially after long-duration hearing loss. This form of deficit would also be viewed as sub-optimal integration across frequencies and modalities. Future investigation on response error patterns to individual stimuli and detailed comparisons between the perceptual and stimulus space may provide insight into the underlying cause for the reduced bimodal benefits in CI listeners. Additionally, investigations with non-speech stimuli and/or with different tasks may further reveal the limiting factors for cross-frequency integration in CI listeners.

(5) Deficits in the across-frequency processing of temporal speech cues. As suggested by Grant et al. (2007), the reduced integration efficiency for HI listeners may not be simply the product of degraded extraction or integration of cues across frequency bands,

but instead be a true deficit in dealing with temporal speech cues such as reported in Healy and Bacon (2002). Healy and Bacon presented two speech-modulated tones at 750 and 3000 Hz to NH and HI listeners. In one condition, the low-frequency band led or lagged the high-frequency band by 12.5- 100 ms. They found that NH listeners could tolerate small disruptions in across-frequency timing (i.e., 12.5 ms) but performance decreased as the cross-frequency asynchrony increased. However, HI listeners' performance dropped more precipitously than their NH counterparts. For consonants, cross-frequency speech cues may not be presented concurrently, e.g., stop voicing relies heavily on voice onset time (VOT) cues, (the time interval between burst and the onset of voicing). The reduced bimodal benefit in our CI listeners may be due to their difficulty in comparing temporal speech cues across frequencies. Their ability to integrate speech cues across ears and across frequency bands could also be hindered by the processing delays between their HA and CI, as well as differences in processing time between frequency channels within a single device.

While the deficit in spectro-temporal processing may seem a likely explanation for the integration deficits in consonant identification in CI listeners, it is still unclear how this can account for CI users' reduced bimodal benefit in vowel identification. Unlike the dynamic properties of consonants, spectral cues for vowel recognition are relatively static. In addition to durational cues (which we did not control for in this study), identification of vowels was primarily based on the steady concurrent cues for F1 and F2 and does not require integration of cross-frequency temporal speech cues. Therefore, the lack of bimodal benefits in vowel identification may reflect a deficit in cross-frequency integration in some of our CI subjects.

In addition to the deficit in processing temporal cues across frequencies, bimodal CI users may also encounter a problem of incompatibility of the two devices given that one ear receives acoustic stimulation and the other ear receives electric stimulation. Although all of our CI subjects reported fused auditory images and enjoyed the sound quality better when using both devices, we cannot exclude the possibility that differences in stimulation modes (acoustic vs. electric) and sound quality may also pose difficulties for integration and create potential interference across ears.

D. Future directions

The present study does not suggest that CI users do not receive bimodal or hybrid benefits. It has been well documented that bimodal and hybrid users can achieve better word and sentence recognition (e.g., Turner et al., 2004; Kong et al., 2005; Mok et al., 2006; Gifford et al., 2007a; Dorman et al., 2008; Reiss et al., 2008; Mok et al., 2009; Zhang et al., 2010). For word and sentence recognition, listeners could be better able to use phonotactic, prosodic, and contextual cues when low frequencies are presented (Brown and Bacon, 2009b; Spitzer et al., 2009; Zhang et al., 2010). A recent report by Zhang et al. (2010) argues that bimodal benefit for word recognition in quiet can be accounted for entirely due to the presence of fundamental frequency cues alone and that the benefit was not significantly different between conditions where the non-implanted ear received only very low-frequency (<125 Hz) versus wideband speech. Also, it has been suggested that the addition of low-frequency cues can aid listeners in glimpsing the target sentence during the spectral and/or temporal dips of the masker (Kong and Carlyon, 2007; Li and Loizou, 2008; Brown and Bacon, 2009a; Zhang et al., 2010).

