Do adults with cochlear implants rely on different acoustic cues for phoneme perception than adults with normal hearing?

Aaron C. Moberly, **Joanna H. Lowenstein**, **Eric Tarr**, **Amanda Caldwell-Tarr**, **D. Bradley Welling**, **Antoine J. Shahin**, and **Susan Nittrouer**
The Ohio State University, Wexner Medical Center

Abstract

Purpose—Several acoustic cues specify any single phonemic contrast. Nonetheless, adult, native speakers of a language share weighting strategies, showing preferential attention to some properties over others. Cochlear implant (CI) signal processing disrupts the salience of some cues: in general, amplitude structure remains readily available, but spectral structure less so. This study asked how well speech recognition is supported if CI users shift attention to salient cues not weighted strongly by native speakers.

Method—20 adults with CIs participated. The /bɑ/-/wɑ/ contrast was used because spectral and amplitude structure varies in correlated fashion for this contrast. Normal-hearing adults weight the spectral cue strongly, but the amplitude cue negligibly. Three measurements were made: labeling decisions, spectral and amplitude discrimination, and word recognition.

Results—Outcomes varied across listeners: some weighted the spectral cue strongly, some weighted the amplitude cue, and some weighted neither. Spectral discrimination predicted spectral weighting. Spectral weighting explained the most variance in word recognition. Age of onset of hearing loss predicted spectral weighting, but not unique variance in word recognition.

Conclusions—The weighting strategies of listeners with normal hearing likely support speech recognition best, so efforts in implant design, fitting, and training should focus on developing those strategies.

The cochlear implant (CI) is a widely accepted and effective treatment for rehabilitation of patients with severe-to-profound hearing loss. Speech recognition outcomes have improved since the introduction of the single-channel device (Ambrosch et al., 2010; David et al., 2003), but, on average, adult CI users still score only 40–50 percent correct on word recognition tests in quiet (Firszt et al., 2004; Hamzavi, Baumgartner, Pok, Franz, & Gstoettner, 2003), a common metric of CI performance. Furthermore, substantial variability in speech recognition remains among CI users (Keifer, von Ilberg, & Reimer, 1998; Peterson, Pisoni, & Miyamoto, 2010; Shipp & Nedzelski, 1995; Zeng, 2004). Improving speech recognition is a central focus of CI research.

Corresponding author: Aaron C. Moberly, Aaron.Moberly@osumc.edu, Office phone: 614-366-3687, Cell phone: 317-345-1662, FAX: 614-293-7292.

Without question, the speech perception deficits exhibited by CI users arise primarily due to signal degradation. Although speech signals consist of structure in the spectral and amplitude domains, CIs likely do not provide equivalent representations of spectral and amplitude changes. Current implant speech processors operate by recovering amplitude structure in up to 22 independent frequency channels, but the effective number of available channels is typically limited to four to seven (Friesen, Shannon, Baskent, & Wang, 2001). Because instantaneous amplitude is well preserved in each channel, the gross temporal envelope of the entire signal should be well preserved. Spectral structure within each of those channels, however, is significantly degraded (Wilson & Dorman, 2008), which means that trajectories of individual formants are only poorly represented. To illustrate this problem, Figure 1 shows spectrograms of a natural token of the word "kite" spoken by a man (left side) and of that same token after being vocoded using eight channels (right side). Although not a perfect model of CI speech processing, it shows that amplitude structure, represented by the gross temporal envelope of the waveform shown at the top, is fairly well preserved, but formant transitions are significantly degraded. Because of this processing limitation, it seems likely that CI users could have difficulty with phonetic decisions that typically rely on formant transitions in some way. At the same time, any decisions that can be made based on amplitude structure should be well supported for these individuals.

Perceptual weighting strategies of first- and second-language learners

Although the term had been used for decades, in 1982 Repp defined acoustic cues as isolable properties of the speech signal that when separately manipulated influence phonemic decisions. Several acoustic cues might be associated with a single phonemic decision, in a correlated manner. Nonetheless, the amount of attention, or weight, given to those various cues usually differs. Individual speakers of a given language show similar weighting strategies for phonemic decisions – weighting some cues strongly and others only weakly – most likely because those strategies allow for the most accurate and efficient speech perception in that language (Best, 1994; Jusczyk, Hohne, & Mandel, 1995; Nittrouer, 2005). Looking beyond the population of adult, native speakers, however, perceptual weighting strategies are found to vary across listeners depending on factors such as age and language experience. And those weighting strategies can change for individual listeners.

The notion of a developmental weighting shift grew out of work showing that children weight the cues to phonemic categories differently than adults (Nittrouer, 1992; Nittrouer & Studdert-Kennedy, 1987). This model suggests that children modify the perceptual weight they assign to acoustic cues as they discover which properties are phonetically informative in their native language (Nittrouer, Manning, & Meyer, 1993). Initially, children do not have highly refined phonetic representations (Liberman, Shankweiler, Fischer, & Carter, 1974). As those representations become more refined, children must learn what structure within the speech signal best supports the recovery of those phonemic units, and they come to weight that structure strongly (Greenlee & Ohala, 1980; Hicks & Ohde, 2005; Mayo, Scobbie, Hewlitt, & Waters, 2003; Nittrouer, 2002; Nittrouer & Miller, 1997; Wardrip-Fruin & Peach, 1984).

Adult, second-language learners possess efficient weighting strategies for their native language, but have difficulty perceiving non-native phonetic contrasts when they begin learning a second language (Beddor & Strange, 1982; Best, McRoberts, & Goodell, 2001; Gottfried, 1984). These perceptual difficulties do not appear to stem from a loss of auditory sensitivity to the acoustic cues that underlie those phonetic decisions (Miyawaki et al., 1975; Werker & Logan, 1985). Instead, language-specific perceptual patterns reflect selective attention to the acoustic cues that are most phonetically informative in the first language (Strange, 1986; 1992). Second-language learners may differentiate non-native contrasts using weighting strategies other than those used by native listeners of that language and more like listeners of their own language (Underbakke, Polka, Gottfried, & Strange, 1988), a strategy that is often not highly effective. Over time and with increasing experience in the second language, individuals may shift their weighting strategies for more efficient perception of the second language (Flege, 1995; Miyawaki et al., 1975; Strange, 1992). Thus, both developmental studies and experiments with second-language learners demonstrate that listeners can modify their weighting strategies to achieve more efficient speech perception. The question addressed in the current study was whether people who get CIs modify their perceptual weighting strategies based on the kinds of structure most readily available through their implants, and what the effects on speech recognition are of either doing so or not doing so.

Adult, post-lingually deafened individuals with CIs presumably have developed highly refined perceptual weighting strategies for their primary language prior to losing their hearing. At the time of implant activation, they must begin to make sense of the degraded electrical representations of speech signals. Like children learning a first language or adults learning a second language, it is conceivable that these individuals with new CIs may shift their perceptual weighting strategies. In this case, however, the availability of cues from their implants may constrain the choice of cues they can select to weight. The cues that are available are not necessarily the most phonetically informative. Since the gross temporal envelope of speech should be well represented by their implants, they may weight amplitude structure heavily. Conversely, they may selectively extract the degraded spectral structure, and continue to weight it heavily in phonetic decisions where it is typically used.

Exploring cue weighting with individuals with CIs

Several authors have examined the topic of cue weighting by adults who lost their hearing after learning language and used CIs. Both spectral and amplitude cues – as well as duration cues – appear to be important in phonemic judgments for these listeners (Dorman, Dankowski, McCandless, Parkin, & Smith, 1991; Nie, Barco, & Zeng, 2006; Xu, Thompson, & Pfingst, 2005), but the relative importance of these cues is unclear. Some authors have found that CI users perceive contrasts using cue-weighting strategies similar to those of normal-hearing individuals, and have suggested that this similarity in cue weighting may be related to accuracy in speech recognition. In a small group of CI users with relatively good word recognition abilities, for example, Dorman and colleagues (1991) observed that most had labeling functions for a voice-onset-time continuum that were similar to individuals with normal hearing. Iverson (2003) found the best speech recognition in adults with CIs who had peak sensitivity for a voice-onset-time continuum at the typical

location, even if overall level of sensitivity was poorer than that of listeners with normal hearing. Iverson and colleagues (2006) examined cue-weighting strategies in vowel recognition by post-lingually deafened CI users as well as normal-hearing individuals listening to noise-vocoded simulations of CI processing. Implant users and normal-hearing subjects with CI simulations used formant movement and duration cues to the same extent in vowel recognition. Looking at pre-lingually deafened children who use CIs, Giezen and colleagues (2010) found that these children used similar cue-weighting strategies as children with normal hearing for three of four contrasts, /ɑ/-/a/, /ɪ/-/i/, and /bu/-/pu/. Thus, in some instances listeners using CIs rely on the same cues as listeners with normal hearing for phonemic judgments.

On the other hand, there is some evidence supporting the idea that individuals with CIs use different perceptual weighting strategies than listeners with normal hearing. For example, in the study by Giezen and colleagues (2010), the children with CIs used spectral cues in the /fu/-/su/ contrast less effectively than normal-hearing children. Hedrick and Carney (1997) conducted a study examining the relative contributions of formant transitions and amplitude structure to labeling decisions for synthetic fricative-vowel syllables. Four post-lingually deafened adults with CIs were found to weight the amplitude cue significantly more strongly than listeners with normal hearing. Therefore, the relationship between the weighting strategies used by individuals with CIs and their speech recognition abilities remains unclear. The purpose of this current study was to further investigate this relationship.