The lack of improvement in consonant identification in our CI listeners should not discourage the use of bimodal hearing. Most of our CI subjects exhibited severe to profound hearing loss at the low frequencies, a greater loss than some of CI candidates under the current FDA guidelines for cochlear implantation. It is possible that bimodal users will receive more benefits for phoneme recognition if they have more residual hearing in the

non-implanted ear. Also, our results clearly showed the benefits of using F1 cues from the HA to improve vowel identification in at least some of our CI listeners. F1 formant transitions also provide cues for consonant recognition and the use of these cues has not been fully investigated. In addition, there is evidence in our study suggesting that listeners who have reduced integration ability could still benefit from bimodal stimulation if they receive useful cues from the HA (e.g., C3). Future research could look into ways to improve HA performance to deliver more useful speech cues.

IV. Conclusions

(1) PreL model of integration is capable of predicting combined-band performance with speech signals that have undergone different signal processing (LP filtering and channel vocoding). Results of model fits for NH listeners for consonant and vowel identification were consistent with those reported in Ronan et al. (2004).

(2) When tested with their everyday program/map, bimodal CI listeners received largely redundant information between the two devices for consonants. Both HA and CI provided greater voicing and manner of articulation information, but less place of articulation information. However, information received between HA and CI was somewhat complementary for vowels with HA providing mostly F1 information and CI providing more F2 information.

(3) While NH listeners showed significant improvement in a combined LP+vocoder condition for both consonant and vowel identification, the majority of CI listeners did not show bimodal benefits compared to the performance in the better ear for consonant identification and only half of the CI listeners showed bimodal benefits for vowel identification.

(4) Predictions from the PreL model were different between NH and CI groups. NH consonant and vowel identification performance was best fit with the PreLIH assumption, in which the response centers in the multi-source case were located half-way between the response centers in the single-source case and multi-source stimulus centers. However, the majority of the CI data was best fit with the PreLI0 assumption, in which the response centers in the multi-source case were located at the same locations they had in the single-source conditions. The differences in modeling results between NH and CI listeners suggest that CI listeners may have reduced integration ability.

(5) The lack of bimodal benefits for phoneme identification in CI listeners can be attributed to a combination of factors, including the insufficient or redundant information provided by an individual ear, reduced ability to integrate speech cues across ears, or both.

Acknowledgments

Appendix

Preliminary Study: Model Predictions for non-overlapping frequency speech cues in NH listeners

Methods

A total of 12 NH subjects (10 females and 2 males), aged 19 to 46 years (mean 24 years) participated in the study. Nine of them participated in the consonant identification experiment. Seven subjects participated in the vowel identification task. Four of these seven subjects also participated in the consonant identification task.

Two sets of speech stimuli (consonants and vowels) identical to those in Experiment 1 were used. These stimuli were subjected to LP filtering or channel-vocoding processing. The only difference in processing parameters between this preliminary study and Experiment 1 was that the channel-vocoding processing preserved only high-frequency cues (>900 Hz), thus there was no (or very minimal) overlapping in frequencies between the LP and vocoder speech. In this vocoder system, speech signal in a frequency range from 900 to 6000 Hz was band-pass filtered into two, four, or six logarithmic frequency bands (corresponding to the 2ch, 4ch, and 6ch vocoders).

Experimental procedures were the same as those employed in Experiment 1. Since there were three vocoder conditions (6ch vocoder, 4ch vocoder, and 2ch vocoder), there were a total of seven testing conditions (one LP-alone, three vocoder-alone; and three LP + vocoder). Each subject was tested with 6ch vocoder first, followed by 4ch and then 2ch. The order of presentation of the listening conditions (LP-alone, vocoder-alone, LP+vocoder) was counterbalanced across subjects.

Model predictions

Integration ability across frequencies was evaluated using the PreLI0, PreLI1, and PreLIH integration models that differ in the location of the response centers in the multi-source condition described in Ronan et al. (2004). Predictions were made separately for each subject, condition, and task. They were computed by first fitting the vocoder-alone and LP-alone matrices in $D = 3$ dimensions and then predicting the scores for the combined condition from a 6-dimensional model. Figure 6 shows the predicted versus observed combined-source consonant (upper panel) and vowel identification (lower panel) scores for each response center location (triangles: PreLI0; squares: PreLI1; circles: PreLIH). On average, PreLI0 underpredicted the combined performance by about 5 percentage points for consonant identification and 9 percentage points for vowel identification. PreLIH overpredicted the combined performance by an average of 4 percentage points for consonant identification and 3 percentage points for vowel identification. Compared to PreLIH, PreLI1 overpredicted the combined performance to a greater extent with an average of 7 percentage points for consonant identification and 4 percentage points for vowel identification. The root-mean-squared-error (RMSE) of the fit was greater for PreLI0 (consonant: 6; vowel: 10) than for PreLIH (consonant: 5; vowel: 4). In general, this pattern of results is similar to those reported in Ronan et al. (2004) for cross-frequency integration for consonant identification in NH listeners, which also showed that PreL0 consistently underestimated the multiband scores and that PreLIH provided better predictions for cross-frequency consonant identification compared to PreLI1.

The consistent pattern of results from the model fits between Ronan et al. (2004) and this preliminary study suggests that the PreL model of integration is capable of predicting

combined-band performance when speech signals from different frequency regions undergo different signal processing (LP filtering and channel vocoding).

REFERENCES

Armstrong M, Pegg P, James C, Blamey P. Speech perception in noise with implant and hearing aid. American Journal of Otology. 1997; 18:S140–141. [PubMed: 9391635]

Boothroyd A. Auditory perception of speech contrasts by subjects with sensorineural hearing loss. Journal of Speech and Hearing Research. 1984; 27:134–144. [PubMed: 6716999]

Braida LD. Crossmodal integration in the identification of consonant segments. Quarterly Journal of Experimental Psychology. 1991; 43:647–677. [PubMed: 1775661]

Braida LD. Integration models of intelligibility. Proceedings of N.A.S.-C.H.A.B.A. Symposium on Speech Communication Metrics and Human Performance, Washington, DC. 1993:1–20.

Brown CA, Bacon SP. Low-frequency speech cues and simulated electric-acoustic hearing. Journal of the Acoustical Society of America. 2009a; 125:1658–1665. [PubMed: 19275323]

Brown CA, Bacon SP. Achieving electric-acoustic benefit with a modulated tone. Ear and Hearing. 2009b; 30:489–493. [PubMed: 19546806]

Ching TY, Incerti P, Hill M. Binaural benefits for adults who use hearing aids and cochlear implants in opposite ears. Ear and Hearing. 2004; 25:9–21. [PubMed: 14770014]

Ching TY, Psarros C, Hill M, Dillon H, Incerti P. Should children who use cochlear implants wear hearing aids in the opposite ear? Ear and Hearing. 2001; 22:365–380. [PubMed: 11605945]

Ching TYC, van Wanrooy E, Dillon H. Binaural-bimodal fitting or bilateral implantation for managing severe to profound deafness: a review. Trends in Amplification. 2007; 11:161–192. [PubMed: 17709573]

Chomsky, N.; Halle, M. The sound pattern of English. Harper & Row; New York: 1968.

Dooley GJ, Blamey PJ, Seligman PM, Alcantara JI, Clark GM, Shallop J, Arndt P, Heller JW, Menapace CM. Combined electrical and acoustical stimulation using a bimodal prosthesis. Archives of Otolaryngology – Head & Neck Surgery. 1993; 119:55–60. [PubMed: 8417744]

Dorman MF, Gifford R, Spahr A, McKarns SA. The benefits of combining acoustic and electric stimulation for the recognition of speech, voice, and melodies. Audiology and Neurotology. 2008; 13:105–112. [PubMed: 18057874]

Dunn CD, Tyler RS, Witt SA. Benefit of wearing a hearing aid on unimplanted ear in adult users of a cochlear implant. Journal of Speech, Language, and Hearing Research. 2005; 48:668–680.