Evaluating cue weighting in adults with CIs using the /bɑ/-/wɑ/ contrast

To examine the perceptual weighting of spectral and amplitude structure for individuals with CIs in the current study, a contrast was chosen in which cues regarding both types of structure are salient for listeners with normal hearing, but the weighting of those cues differs drastically. Fitting this description were synthetic /bɑ/ and /wɑ/ syllables. These two syllables form a minimal pair, with the only difference being manner of production: /b/ is a stop and /w/ is a glide. The articulatory gestures involved in producing both syllables are essentially the same: the vocal tract is closed at the lips and then opens. The primary difference in production is the rate of mouth opening: it is rapid for /bɑ/ and slow for /wɑ/. Consequently, onset and steady-state formant frequencies are similar, but the time it takes to reach steady-state values differs for /bɑ/ and /wɑ/. That interval is termed the formant rise time (FRT). Similarly, the time required to reach peak amplitude differs for the two syllables, and that interval is termed the amplitude rise time (ART). Both FRT and ART are shorter for /bɑ/ than for /wɑ/.

It would seem that both FRT and ART could serve as cues to the /bɑ/-/wɑ/ contrast because both cues are equally salient for listeners with normal hearing. For adult, native English speakers, however, FRT has been found to be the primary cue, with ART receiving little, if any, weight. For example, Nittrouer and Studdert-Kennedy (1986) switched ARTs across syllable types using natural speech tokens and found that adults continued to label consonants based on FRT. Both Walsh and Diehl (1991) and Nittrouer, Lowenstein, and Tarr (2013) replicated those results using synthetic stimuli, which provides an important control. In natural speech, fundamental frequency differs between the two consonantal

contexts and could serve as another cue to consonant identity. Using synthetic stimuli eliminates that possibility because fundamental frequency can be held constant across stimuli, so that only ART and FRT are permitted to vary in a controlled manner. Vowel duration, which could be impacted by speaking rate, also plays a minor role in perception of the /bɑ/-/wɑ/ contrast (Miller & Liberman, 1979; Miller & Wayland, 1993; Shinn, Blumstein, & Jongman, 1985) by shifting the phoneme boundary for formant transitions by a few milliseconds, but the effect is only found at syllable lengths of less than 200 ms. In order to eliminate vowel duration as a conditioning context in this study, duration was kept constant for all stimuli at a length greater than 300 ms.

The /bɑ/-/wɑ/ contrast provides an excellent opportunity to investigate weighting strategies in CI users because these devices likely fail to provide a veridical representation of spectral structure, but represent amplitude structure well. Several related questions can therefore be asked with these stimuli: Do listeners with CIs adjust their weighting strategies after implantation according to the kind of structure that is best represented by their implant processing? And is their speech recognition affected by any changes in perceptual weighting strategies that are observed?

Two contrasting hypotheses were developed concerning what perceptual weighting strategies might be for CI users with these /bɑ/-/wɑ/ stimuli. The first hypothesis was the ***different, but effective, strategies*** hypothesis. According to this hypothesis, the best performing CI users would shift their perceptual strategies to weight the most salient properties delivered by their implants. In the case of the /bɑ/-/wɑ/ contrast, this would suggest that amplitude structure (i.e., ART) would be weighted more strongly than what has been observed for adults with normal hearing.

The second hypothesis developed was the ***same, but adapted, strategies*** hypothesis. If this hypothesis is correct, the best performing CI users do not shift their perceptual strategies to the most salient property. Instead, they continue to weight spectral structure (i.e., FRT) heavily in the /bɑ/-/wɑ/ contrast, as they did prior to losing their hearing and getting an implant, even though this information is highly degraded by their implants. Although the representation of spectral structure may not be as refined as it is for normal-hearing listeners, it nonetheless remains the primary cue.

In the current experiment, the relative weighting of spectral and amplitude structure was examined using just this one phonemic contrast. Although perhaps slightly over-reaching, the weighting strategies found for this contrast were taken to characterize more general strategies used by these listeners. Accordingly, weighting factors for these stimuli were correlated with word recognition scores to assess whether listeners with CIs fare better when they manage to retain the same weighting strategies as those used by listeners with normal hearing, or when they adjust weighting strategies to accommodate the kinds of signal structure most accessible through their implants. In order for the ***same, but adapted, strategies*** hypothesis to receive support, strong and positive correlations should be found between strategies that weight FRT strongly and word recognition scores. In order for the ***different, but effective, strategies*** hypothesis to be supported, perceptual strategies that

strongly weight ART should be positively and strongly correlated with word recognition scores.

Indexing the weights assigned to acoustic cues

In the current study, the same stimuli as those used by Nittrouer et al. (2013) were presented to adults with CIs, and perceptual weighting strategies were indexed using calculated weighting factors. For illustrative purposes, Figure 2 shows labeling functions for adults with normal hearing in that study, obtained with two complementary sets of stimuli. On the top panel, labeling functions are shown for stimuli in which FRT was manipulated along a nine-step continuum. In this condition, FRT was the *continuous* cue. Each stimulus along that continuum was fit with each of two ARTs: a short ART, as is found for /bɑ/, and a long ART, as is found for /wɑ/. Consequently, ART is termed the *binary* cue. Probit (similar to logit) functions were subsequently fit to recognition probabilities obtained for stimuli along each continuum. An estimate of the relative weighting of FRT (the continuous cue) and ART (the binary cue) could be gathered, respectively, from the slopes of the labeling functions and the separation between those functions based on ART. The steeper the functions, the more weight that was assigned to FRT; the greater the separation between functions for stimuli with /bɑ/ and /wɑ/ ARTs, the more weight that was assigned to ART. Thus, steep functions that are close together suggest that listeners largely made their phonemic decisions based on FRT. Shallower functions that are well separated (based on ART) would suggest that the listeners based their decisions on whether gross temporal envelopes were appropriate for a syllable-initial /b/ or /w/. As can be seen in the top panel of Figure 2, adults strongly weighted FRT and barely weighted ART.

In a complementary but separate condition in that same study, ART was manipulated along a continuum, and two different FRTs were applied in a binary manner: one appropriate for /bɑ/ and one appropriate for /wɑ/. The bottom panel of Figure 2 shows labeling functions for this condition. Here, labeling functions are shallow and widely separated. This labeling pattern again reflects strong weighting of FRT and almost no weighting of ART.

Although the above technique - using steepness of and separation between labeling functions - is useful for illustrating data, it is not possible to compare weighting for each cue in an "apples-to-apples" fashion with this approach. More equivalent metrics are obtained by using recognition probabilities in logistic regression, and this method has been applied in a number of speech perception studies (Benkí, 2001; Giezen, Escudero, & Baker, 2010; McMurray & Jongman, 2011; Morrison, 2009; Smits, Sereno, & Jongman, 2006). The calculated regression coefficients can be used as metrics of the extent to which phonemic labels are explained by each cue. Hereafter, these regression coefficients are referred to as "weighting factors." Logistic regression was used to obtain weighting factors for FRT and ART in each condition separately in the data from Nittrouer et al. (2013), and mean FRT and ART weighting factors were computed across conditions. Reflecting the observations from the labeling functions in Figure 2, these weighting factors for the adults with normal hearing were 9.52 for FRT and 1.10 for ART.

Discrimination abilities of adults with CIs

The focus of the study reported here was on the perceptual weighting of spectral and amplitude cues to a specific phonemic contrast, and the potential effect of that weighting on speech recognition. However, this approach makes the assumption that the weighting of acoustic cues is at least quasi-independent of a listener's sensitivity to those cues. Certainly that is a fair assumption for children learning a first language (e.g. Nittrouer, 1996) and adults learning a second language (Miyawaki et al., 1975). It is not as obvious for listeners with hearing loss because there is good reason to suspect that their sensitivity is compromised. Previous studies have found a relationship between the spectral and temporal discrimination abilities displayed by individuals with cochlear implants and their speech recognition (Cazals, Pelizzone, Saudan, & Boex, 1994; Dorman, Smith, Smith, & Parkin, 1996). A conflicting finding was found by Iverson (2003), however, who reported that word recognition accuracy was not related to overall sensitivity to a temporal cue, voice onset time, for a group of adults with CIs. In the current study, discrimination measures for spectral and amplitude cues were collected using an AX paradigm to examine the effects of sensitivity to acoustic structure on cue weighting strategies. Sensitivities to the spectral and amplitude structure of these stimuli were also examined as contributors to word recognition, independent of cue weighting.

An important factor in designing stimuli for discrimination testing is a concern for how "speech-like" the stimuli are. Although discrimination tasks with simple, non-speech stimuli are often considered adequate for revealing absolute sensitivity to acoustic properties, that perspective may not be strictly correct. Previous studies have shown that listeners use different strategies for integrating spectral components based on whether stimuli are recognized as being speech-like or not. Consequently, discrimination testing may be affected by stimulus selection (Nittrouer et al., 2013; Remez, Rubin, Berns, Pardo, & Lang, 1994). Because the goal of discrimination testing in this study was to examine sensitivity to acoustic cues for participants with CIs, regardless of how "speech-like" the stimuli were, two kinds of stimuli were used: stimuli that retained some speech-like quality because they were generated with a software speech synthesizer (i.e. *formant stimuli*) and stimuli comprised of sine waves replicating those formants (i.e. *sine-wave stimuli*). The same fundamental frequency and starting and steady-state formant frequencies were used in creating the speech-like stimuli for the discrimination testing and labeling tasks in this study. Sine-wave stimuli were created for discrimination testing from the formant tracks of the formant stimuli. The similarities between stimuli would ensure that the relevant spectral changes would be treated the same by participants' implant processors during both the labeling and discrimination tasks.

Summary

The current experiment was designed to evaluate whether post-lingually deafened adults who get CIs modify their perceptual weighting strategies for speech, and if any modifications affect speech recognition. To investigate these questions, stimuli forming a stop-glide manner contrast were used in a phonemic labeling task. Discrimination testing was performed to examine sensitivity to spectral and amplitude structure delivered through

participants' implants. Word recognition scores were obtained and examined relative to derived cue-weighting factors and discrimination abilities. Finally, vocabulary skills were measured because these skills may influence word recognition.

Method

Participants

Twenty-one adults who wore CIs and were between 18 and 62 years of age enrolled in the study. All were native English speakers and had varying etiologies of hearing loss and ages of implantation. Twenty of the CI users had Cochlear devices, and one used an Advanced Bionics device. All participants had CI-aided thresholds measured by certified audiologists within the 12 months prior to testing. Mean aided thresholds for the frequencies of .25 to 4 kHz were better than 35 dB hearing level for all participants. Six participants had bilateral implants, five used a hearing aid on the ear contralateral to the CI in everyday settings, and ten did not use additional amplification. All participants with Cochlear devices used an Advanced Combined Encoder (ACE) speech processing strategy, except one, who used a spectral peak (SPEAK) strategy. The participant with an Advanced Bionics device used the HiRes Fidelity 120 sound processing strategy. The Appendix shows relevant demographic and treatment data for individual participants.

Socioeconomic status (SES) was collected because it may predict speech and language abilities. SES was determined using a metric described by Nittrouer and Burton (2005), which involved indexing occupational status and educational level. Two eight-point scales were used, with a score of "8" representing the highest occupational status and the highest educational level achieved by the individual. These two scores are then multiplied to give a SES score between 1 and 64. In addition to SES, participants were asked to report on their annual family income. For this purpose, a 4-point scale was used: (1) less than $20,000; (2) between $20,000 and $30,000; (3) between $30,000 and $45,000; or (4) greater than $45,000. Data regarding SES and income are also shown in the Appendix.

Equipment and Materials

All testing took place in a soundproof booth, with the computer that controlled stimulus presentation in an adjacent room. Audiometric testing was done with a Welch Allyn TM262 audiometer using TDH-39 headphones. Stimuli were stored on a computer and presented through a Creative Labs Soundblaster card, a Samson amplifier, and a Roland MA-12C powered speaker. This system has a flat frequency response and low noise. Custom-written software controlled the presentation of the test stimuli. For the labeling and discrimination tasks, the experimenter recorded responses with a keyboard connected to the computer. Two drawings (on 8×8 in. cards) were used to represent each response label for the labeling task: for /bɑ/, a picture of a baby, and for /wɑ/, a picture of the ocean (water). Two additional drawings (on 8×8 in. cards) were used to represent each response label for the discrimination task: for *same*, a picture of two black squares, and for *different*, a picture of a black square and a red circle. These response cards were the same as those used by Nittrouer et al. (2013).

For word recognition testing, the CID W-22 Word List 1 was utilized. This measure is used clinically and consists of 50 monosyllabic words recorded by a male talker. These words were also stored on the computer and played through the same hardware as the test stimuli.

For a measure of receptive vocabulary, the Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4; Dunn & Dunn, 2007), Form A was utilized. This test consists of a booklet that has separate pages for each word. Each page consists of four pictures, one of which represents the target word and three foils.

Stimuli

Synthetic speech stimuli for labeling—For testing, the synthetic stimuli created by Nittrouer et al. (2013) were used. These stimuli consisted of four continua: two continua on which stimuli varied along a continuum in formant rise time (FRT), with two settings of amplitude rise time (ART), and two continua on which stimuli varied along a continuum in ART, with different FRTs. To make these stimuli, /bɑ/-/wɑ/ continua were created using a Klatt synthesizer (Sensyn) with a sampling rate of 10 kHz. The vowel /ɑ/ was used in order to provide onsets with maximum amplitude and formant (at least for F1) excursions: less open vowels would not have had excursions as extensive. All tokens were 370 ms in duration, considerably longer than syllable lengths that evoked shifts in phoneme boundaries along formant transition continua in past experiments (e.g., Miller & Liberman, 1979; Pisoni et al., 1983; Shinn et al., 1985). Fundamental frequency was constant across all stimuli at 100 Hz. Starting and steady-state frequencies of the first two formants were the same for all stimuli, even though the time to reach steady-state frequencies varied. F1 started at 450 Hz and rose to 760 Hz at steady state. F2 started at 800 Hz and rose to 1150 Hz at steady state. F3 was kept constant at 2400 Hz. Controlling ART required that the signals be processed in Matlab because there is no way to reliably control amplitude in Klatt-based speech synthesizers. More details about stimulus generation can be found in Nittrouer et al. (2013), but are described in brief below.

For the two FRT continua (i.e., the "FRT condition"), rise time of the first two formant transitions varied along a 9-step continuum from 30 ms to 110 ms, in 10-ms steps. All stimuli along one of these continua had an ART of 10 ms (the most /bɑ/-like ART) and all stimuli along the other continuum had an ART of 70 ms (the most /wɑ/-like ART). These continua will be termed the /bɑ/-FRT and the /wɑ/-FRT continua, respectively.

For the two ART continua (i.e., the "ART condition"), rise time of the gross temporal envelope of the syllable varied along a 7-step continuum from 10 ms to 70 ms, in 10-ms steps. For one continuum, all stimuli had an FRT of 30 ms (the most /bɑ/-like); for the other continuum, all stimuli had an FRT of 110 ms (the most /wɑ/-like). These continua will be termed the /bɑ/-ART and the /wɑ/-ART continua, respectively.

Figure 3 shows waveforms and spectrograms of a stimulus with the most /bɑ/-like ART and FRT (left panel) and a stimulus with the most /wɑ/-like ART and FRT (right panel). During testing, stimuli were played 10 times each in blocks of however many stimuli there were, so that listeners heard a total of 180 FRT stimuli (9 steps × 2 ARTs × 10 trials) and 140 ART stimuli (7 steps × 2 FRTs × 10 trials). Stimuli in the FRT and ART conditions were

presented separately. In addition to these synthetic stimuli used during testing, five samples each of natural /bɑ/ and /wɑ/ syllables were used for training.

Synthetic stimuli for discrimination—Stimuli were created for four continua. For two continua, stimuli varied in FRT. For the other two continua, stimuli varied in ART. In both cases, one set of stimuli was created with a software speech synthesizer, and the other set was created with sine wave synthesis.

For the "formant FRT continuum," the /bɑ/-FRT stimuli were used, with ART held constant (at 10 ms), so only FRT varied across stimuli. The "sine wave FRT continuum" consisted of sine waves synthesized with Tone (Tice & Carrell, 1997), using the duration, formant frequencies, and FRTs from the formant stimuli. Each FRT continuum consisted of 9 stimuli.

For the "formant ART continuum," stimuli were generated with steady-state formants set to the values used for the labeling stimuli: F1 = 760 Hz; F2 = 1150 Hz; and F3 = 2400 Hz. The ART of the gross temporal envelope of the stimuli varied along an 11-step continuum from 0 ms to 250 ms, in 25-ms steps. The tokens were all 370 ms in duration, with a constant fundamental frequency of 100 Hz, matching values from the labeling stimuli. The "sine wave ART continuum" consisted of sine waves synthesized with Tone (Tice & Carrell, 1997), using the duration, formant frequencies, and ARTs from the formant stimuli. Each ART continuum consisted of 11 stimuli.

Procedures

All procedures were approved by the Institutional Review Board of The Ohio State University. Informed, written consent was obtained from each participant. All participants were tested while wearing just one CI. The participants who typically wore hearing aids did not wear them during testing to avoid the confounding factor of using acoustic hearing for perceptual decisions. For bilateral CI users, testing was performed using their first implant only. This approach was used to ensure that all participants were relying on one implant for perceptual decision-making. As a result of this procedure, 12 participants were tested through CIs on their right ears, and 9 were tested through CIs on their left ears. All stimuli were presented at 68 dB SPL, measured at ear level, via a speaker positioned one meter from the participant at 0 degrees azimuth. The decision was made to present stimuli in this way, rather than directly to participants' CIs, because a free-field presentation would best allow the tester to confirm loudness of presentation and ensure that stimuli were being delivered equally across participants. Breaks in testing could be taken at participants' request.

Before testing, participants completed questionnaires regarding hearing history, education, and socioeconomic status. Audiometry for assessment of residual hearing in the implanted and nonimplanted ears was then performed. This was done to determine whether either ear needed to be plugged during testing. None of the participants had pure-tone average (PTA) thresholds in the nonimplanted or implanted ear better than 68 dB HL for the frequencies of 0.5, 1 and 2 kHz. Therefore, ear plugging was not necessary for any participant during the presentation of stimuli, which was performed at 68 dB SPL.

Participants were tested during two sessions of 60 minutes each on two separate days. In the first session, labeling, word recognition, and vocabulary tasks were presented, in the following order: one of the two labeling tasks (FRT continuum or ART continuum) followed by word recognition testing, then vocabulary assessment, and finally the remaining labeling task. The order of labeling tasks alternated among participants. After all participants completed that first session, they were asked to return for a second session consisting of the discrimination tasks. For discrimination testing, the order of tasks was randomized among participants. Five participants were unable to return for the second testing session.

Labeling tasks—Training tasks were performed immediately prior to testing. For the first training task, the natural /bɑ/-/wɑ/ stimuli were presented. The experimenter introduced each picture separately and told the participant the name of the word associated with that picture. Participants were instructed that they would hear a word and they needed to both point to the picture representing that word and say the word. Having participants both point to the picture and say the word served as a reliability check on responses. The 10 natural tokens were presented in random order, with feedback. Then the stimuli were presented without feedback. Participants were given two opportunities, if needed, to respond at a level of 90% correct within a single block in order to proceed to training with the synthetic stimuli. Any participant who could not respond at this level of accuracy was dismissed without testing.