Fishman KE, Shannon RV, Slattery WH. Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. Journal of Speech, Language, and Hearing Research. 1997; 40:1201–1215.

Gifford RH, Dorman MF, McKarns SA, Spahr AJ. Combined electric and contralateral acoustic hearing: Word and sentence recognition with bimodal hearing. Journal of Speech, Language, and Hearing Research. 2007a; 50:835–843.

Gifford RH, Dorman MF, Spahr AJ, Bacon SP. Auditory function and speech understanding in listeners who qualify for EAS surgery. Ear and Hearing. 2007b; 28:114S–118S. [PubMed: 17496661]

Grant KW. Measures of auditory-visual integration for speech understanding: A theoretical perspective (L). Journal of the Acoustical Society of America. 2002; 112:30–33. [PubMed: 12141356]

Grant KW, Seitz PF. Measures of auditory-visual integration in nonsense syllables and sentences. - Journal of the Acoustical Society of America. 1998; 104:2438–2450. [PubMed: 10491705]

Grant KW, Tufts JB, Greenberg S. Integration efficiency for speech perception within and across sensory modalities by normal-hearing and hearing impaired individuals. Journal of the Acoustical Society of America. 2007; 121:1164–1176. [PubMed: 17348537]

Grant KW, Walden BE, Seitz PF. Auditory-visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory-visual integration. Journal of the Acoustical Society of America. 1998; 103:2677–2690. [PubMed: 9604361]

Hamzavi J, Pok SM, Gstoettner W, Baumgartner WD. Speech perception with a cochlear implant used in conjunction with a hearing aid in the opposite ear. International Journal of Audiology. 2004; 43:61–65. [PubMed: 15035557]

Healy EW, Bacon SP. Across-frequency comparison of temporal speech information by listeners with normal and impaired hearing. Journal of Speech, Language, and Hearing Research. 2002; 45:1262–1275.

Hillenbrand J, Getty LA, Clark MJ, Wheeler K. Acoustic characteristics of American English vowels. Journal of the Acoustical Society of America. 1995; 97:3099–3111. [PubMed: 7759650]

Kong Y-Y, Carlyon RP. Improved speech recognition in noise in simulated binaurally combined acoustic and electric stimulation. Journal of the Acoustical Society of America. 2007; 121:3717–3727. [PubMed: 17552722]

Kong Y-Y, Stickney GS, Zeng F-G. Speech and melody recognition in binaurally combined acoustic and electric hearing. Journal of the Acoustical Society of America. 2005; 117:1351–1361. [PubMed: 15807023]

Li N, Loizou PC. A glimpsing account for the benefit of simulated combined acoustic and electric hearing. Journal of the Acoustical Society of America. 2008; 123:2287–2294. [PubMed: 18397033]

Massaro, DW. Speech Perception by Ear and Eye: A Paradigm for Psychological Inquiry. Lawrence Erlbaum Hillsdale; NJ: 1987.

Massaro, DW. Perceiving Talking Faces: From Speech Perception to a Behavioral Principle. MIT Press; Cambrige, MA: 1998.

Massaro DW, Cohen MM. Test of auditory-visual integration efficiency within the framework of the fuzzy logical model of perception. Journal of the Acoustical Society of America. 2000; 108:784–789. [PubMed: 10955645]

McDermott HJ, Dorkos VP, Dean MR, Ching TY. Improvements in speech perception with use of the AVR TranSonic frequency-transposing hearing aid. Journal of Speech, Language, and Hearing Research. 1999; 42:1323–1335.

Miller GA, Nicely PE. An analysis of perceptual confusions among some English consonant. Journal of the Acoustical Society of America. 1955; 27:338–352.