Training with the synthetic /bɑ/-/wɑ/ stimuli was performed next, using only endpoint stimuli. These endpoints are stimuli with the most /bɑ/-like FRT and ART (30 ms FRT, 10 ms ART) and the most /wɑ/-like FRT and ART (110 ms FRT, 70 ms ART). Training was done as it had been with the natural tokens, and participants were required to obtain 70% correct responses without feedback to move to testing. They were given three opportunities in which to reach that goal. Next, the first test with synthetic /bɑ/-/wɑ/ stimuli along either the FRT or ART continuum was given, with the second test condition given after word recognition and vocabulary testing.

Word recognition—The CID word list was presented. For this task, the words were played over the speaker used for the /bɑ/-/wɑ/ labeling task, and the participant repeated each word. The proportion of words repeated correctly was measured and used as the word recognition score.

Vocabulary—The PPVT-4 was presented next. For each item, the examiner said a word, and the participant responded by repeating the word and then pointing to and stating the number (1 through 4) associated with the picture that represented that word. If the participant did not repeat the word correctly, the examiner would say the word until it was repeated correctly by the participant prior to selecting a picture. Raw scores were calculated along with standard scores.

Discrimination tasks—For discrimination testing, an AX procedure was used. In this procedure, listeners compare a stimulus, varying across trials (X), to a constant standard stimulus (A). For both the formant and sine wave FRT stimuli, the A stimulus was the one with the 30-ms FRT. For both the formant and sine wave ART stimuli, the A stimulus was

the one with the 0-ms ART. All stimuli along the continuum, including the one serving as A, served as X (comparison) stimuli, and 10 trials for each were administered in random order. The interstimulus interval between standard and comparison stimuli was 450 ms. The participant responded by pointing to the card with two black squares and saying *same* if stimuli were judged as the same, and by pointing to the card with the black square and the red circle and saying *different* if stimuli were judged as different. Both pointing and verbal responses were used as a reliability check.

Before testing, participants were shown the cards and required to point to *same* and *different*. Five pairs of stimuli that were identical and five pairs that were maximally different were presented. Participants reported whether the stimuli were the same or different and received feedback. The same training stimuli were then presented without feedback. Correct responses to nine of the ten training trials without feedback were required to proceed to testing.

Analyses

Computation of weighting factors from labeling tasks—To calculate measures of the perceptual weights for the FRT and ART cues, logistic regression was performed. The proportion of /w/ responses given to each stimulus on each continuum was used in the computation of these weighting factors. Values for the continuous property were normalized to values between 0 and 1 to match settings on the binary property. Weighting factors were calculated for both the FRT cue and the ART cue, for each test condition, and averaged across conditions. These FRT and ART weighting factors were used in subsequent statistical analyses.

Computation of d′ values from discrimination tasks—The discrimination functions of each listener were used to compute an average d′ value for each condition (Holt & Carney, 2005; Macmillan & Creelman, 2005). The d′ value was selected as the discrimination measure because it is bias-free. The d′ value is defined in terms of z-values based on a Gaussian normal distribution which converts values into standard deviation units. The d′ value was calculated at each step along the continuum as the difference between the z-value for the "hit" rate (proportion of *different* responses when A and X stimuli were different) and the z-value for the "false alarm" rate (proportion of *different* responses when A and X stimuli were identical). A hit rate of 1.0 and a false alarm rate of 0.0 require a correction in the calculation of d′ and were assigned values of .99 and .01, respectively. A value of zero for d′ means participants cannot discriminate the difference between stimuli. A positive d′ value suggests the "hit" rate was greater than the "false alarm" rate. Using the above correction values, the minimum d′ value would be 0, and the maximum d′ value would be 4.65. The average d′ value was then calculated across all steps of the continuum, and this value was used in regression analyses.

Regression analyses—A series of separate linear regression analyses was performed to try to uncover the pattern and nature of relationships among the various demographic and dependent measures. Standardized β coefficients are reported when $p < .10$; otherwise results are reported as *not significant*.

Results

One participant was unable to recognize nine out of ten of the natural speech tokens correctly, so that individual was not tested. Twenty participants were included in the labeling tasks, word recognition, and vocabulary testing. Sixteen participants underwent testing on the discrimination tasks. One participant was blind and could not perform the vocabulary testing, so that individual was not included in analyses involving the vocabulary score.

First, two-sample t tests were performed to see if side of implant influenced scores for word recognition, vocabulary score, FRT weighting factor, ART weighting factor, d' values for FRT continua, or d' values for ART continua. No differences were found. Next, potential effects of typically using one CI, two CIs, or a CI plus hearing aid were examined. One-way ANOVAs found no differences in group means for word recognition, vocabulary score, FRT weighting factor, ART weighting factor, d' values for FRT continua, or d' values for ART continua based on whether participants used one CI, two CIs, or a CI plus hearing aid. Therefore, data were combined across all participants in subsequent analyses, regardless of side of implantation and typical device use.

Weighting factors for FRT and ART

Table 1 shows means and standard deviations of FRT and ART weighting factors, d' values for FRT and ART continua, word recognition scores, and vocabulary scores. For comparison, the listeners from Nittrouer et al. (2013) had a mean FRT weighting factor across the two stimulus conditions of 9.52 ($SD = 3.10$). Those same listeners had a mean ART weighting factor of 1.10 ($SD = 1.34$). To illustrate the relationship between computed weighting factors and labeling functions in this study, Figure 4 shows labeling functions for the three CI users with the highest FRT weighting factors, and Figure 5 shows labeling functions for the three CI users with the highest ART weighting factors. Individual weighting factors are shown beneath each set of labeling functions. As with Figure 2 from Nittrouer et al., the comparison of labeling functions and weighting factors demonstrates that these weighting factors reflect the labeling patterns of the listeners. Inspection of both labeling functions and weighting factors leads to the conclusion that many of the CI users in this study showed different weighting strategies from those observed for adults with normal hearing in the study by Nittrouer et al. In addition, significant variability was seen in FRT and ART weighting factors among individuals, as can be seen in Figure 6. No correlation was found between FRT and ART weighting factors, which suggests that individuals with CIs do not necessarily attend to one cue at the exclusion of the other. The next question of interest concerned how those weighting strategies affect word recognition.

Vocabulary and word recognition

Prior to addressing the question of the impact of weighting strategies on word recognition, it was important to examine the role that might be played by linguistic knowledge in word recognition. In particular, vocabulary abilities might be expected to influence word recognition independently of weighting strategies. If so, that relationship would need to be taken into account in any analyses of the effects of weighting strategies on word recognition.

Therefore, the question was asked of how strongly vocabulary abilities predicted word recognition. To answer that question, vocabulary score was used as the predictor variable in regression analysis with word recognition score as the dependent measure. No significant relationship was found between these two factors, so examination of the relationship between weighting factors and word recognition proceeded without regard to vocabulary knowledge.

Weighting factors and word recognition

Using the FRT and ART weighting factors derived from the /bɑ/-/wɑ/ labeling task as indicators of the extent to which time-varying spectral and amplitude structure are relied upon in listening situations, it was asked if these weighting factors explain significant proportions of variability in word recognition. To do that, linear regression analyses were performed separately for each weighting factor, using word recognition as the dependent measure. These analyses revealed a significant relationship between FRT weighting factors and word recognition scores, standardized $\beta = .77$, $p < .001$. To illustrate this relationship, a scatter plot of the two scores is shown in Figure 7. No significant relationship was found between ART weighting factors and word recognition scores. Thus it was concluded that the extent to which FRT was weighted explained word recognition to a considerable extent, but weighting of ART had no influence on word recognition.

Discrimination of FRT and ART

From the previous analyses, it was still unclear whether the individuals with CIs tested in this study were sensitive to the acoustic cues presented. Sensitivity would be necessary in order for a cue to be weighted in a perceptual decision. Means and standard deviations of d′ values for FRT discrimination and ART discrimination were calculated for both formant stimuli and sine-wave stimuli, and are presented in Table 1. For comparison, a d′ value of 0 suggests no ability to discriminate the cue, a d′ value of 2.33 suggests the individual could discriminate 50% of different stimuli as different, and a d′ value of 4.65 suggests 100% discrimination of different stimuli as different. Overall, participants showed significant variability, but could discriminate about 50% of FRT and ART cues, except for the FRT sine-wave condition, in which they did more poorly as a group and could discriminate less than 5% of the FRT differences. To illustrate the relationship between computed d′ values and discrimination functions, Figure 8 shows discrimination functions for a participant with relatively good sensitivity to both FRT and ART (left side), as well as for a participant with relatively poor sensitivity to FRT, but good sensitivity to ART (right side).

The obtained d′ values were used as predictor variables in linear regression analyses with word recognition scores as the dependent measure. The d′ value for the formant FRT continuum predicted word recognition, $\beta = .55$, $p = .034$, and a similar trend was seen for the d′ value for the sine wave FRT continuum and word recognition, $\beta = .50$, $p = .055$, although it did not reach statistical significance. No significant relationships were found for the d′ values for the ART cues and word recognition. Next, d′ values for FRT discrimination and ART discrimination were used as predictor variables, with FRT weighting factor and ART weighting factor as dependent measures. An effect was seen between FRT weighting factor and the d′ value for FRT discrimination for formant stimuli,

standardized β = .56, p = .030, with a trend towards a significant relationship between FRT weighting and the d′ value for FRT discrimination for sine wave stimuli, standardized β = .49, p = .065. No effect was seen between ART weighting factor and d′ values for ART discrimination of either formant or sine wave stimuli. In addition, no significant relationships were found between d′ values for FRT discrimination and ART weighting factor, or between d′ values for ART discrimination and FRT weighting factor.