Mok M, Galvin KL, Dowell RC, McKay CM. Speech perception benefit for children with cochlear implant and a hearing aid in opposite ears and children with bilateral cochlear implants. Audiology and Neurotology. 2009; 15:44–56. [PubMed: 19468210]

Mok M, Grayden D, Dowell RC, Lawrence D. Speech perception for adults who use hearing aids in conjunction with cochlear implants in opposite ears. Journal of Speech, Language, and Hearing Research. 2006; 49:338–351.

Palva AK, Jokinen K. Role of the binaural test in filtered speech audiometry. Acta Oto-Laryngologica. 1975; 79:310–313. [PubMed: 1136771]

Reiss LAJ, Gantz BJ, Turner CW. Cochlear implant speech processor frequency allocations may influence pitch perception. Otology and Neurotology. 2008; 29:160–167. [PubMed: 18025998]

Ronan D, Dix AK, Shah P, Braida LD. Integration across frequency bands for consonant identification. Journal of the Acoustical Society of America. 2004; 116:1749–1762. [PubMed: 15478442]

Ross LA, Saint-Amour D, Leavitt VM, Javitt DC, Foxe JJ. Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. Cerebral Cortex. 2006; 17:1147–1153. [PubMed: 16785256]

Sagi E, Svirsky MA. Information transfer analysis: a first look at estimation bias. Journal of the Acoustical Society of America. 2008; 123:2848–2857. [PubMed: 18529200]

Sagi E, Meyer TA, Kaiser AR, Teoh SW, Svirsky MA. A mathematical model of vowel identification by users of cochlear implants. Journal of the Acoustical Society of America. 2010; 127:1069–1083. [PubMed: 20136228]

Shallop JK, Arndt P, Turnacliff KA. Expanded indications for cochlear implantation: perceptual results in seven adults with residual hearing. Journal of Speech-Language Pathology and Audiology. 1992; 16:141–148.

Shannon RV, Jensvold A, Padilla M, Robert ME, Wang X. Consonant recordings for speech testing. Journal of the Acoustical Society of America. 1999; 106:L71–74. [PubMed: 10615713]

Souza PE, Boike KT. Combining temporal-envelope cues across channels: effects of age and hearing loss. Journal of Speech, Language, and Hearing Research. 2006; 49:138–149.

Spitzer S, Liss J, Spahr T, Dorman M, Lansford K. The use of fundamental frequency for lexical segmentation in listeners with cochlear implants. Journal of the Acoustical Society of America. 2009; 125:EL236–241. [PubMed: 19507928]

Turner CW, Henry BA. Benefits of amplification for speech recognition in background noise. Journal of the Acoustical Society of America. 2002; 112:1675–1680. [PubMed: 12398472]

Turner CW, Chi S-L, Flock S. Limiting spectral resolution in speech for listeners with sensorineural hearing loss. Journal of Speech, Language, and Hearing Research. 1999; 42:773–784.

Turner CW, Gantz BJ, Vidal C, Behrens A, Henry BA. Speech recognition in noise for cochlear implant listeners: benefits of residual acoustic hearing. Journal of the Acoustical Society of America. 2004; 115:1729–1735. [PubMed: 15101651]

Tyler RS, Parkinson AJ, Wilson BS, Witt S, Preece JP, Noble W. Patients utilizing a hearing aid and a cochlear implant: speech perception and localization. Ear and Hearing. 2002; 23:98–105. [PubMed: 11951854]

Zhang T, Dorman MF, Spahr AJ. Information from the voice fundamental frequency (F0) region accounts for the majority of the benefit when acoustic stimulation is added to electric stimulation. Ear and Hearing. 2010; 31:63–69. [PubMed: 20050394]

Figure 1.
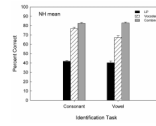Average overall consonant and vowel identification scores in NH listeners for the three listening conditions (black bars: LP speech alone; slated bars: vocoder speech alone; gray bars: LP+vocoder). Error bars represent standard error of the mean.
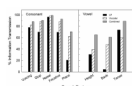
Figure 2.
Average percent information transmission for five consonant features (left) and three vowel features (right) in three listening conditions for NH listeners.
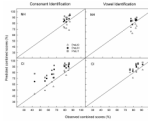
Figure 3.
PreL model predictions for consonant (left) and vowel (right) identification for NH (top) and CI (bottom) listeners. Predictions made by the PreLI0, PreLI1, and PreLIH methods are represented by open triangles, gray squares, and dark circles, respectively. Note that some data points overlap in the vowel identification results for NH listeners (upper right panel).
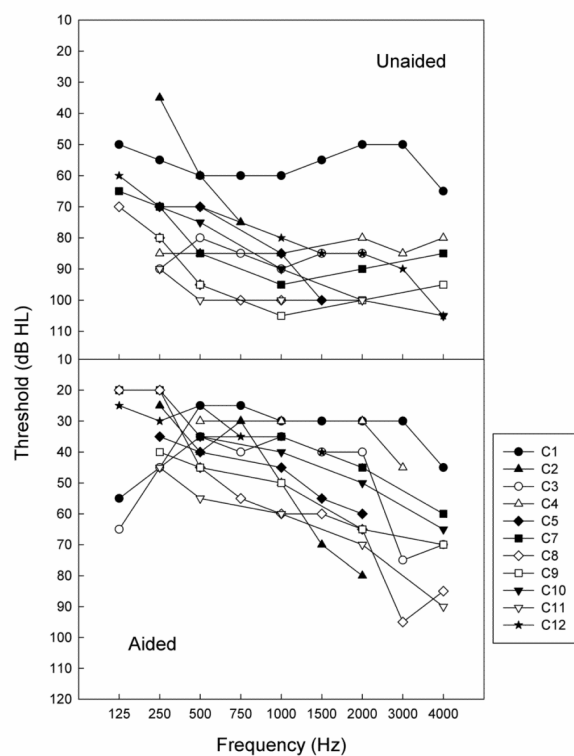
Figure 4.
Unaided (top) and aided (bottom) thresholds in the non-implanted ear for each individual CI subject.
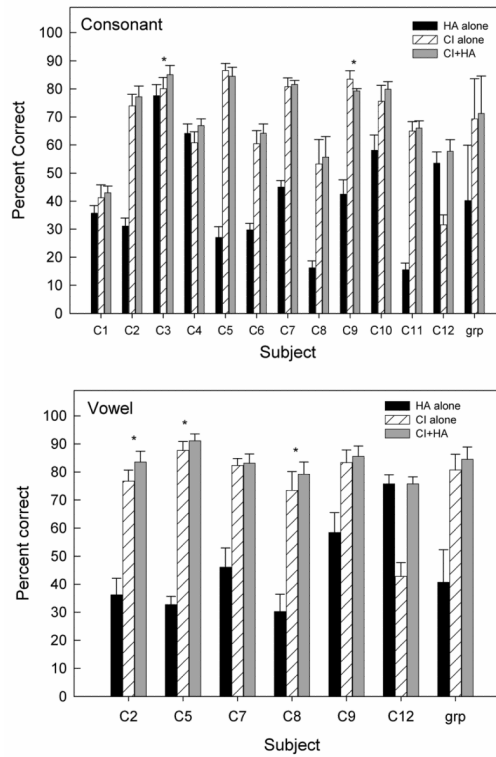
Figure 5.
Overall consonant (top) and vowel (bottom) identification scores for individual CI listeners for three listening conditions. Error bars represent standard error of the mean. Significant difference between the CI+HA score and the better ear score is marked with an asterisk (*).
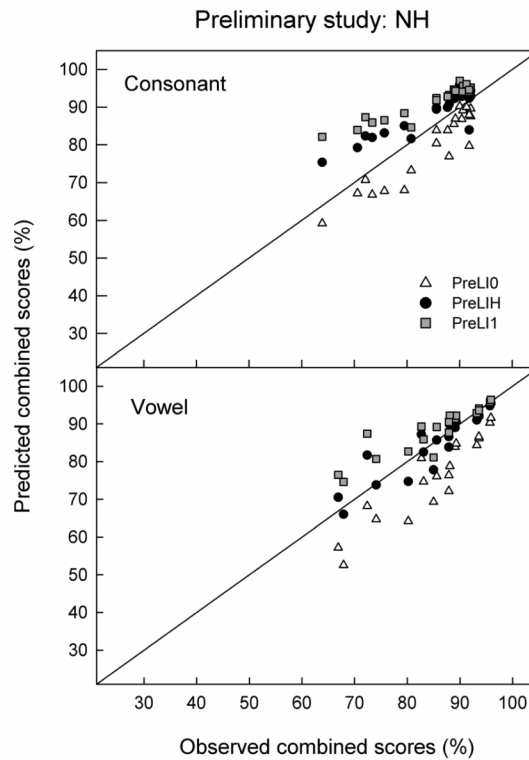
Figure 6.
PreL model predictions for NH listeners in the preliminary study with overlapping frequency speech cues for consonant (top) and vowel (bottom) identification. Predictions made by the PreLI0, PreLI1, and PreLIH models are represented by open triangles, gray squares, and dark circles, respectively.