Because the tendency to weight spectral structure heavily and the discrimination of spectral structure for the formant stimuli were both predictors of word recognition, it was important to examine whether sensitivity to spectral structure would independently lead to better word recognition. Therefore, a stepwise linear regression was performed with FRT weighting factor and the d′ value for FRT discrimination for formant stimuli as predictor variables, with word recognition as the dependent measure. After the variance in word recognition predicted by FRT weighting was removed, no unique variance in word recognition could be attributed to sensitivity.

Cue weighting, age of onset of hearing loss, and word recognition

Next, beyond sensitivity to spectral structure, the question of what determined weighting strategies for these CI users was examined. To address this question, independent demographic factors as well as vocabulary score were used as predictor variables in separate linear regression analyses with FRT weighting factor as the dependent measure. Socioeconomic status, income level, vocabulary abilities, and gender all failed to predict FRT weighting factors. However, age of onset of hearing loss was found to predict significant variance in FRT weighting, standardized β = .525, p = .017. According to this result, the longer an individual had normal, or close to normal hearing, the more weight that individual assigned to FRT. Even though weighting of ART was not found to explain any significant proportion of the variance in word recognition, age of onset of hearing loss was used in regression analysis with ART weighting. This was done to see if CI users who lost their hearing early in life tended to weight ART heavily. No significant effect was found.

The results described so far suggest that having perceptual weighting strategies similar to those of adults with normal hearing is important to having good word recognition for listeners with CIs. Longer periods of time with normal, or near-normal hearing seem to be responsible for the development of those typical weighting strategies. But there is one potential problem in too readily accepting these suggestions: It could be that longer periods of normal hearing simply improve word recognition abilities, independent of perceptual weighting strategies for acoustic cues. That seemed fairly unlikely in this case, because if such an effect were to exist it would probably have its influence through vocabulary knowledge and vocabulary scores were not found to explain any significant variance in word recognition. Nonetheless, the proposal warranted investigation. To do so, word recognition score was used as the dependent measure in regression analysis with age of onset of hearing loss as the predictor variable. Indeed a significant effect was found, standardized β = .57, p = .009. However, it was possible that this relationship could reflect the dependence of weighting factors on experience with normal hearing. In other words, the variance explained in word recognition scores by age of onset of hearing loss could be redundant with that

explained by weighting strategies. To answer that question, a stepwise linear regression was performed, using FRT weighting factor and age of onset of hearing loss as predictor variables, and word recognition score as the dependent measure. Once the variance attributed to FRT weighting was removed, no unique variance in word recognition could be explained by age of onset of hearing loss. Thus, it was reasonable to conclude that age of onset of hearing loss had its primary effect on perceptual weighting strategies for acoustic cues, which in turn explained how well these CI users could recognize words.

Subset analysis based on age of onset of hearing loss

In order to examine the relationships between cue-weighting strategies, word recognition, and discrimination, subjects had been deliberately selected who had a wide range of abilities and ages of onset of hearing loss. Although all individuals were implanted post-lingually, meaning they had some language development prior to their hearing loss advancing to the point of qualification for an implant, the age of onset of hearing loss varied from some hearing loss at birth to age 40 years. This heterogeneity of age of onset of hearing loss could potentially affect whether all participants perceived the stimuli categorically. Therefore, the group was split into an "early onset" hearing loss group, defined as onset of hearing loss before age 13 years, and a "late onset" hearing loss group, defined as onset of hearing loss at or after age 13. Although somewhat arbitrary, this age was selected based on evidence suggesting that a sensitive period for first language acquisition exists and ends around puberty (Hurford, 1991; Ingram, 1989). Twelve participants fit the criterion for "early onset" hearing loss, and eight fit the criterion for "late onset" hearing loss.

For each group separately, linear regression analyses were performed for each weighting factor, using word recognition as the dependent measure. These analyses revealed a significant relationship between FRT weighting and word recognition score for both the "early onset" group, standardized $\beta = .64$, $p = .024$, and the "late onset" group, standardized $\beta = .90$, $p = .002$. No relationship was found between ART weighting and word recognition for either group. Within the "early onset" group, age of onset of hearing loss predicted FRT weighting, standardized $\beta = .67$, $p = .017$. The same was not found for the "late onset" group. No relationship was found between age of onset of hearing loss and word recognition for either the "early onset" or "late onset" group separately. The d′ values for FRT continua or ART continua did not predict word recognition or weighting factors for either the "early onset" or "late onset" group. Lastly, no relationship was found between chronological age and word recognition for the "late onset," "early onset," or entire group.

Discussion

The experiment presented here was conducted to examine whether adult CI users demonstrate different weighting strategies for acoustic cues to phonemic categorization than do listeners with normal hearing, and whether any adjustments in those weighting strategies affect word recognition in general. These two goals were achieved using a labeling task for the /bɑ/-/wɑ/ contrast in which normal-hearing adults heavily weight spectral structure. In addition, data on word recognition and discriminative capacities were collected.

Regarding the first goal, the results of this study revealed that indeed some adult CI users demonstrated different weighting strategies from those of listeners with normal hearing: Instead of weighting the spectral cue almost exclusively, some CI users paid attention to the amplitude cue. Because these two acoustic cues are perfectly correlated in the /bɑ/-/wɑ/ contrast used in this experiment it might be predicted that both cues would be equally as effective in supporting speech perception.

The second goal of this study was to examine just that question: how weighting strategies affected word recognition. In this study, listeners' weighting of formant and amplitude structure was viewed as a metric of whether they tended to use the weighting strategies commonly employed by listeners with normal hearing or use strategies based on the salience of acoustic structure through their implants. Accordingly, weighting factors for spectral and amplitude structure were used as predictor variables in regression analyses with word recognition scores. It was found that the weighting of spectral structure (FRT) was the only independent predictor of word recognition for adult CI users. The individuals who continued to use the weighting strategies typically used by normal-hearing adults had the best word recognition. These findings provided support for the ***same, but adapted, strategies*** hypothesis: the ability to perceptually weight acoustic cues similarly to adults with normal hearing, even with the degraded spectral cues delivered by cochlear implants, was associated with the best word recognition outcomes.

In addition to those two primary goals, this study examined the relationship between weighting strategies and sensitivity to spectral and amplitude structure. It was found that sensitivity to FRT, as assessed with the discrimination tasks, did explain CI users' weighting of FRT cues to some extent. These findings may simply confirm that sensitivity to FRT is a requisite for weighting FRT strongly. Certainly individuals would have to be sensitive to a spectral cue to be able to use it in a phonemic decision. Another possible explanation is that greater sensitivity to a spectral cue may encourage use of this cue in perceptual weighting. However, this explanation can be refuted because sensitivity to ART did not explain ART cue weighting. Basic sensitivity to spectral and amplitude structure did not explain variance in word recognition, whereas weighting of spectral structure strongly explained variance in word recognition.

Age of onset of hearing loss is known to be a factor in speech perception outcomes for CI users. The findings of this study are in agreement, with age of onset of hearing loss serving as a significant predictor of word recognition scores, standardized $\beta = .57$. However, even greater variance in word recognition scores could be attributed to spectral weighting, standardized $\beta = .77$. Spectral cue weighting itself appeared to explain variance in word recognition above and beyond any effect of age of onset of hearing loss, and spectral cue weighting predicted variance in word recognition for both the "early onset" and the "late onset" hearing loss groups. Nonetheless, those individuals who lost their hearing later in life were more likely to weight spectral structure heavily. These findings suggest that individuals with later onset of hearing loss had already developed the most efficient perceptual weighting strategies. The ability to continue exercising these strategies as normal-hearing listeners do led to the best outcomes in word recognition.

Even though only one phonemic contrast was examined in this study, the findings are in agreement with the results of previous studies using different phonetic contrasts to examine cue weighting by individuals with CIs. Individual variability in cue weighting has been observed to be high, as was seen in testing of vowel perception based on formant movement and duration cues (Iverson et al., 2006) and perception of the /d/-/t/ contrast based on voice onset time (Iverson, 2003). The current study found that word recognition accuracy was related to using cue-weighting strategies that are most like those used by adults with normal hearing, independent of their discrimination abilities. Similarly, Iverson (2003) found the best speech recognition in adults with CIs who had relatively normally-located peaks on discrimination functions, even if overall level of sensitivity was poor. That finding provides independent support for the *same, but adapted, strategies* hypothesis. Lastly, Giezen and colleagues (2010) found that children with CIs weighted spectral cues more heavily than duration cues, as do children with normal hearing, in spite of the poor spectral resolution delivered by their implants.

The results of the study reported here have significant clinical ramifications. A great deal of research with CI users is focused on improving signal presentation by implants to deliver the necessary cues for speech understanding. The findings of this study show that just having access to the acoustic structure that underlies speech recognition is not sufficient to ensure good recognition; optimal weighting of the various components of that structure is equally important. These perceptual weighting strategies form one aspect of what might be termed *top-down* effects, and all results of the current study emphasize the fact that speech perception is as much determined by these effects as by the kinds of signal properties that are available. This study does not irrefutably demonstrate that adopting weighting strategies most like those used by listeners with normal hearing leads to improved speech recognition for those using CIs. Nonetheless, the demonstration of the value of top-down effects opens the door for the development of auditory training protocols that may encourage CI users to shift their weighting strategies to those that are most effective in speech perception. Shifts in weighting strategies have been seen as a result of auditory training using nonspeech stimuli (Holt & Lotto, 2006), and it is likely that CI users could benefit from training using speech stimuli.