Table I

Lower, center, and upper frequencies of filters used to create simulations of processor with 4 channels.

	Analysis & carrier band			
	1	2	3	4
Lower	200	575	1336	2877
Center	355	889	1971	4165
Upper	575	1336	2877	6000

Table II

Feature classification for voicing, manner of articulation, and place of articulation categories for consonants, and height, back, and tense categories for vowels.

Consonants	Vowels
VOICING	HEIGHT
Voiced: /b, d, g, v, ð, z, ʒ, m, n/	High: /i, ɪ, u, ʊ/
Unvoiced: /p, t, k, f, θ, s, ʃ /	Mid: /ɛ, ɝ, ʌ, ɔ/ Low: /æ/
MANNER	BACK
Stop: /p, t, k, b, d, g/	Front: /i, ɪ, ɛ, æ/
Nasal: /m, n/	Central: /ɝ, ʌ/
Fricative: /f, θ, s, ʃ, v, ð, z, ʃ/	Back: /u, ʊ, ɔ/
PLACE	TENSE
Bilabial: /p, b, m/	Tense: /i, u/
Alveolar: /t, d, s, z, n/	Lax: /i, ɛ, æ, ɝ, ʌ, ʊ, ɔ/
Labio-Dental: /f, v/	
Dental: /θ, ð/	
Palatal: /ʃ, ʒ/	
Velar: /k, g/	

Table III

CI subjects detailed demographic information

Subject	Age	Gender	CI Ear	Onset HL* (CI ear)	Etiology (CI ear)	Dur (yrs) HL prior CI**	CI Processor	Yrs of CI use
C1	46	M	L	mid-20s	Unknown	5	Harmony	1.5
C2	64	F	L	37	Unknown	10	Freedom	2
C3	21	M	R	3	Morquio Syndrome	6	Harmony	1
C4	16	F	R	birth	Enlarged aqueduct	14	Freedom	1
C5	19	F	L	birth	Unknown	9	ESPrit 3G	10
C6	69	F	L	45	Unknown	5	Harmony	2.5
C7	26	M	R	birth	Genetic	22	Freedom	4.5
C8	57	M	L	5	Premature birth	6	Harmony	5
C9	16	F	R	birth	Mondini	3	Freedom	1
C10	15	F	L	birth	Cogan Syndrome	10	Freedom	5
C11	16	M	R	birth	Genetic	11	Freedom	5
C12	64	F	R	2	Meningitis	60	Harmony	5.5

*
Age of onset of hearing loss in the CI ear.

**
Duration of severe to profound hearing loss >1000 Hz in the CI ear prior to implantation.

Table IV

Percent information transmission for consonant identification

Subject	Condition	Voicing	Stop	Nasal	Fricative	Place
C1	HA-alone	34.75	24.47	47.85	8.55	27.78
	CI-alone	42.37	30.67	64.73	20.70	28.1
	CI+HA	45.68	30.29	70.66	24.96	27.8
C2	HA-alone	38.29	17.93	90.93	19.68	12.29
	CI-alone	55.96	78.08	90.25	74.17	67.09
	CI+HA	56.72	78.94	95.76	77.47	72.35
C3	HA-alone	79.15	71.80	98.25	71.64	58.33
	CI-alone	84.08	79.09	90.25	76.31	67.17
	CI+HA	88.66	83.45	100.00	83.38	74.60
C4	HA-alone	76.93	65.65	94.70	63.52	46.75
	CI-alone	65.22	69.65	73.39	61.23	40.19
	CI+HA	76.54	74.16	93.71	72.23	47.61
C5	HA-alone	31.35	19.78	66.69	20.76	8.31
	CI-alone	70.98	83.02	100.00	83.25	82.16
	CI+HA	63.67	82.55	100.00	82.17	81.88
C6	HA-alone	32.74	22.57	71.65	20.85	10.55
	CI-alone	45.93	63.45	72.89	63.23	47.84
	CI+HA	50.89	59.43	93.57	61.62	49.86
C7	HA-alone	57.75	37.22	94.59	36.68	22.07
	CI-alone	77.05	78.31	97.66	77.77	72.10
	CI+HA	81.64	79.28	100.00	78.99	72.06
C8	HA-alone	14.30	12.55	42.20	8.77	2.78
	CI-alone	18.92	52.02	68.84	48.71	40.90
	CI+HA	26.93	66.23	82.82	61.94	42.12
C9	HA-alone	39.69	17.01	72.46	17.29	25.17
	CI-alone	65.30	91.53	100.00	91.40	78.73
	CI+HA	54.75	90.64	100.00	90.93	76.89
C10	HA-alone	78.55	63.29	96.93	63.37	37.18
	CI-alone	81.01	82.45	94.59	80.40	62.39

Subject	Condition	Voicing	Stop	Nasal	Fricative	Place
	CI+HA	71.64	77.81	95.91	78.11	68.07
C11	HA-alone	13.97	2.22	76.09	4.68	2.15
	CI-alone	55.64	47.74	84.75	47.03	46.88
	CI+HA	53.33	43.9	90.46	42.38	47.00
C12	HA-alone	69.33	56.94	97.66	59.62	38.58
	CI-alone	27.60	38.55	9.47	24.99	25.51
	CI+HA	69.19	61.31	94.59	64.02	48.71
Group (exclude C12)	HA-alone	39.47	26.78	72.38	25.48	16.27
	CI-alone	56.26	66.44	82.80	63.20	52.32
	CI+HA	58.06	67.53	91.68	66.18	55.21

Table V

Percent information transmission for vowel identification

Subject	Condition	Height	Back	Tense
C2	HA-alone	22.59	1.44	59.79
	CI-alone	44.06	73.79	66.93
	CI+HA	53.45	78.05	84.36
C5	HA-alone	26.09	5.23	70.67
	CI-alone	70.62	84.87	93.43
	CI+HA	74.97	89.36	100.00
C7	HA-alone	44.20	7.24	75.35
	CI-alone	60.68	76.63	79.03
	CI+HA	57.19	77.73	89.64
C8	HA-alone	17.63	1.47	36.47
	CI-alone	48.47	49.61	69.14
	CI+HA	59.07	59.75	79.10
C9	HA-alone	41.19	17.89	73.44
	CI-alone	62.45	75.34	86.07
	CI+HA	65.05	75.84	92.95
C12	HA-alone	64.83	43.47	91.96
	CI-alone	13.23	14.93	43.74
	CI+HA	62.60	46.76	89.73
Group (exclude C12)	HA-alone	27.27	4.35	59.50
	CI-alone	53.32	69.23	77.80
	CI+HA	60.42	73.37	87.88