The results of this study showed that weighting of spectral structure heavily was indeed somewhat related to sensitivity to that structure. Better sensitivity to spectral cues, however, was not sufficient for effective speech recognition. Studies of second-language learners reveal that they have adequate sensitivity to the acoustic cues that underlie phonemic decisions in the second language. Nonetheless, if those cues are not relevant in the first language, they will not be weighted strongly in the second language, at least not without adequate experience. Thus, sensitivity to spectral cues appears necessary but not sufficient for spectral weighting, and it is speculated that only those with sufficient experience through normal, or close to normal hearing weighted spectral cues heavily. Further work will help to clarify the relative roles of peripheral sensitivity to the speech structure delivered through implants and the higher level perceptual organization of CI users.

Conclusion

The ability to weight the acoustic cues to phonemic categories in fashion similar to how normal-hearing adults weight them appears to be necessary for highly accurate word recognition in CI users. The findings of this study could influence the design and programming of auditory prostheses to enhance spectral sensitivity. In addition, auditory training protocols may benefit by placing the emphasis on the retention or development of typical weighting strategies for the acoustic cues to phonemic categories.

Acknowledgments

References

Ambrosch P, Muller-Deile J, Aschendorff A, Anje L, Laszig R, Boermans PP, Sterkers O. European adult multi-centre HiRes[REGISTERED] 120 study--an update on 65 subjects. Cochlear Implants International. 2010; 11(Suppl 1):406–411. [PubMed: 21756661]

Beddor PS, Strange W. Cross-language study of perception of the oral-nasal distinction. Journal of the Acoutical Society of America. 1982; 71:1551–1561.

Benkí JR. Place of articulation and first mormant transition pattern both affect perception of voicing in English. Journal of Phonetics. 2001; 29:1–22.

Best, CT. The emergence of native-language phonological influences in infants: A perceptual assimilation model. In: Goodman, JC.; Nusbaum, HC., editors. The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words. Cambridge: MIT Press; 1994. p. 167-224.

Best CT, McRoberts GW, Goodell E. Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. Journal of the Acoustical Society of America. 2001; 109:775–794. [PubMed: 11248981]

Cazals Y, Pelizzone M, Saudan O, Boex C. Low-pass filgering in amplitude modulation detection associated with vowel and consonant identification in subjects with cochlear implants. Journal of the Acoustical Society of America. 1994; 96:2048–2054. [PubMed: 7963020]

David EE, Ostroff JM, Shipp D, Nedzelski JM, Chen JM, Parnes LS, Sequin C. Speech coding strategies and revised cochlear implant candidacy: An analysis of post-implant performance. Otology & Neurotology. 2003; 24:228–233. [PubMed: 12621337]

Dorman MF, Dankowski K, McCandless G, Parkin JL, Smith L. Vowel and consonant recognition with the aid of a multichannel cochlear implant. Quarterly Journal of Experimental Psychology A. 1991; 43:585–601.

Dorman MF, Smith LM, Smith M, Parkin JL. Frequency discrimination and speech recognition by patients who use the Ineraid and continuous interleaved sampling cochlear-implant signal processors. Journal of the Acoustical Society of America. 1996; 99:1174–1184. [PubMed: 8609301]

Dunn, L.; Dunn, D. Peabody Picture Vocabulary Test. 4. Bloomington: Pearson Education Inc; 2007.

Firszt JB, Holden LK, Skinner MW, Tobey EA, Peterson A, Gaggl W, Wackym PA. Recognition of speech presented at soft to loud levels by adult cochlear implant recipients of three cochlear implant systems. Ear and Hearing. 2004; 25:375–387. [PubMed: 15292777]

Flege, JE. Second language speech learning: Theory, findings, and problems. In: Strange, W., editor. Speech Perception and Linguistic Experience: Issues in Cross-Language Research. Baltimore: York Press; 1995. p. 233-277.

Friesen LM, Shannon RV, Baskent D, Wang X. Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. Journal of the Acoustical Society of America. 2001; 110:1150–1163. [PubMed: 11519582]

Giezen MR, Escudero P, Baker A. Use of acoustic cues by children with cochlear implants. Journal of Speech, Language, and Hearing Research. 2010; 53:1440–1457.

Gottfried TL. Effects of consonant context on the perception of French vowels. Journal of Phonetics. 1984; 12:91–114.

Greenlee M, Ohala JJ. Phonetically motivated parallels between child phonology and historical sound change. Language Sciences. 1980; 2:283–308.

Hamzavi J, Baumgartner W, Pok SM, Franz P, Gstoettner W. Variables affecting speech perception in postlingually deaf adults following cochlear implantation. Acta Otolaryngolica. 2003; 123:493–498.

Hedrick MS, Carney AE. Effect of relative amplitude and formant transitions on perception of place of articulation by adult listeners with cochlear implants. Journal of Speech, Language, and Hearing Research. 1997; 40:1445–1457.

Hicks BC, Ohde RN. Developmental role of static, dynamic, and contextual cues in speech perception. Journal of Speech, Language, and Hearing Research. 2005; 48:960–974.

Holt RF, Carney AE. Multiple looks in speech sound discrimination in adults. Journal of Speech, Language, and Hearing Research. 2005; 48:922–943.

Holt LL, Lotto AJ. Cue weighting in auditory categorization: Implications for first and second language acquisition. Journal of the Acoustical Society of America. 2006; 119:3059–3071. [PubMed: 16708961]

Hurford JR. The evolution of the critical period for language acquisition. Cognition. 1991; 40:159–201. [PubMed: 1786674]

Ingram, D. First language acquisition: method, description, and explanation. New York: Cambridge University Press; 1989.

Iverson P. Evaluating the function of phonetic perceptual phenomena within speech recognition: An examination of the perception of /d/-/t/ by adult cochlear implant users. Journal of the Acoustical Society of America. 2003; 113:1056–1064. [PubMed: 12597198]

Iverson P, Smith CA, Evans BG. Vowel recognition via cochlear implants and noise vocoders: Effects of formant movement and duration. Journal of the Acoustical Society of America. 2006; 120:3998–4006. [PubMed: 17225426]

Jusczyk, PW.; Hohne, EA.; Mandel, DR. Picking up regularities in the sound structure of the native language. In: Strange, W., editor. Speech Perception and Linguistic Experience: Issues in Cross-Language Research. Baltimore: York Press; 1995. p. 91-119.

Keifer J, von Ilberg C, Reimer B. Results of cochlear implantation in patients with severe to profound hearing loss-implications for patient selection. Audiology. 1998; 37:382–395. [PubMed: 9888194]

Liberman AM. Some results of research on speech perception. Journal of the Acoustical Society of America. 1957; 29:117–123.

Liberman IY, Shankweiler D, Fischer FW, Carter B. Explicit syllable and phoneme segmentation in the young child. Journal of Experimental Child Psychology. 1974; 18:201–212.

Macmillan, NA.; Creelman, CD. Detection Theory: A User's Guide. Mahwah, NJ: Lawrence Earlbaum Associates; 2005.

Mayo C, Scobbie JM, Hewlitt N, Waters D. The influence of phonemic awareness development on acoustic cue weighting strategies in children's speech perception. Journal of Speech and Hearing Research. 2003; 46:1184–1196.

McMurray B, Jongman A. What information is necessary for speech categorization? Harnessing variability in the speech signal by integrating cues computed relative to expectations. Psychological Review. 2011; 118:219–246. [PubMed: 21417542]

Miller JL, Liberman AM. Some effects of later-occurring information on the perception of stop consonant and semivowel. Perception & Psychophysics. 1979; 25:457–465. [PubMed: 492910]

Miller JL, Wayland SC. Limits on the limitations of context conditioned effects in the perception of [b] and [w]. Perception & Psychophysics. 1993; 54:205–210. [PubMed: 8361836]

Miyawaki K, Strange W, Verbrugge RR, Liberman AM, Jenkins JJ, Fujimura O. An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. Perception & Psychophysics. 1975; 18:331–340.

Morrison GS. Analysis of categorical response data: Use logistic regression rather than endpoint-difference scores or discriminant analysis (L). Journal of the Acoustical Society of America. 2009; 126:2159–2162. [PubMed: 19894795]

Nie K, Barco A, Zeng FG. Spectral and temporal cues in cochlear implant speech perception. Ear and Hearing. 2006; 27:208–217. [PubMed: 16518146]

Nittrouer S. Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. Journal of Phonetics. 1992; 20:351–382.

Nittrouer S. Discriminability and perceptual weighting of some acoustic cues to speech perception by 3-year-olds. Journal of Speech & Hearing Research. 1996; 39:278–297. [PubMed: 8729917]

Nittrouer S. Learning to perceive speech: How fricative perception changes, and how it stays the same. Journal of the Acoustical Society of America. 2002; 112:711–719. [PubMed: 12186050]

Nittrouer S. Age-related differences in weighting and masking of two cues to word-final stop voicing in noise. Journal of the Acoustical Society of America. 2005; 118:1072–1088. [PubMed: 16158662]

Nittrouer S, Burton LT. The role of early language experience in the development of speech perception and phonological processing abilities: Evidence from 5-year-olds with histories of otitis media with effusion and low socioeconomic status. Journal of Communication Disorders. 2005; 38:29–63. [PubMed: 15475013]

Nittrouer S, Lowenstein JH, Tarr E. Amplitude rise time does not cue the /bɑ/-/wɑ/ contrast for adults or children. Journal of Speech, Language, and Hearing Research. 2013; 56:427–440.

Nittrouer S, Manning C, Meyer G. The perceptual weighting of acoustic cues changes with linguistic experience. Journal of the Acoustical Society of America. 1993; 94:S1865.

Nittrouer S, Miller ME. Predicting developmental shifts in perceptual weighting schemes. Journal of the Acoustical Society of America. 1997; 101:2253–2266. [PubMed: 9104027]

Nittrouer S, Studdert-Kennedy M. The stop-glide distinction: Acoustic analysis and perceptual effect of variation in syllable amplitude envelope for initial /b/ and /w/. Journal of the Acoustical Society of America. 1986; 80:1026–1029. [PubMed: 3771922]

Nittrouer S, Studdert-Kennedy M. The role of coarticulatory effects in the perception of fricatives by children and adults. Journal of Speech & Hearing Research. 1987; 30:319–329. [PubMed: 3669639]

Peterson NR, Pisoni DB, Miyamoto RT. Cochlear implants and spoken language processing abilities: Review and assessment of the literature. Restorative Neurology and Neuroscience. 2010; 28:237–250. [PubMed: 20404411]

Pisoni DB, Carrell TD, Gans SJ. Perception of the duration of rapid spectrum changes in speech and nonspeech signals. Perception & Psychophysics. 1983; 34:314–322. [PubMed: 6657432]

Remez RE, Rubin PE, Berns SM, Pardo JS, Lang JM. On the perceptual organization of speech. Psychological Review. 1994; 101:129–156. [PubMed: 8121955]

Repp BH. Phonetic trading relations and context effects: new evidence for a phonetic mode of perception. Psychological Bulletin. 1982; 92:81–110. [PubMed: 7134330]

Shinn PC, Blumstein SE, Jongman A. Limitations of context-conditioned effects on the perception of [b] and [w]. Perception and Psychophysics. 1985; 38:397–407. [PubMed: 3831918]

Shipp DB, Nedzelski JM. Prognostic indicators of speech recognition performance in adult cochlear implant users: A prospective analysis. Annals of Otology Rhinology and Laryngology. 1995; 166(Suppl 1):194–196.

Smits R, Sereno J, Jongman A. Categorization of sounds. Journal of Experimental Psychology Human Perception and Performance. 2006; 32:733–754. [PubMed: 16822135]

Strange, W. Speech input and the development of speech perception. In: Kavanagh, JF., editor. Otitis Media and Child Development. Parkton: York Press; 1986. p. 12-26.

Strange, W. Learning Non-native Phoneme Contrasts: Interactions Among Subject, Stimulus, and Task Variables. In: Tohkura, E.; Vatikiotis-Bateson, E.; Sagisaka, Y., editors. Speech Perception, Production, and Linguistic Structure. Tokyo: Ohmsha; 1992. p. 197-219.

Tice, B.; Carrell, T. TONE: Tone-analog waveform synthesizer [Computer software]. Lincoln, NE: University of Nebraska; 1997.

Underbakke M, Polka L, Gottfried TL, Strange W. Trading relations in the perception of /r/-/l/ by Japanese learners of English. Journal of the Acoustical Society of America. 1988; 84:90–100. [PubMed: 3411058]

Walsh MA, Diehl RL. Formant transition duration and amplitude rise time as cues to the stop/glide distinction. Quarterly Journal of Experimental Psychology. 1991; 43:603–620. [PubMed: 1775659]

Wardrip-Fruin C, Peach S. Developmental aspects of the perception of acoustic cues in determining the voicing feature of final stop consonants. Language and Speech. 1984; 27:367–379. [PubMed: 6536845]

Werker JF, Logan JS. Cross-language evidence for three factors in speech perception. Perception & Psychophysics. 1985; 37:35–44. [PubMed: 3991316]

Wilson BS, Dorman MF. Cochlear implants: Current designs and future possibilities. Journal of Rehabilitation Research and Development. 2008; 45:695–730. [PubMed: 18816422]

Xu L, Thompson CS, Pfingst BE. Relative contributions of spectral and temporal cues for phoneme recognition. Journal of the Acoustical Society of America. 2005; 117:3255–3267. [PubMed: 15957791]

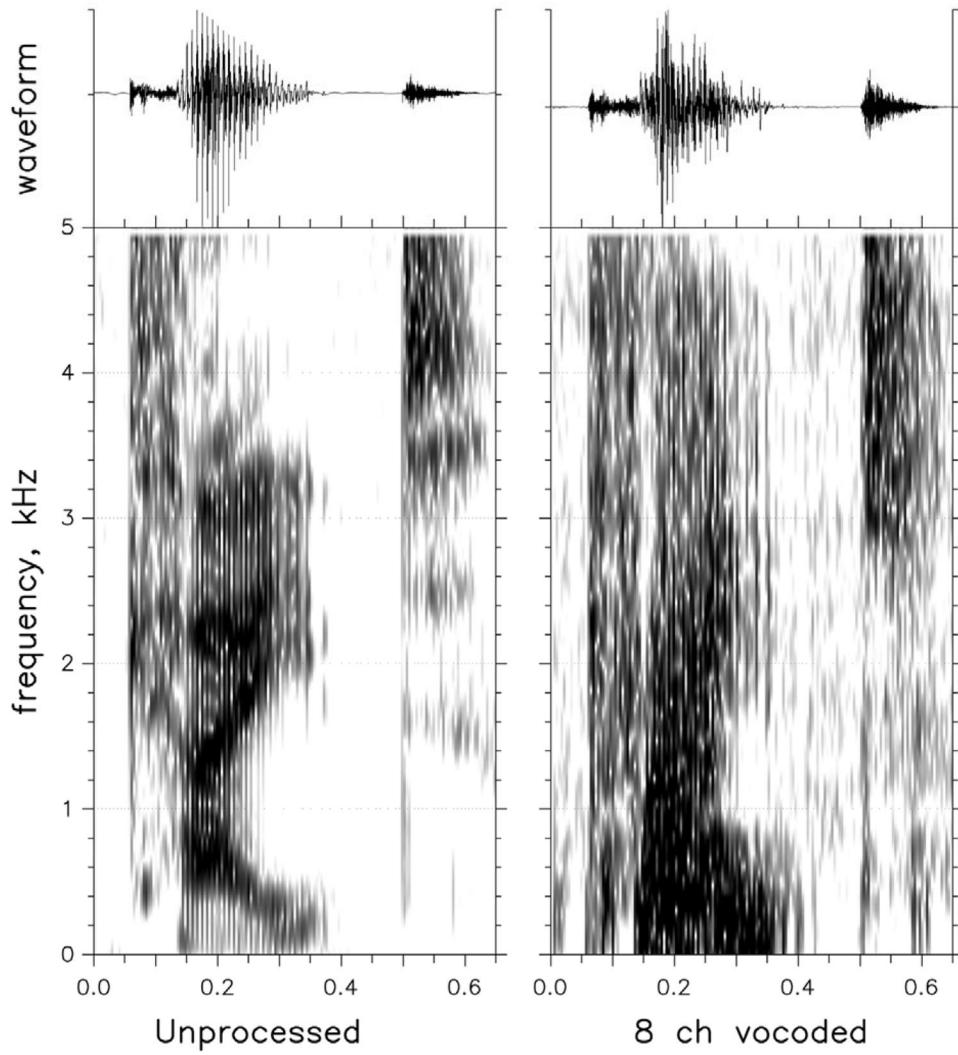Zeng FG. Trends in cochlear implants. Trends in Amplification. 2004; 8:1–34. [PubMed: 15247993]

Figure 1.
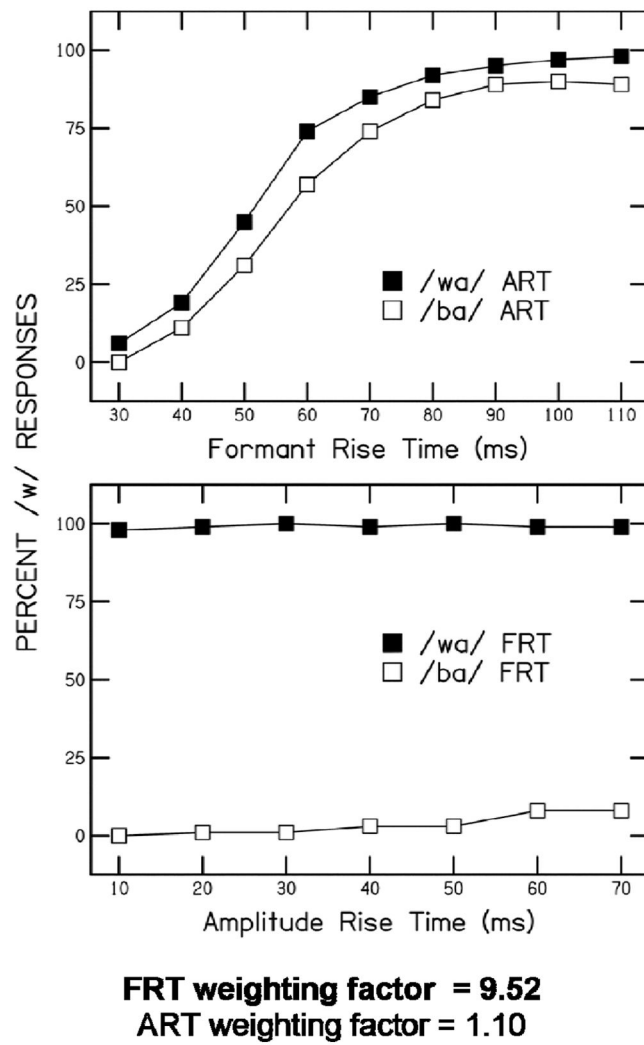Waveforms and spectrograms of the word "kite." Left: Natural stimulus. Right: Vocoded stimulus.

Figure 2.
Labeling functions and weighting factors for adults for synthetic /bɑ/ and /wɑ/ stimuli. (Reprinted from Nittrouer et al., 2013)
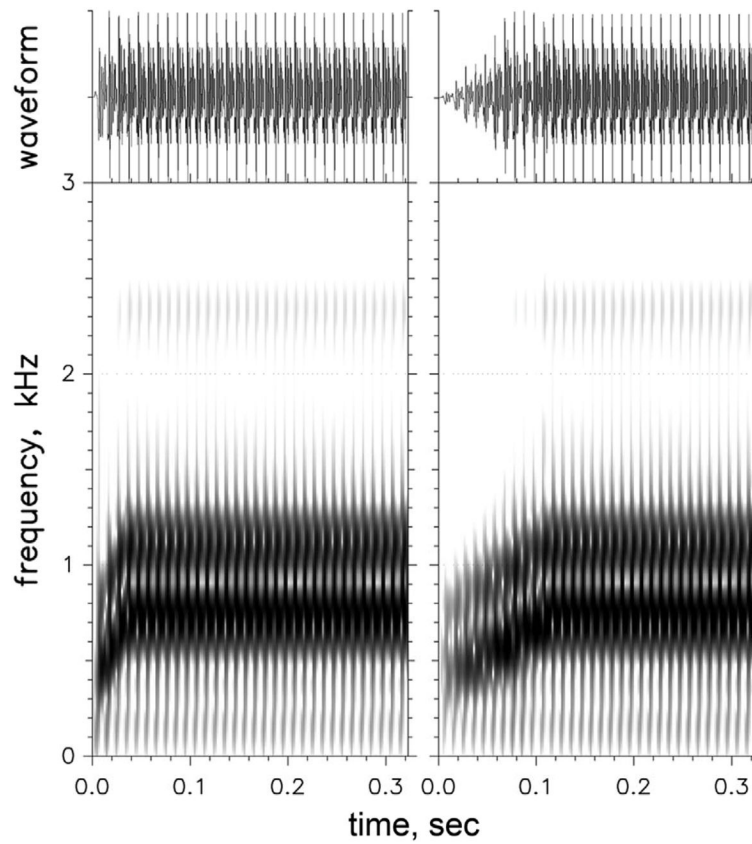
Figure 3.
Waveforms and spectrograms of synthetic stimuli. Left: Stimulus with most /bɑ/-like amplitude rise time (ART) seen in the waveform above and most /bɑ/-like formant rise time (FRT) seen in the spectrogram below; Right: Stimulus with most /wɑ/-like ART and FRT.
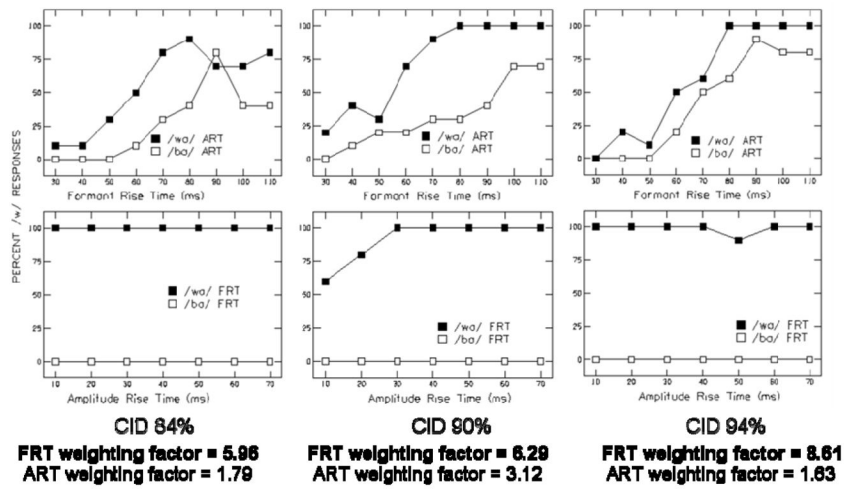
Figure 4.
Results for word recognition (CID W-22 word list) and labeling tasks for the three participants with the strongest weighting of spectral structure. Participants from left to right: CI01, CI05, CI14. FRT: Formant rise time; ART: Amplitude rise time
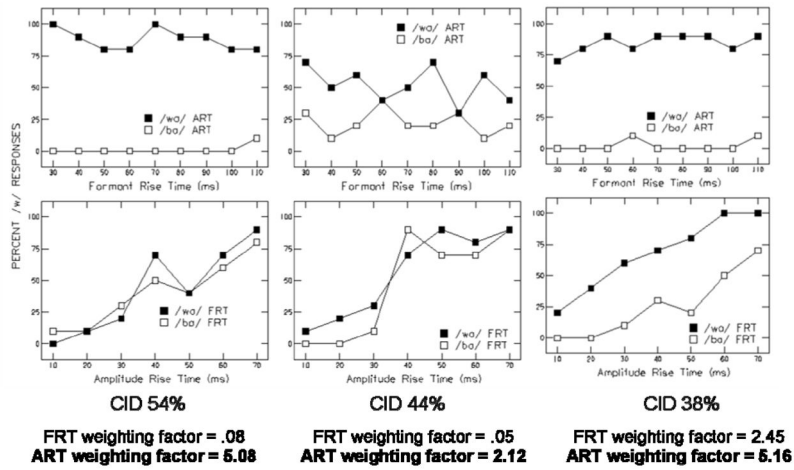
Figure 5.
Results for word recognition (CID W-22 word list) and labeling tasks for the three participants with the strongest weighting of amplitude structure. Participants from left to right: CI03, CI12, CI13. FRT: Formant rise time; ART: Amplitude rise time
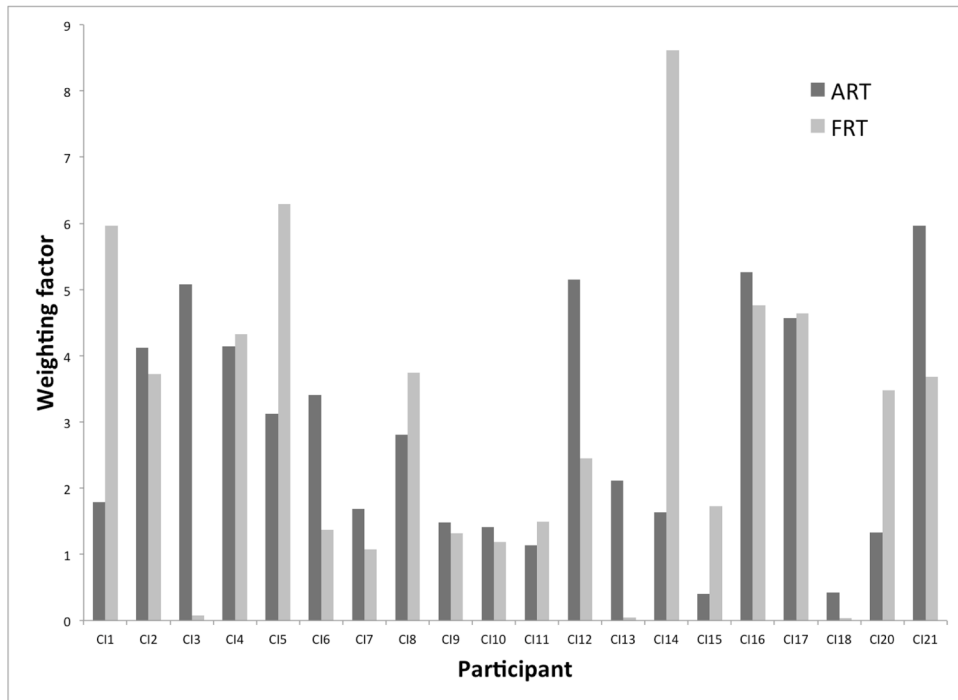
Figure 6.
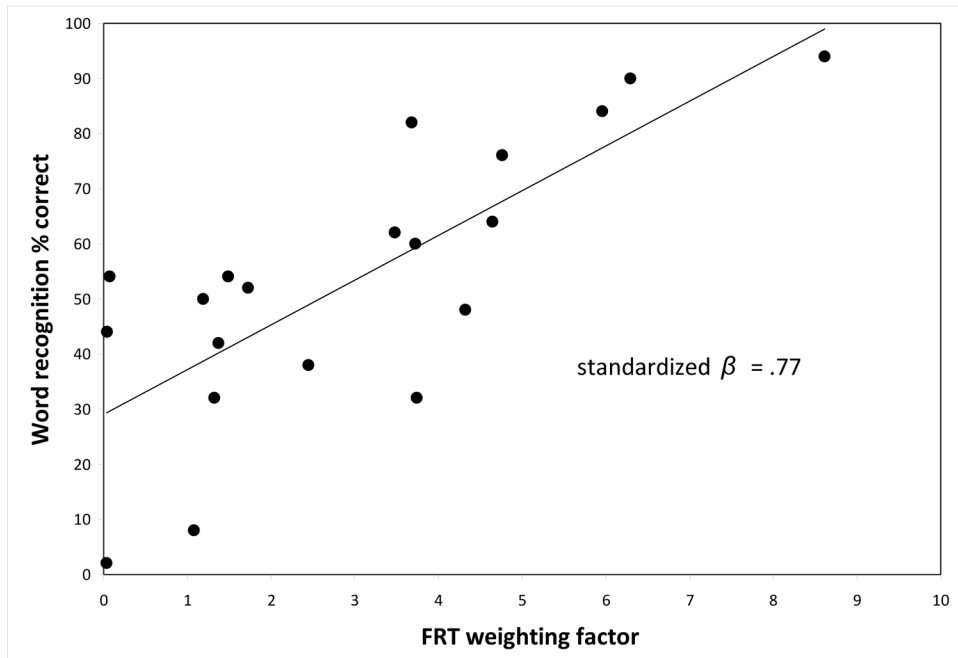ART and FRT weighting factors for individual participants.

Figure 7.
Scatter plot of word recognition (CID W-22 word list) versus FRT weighting. FRT: Formant rise time
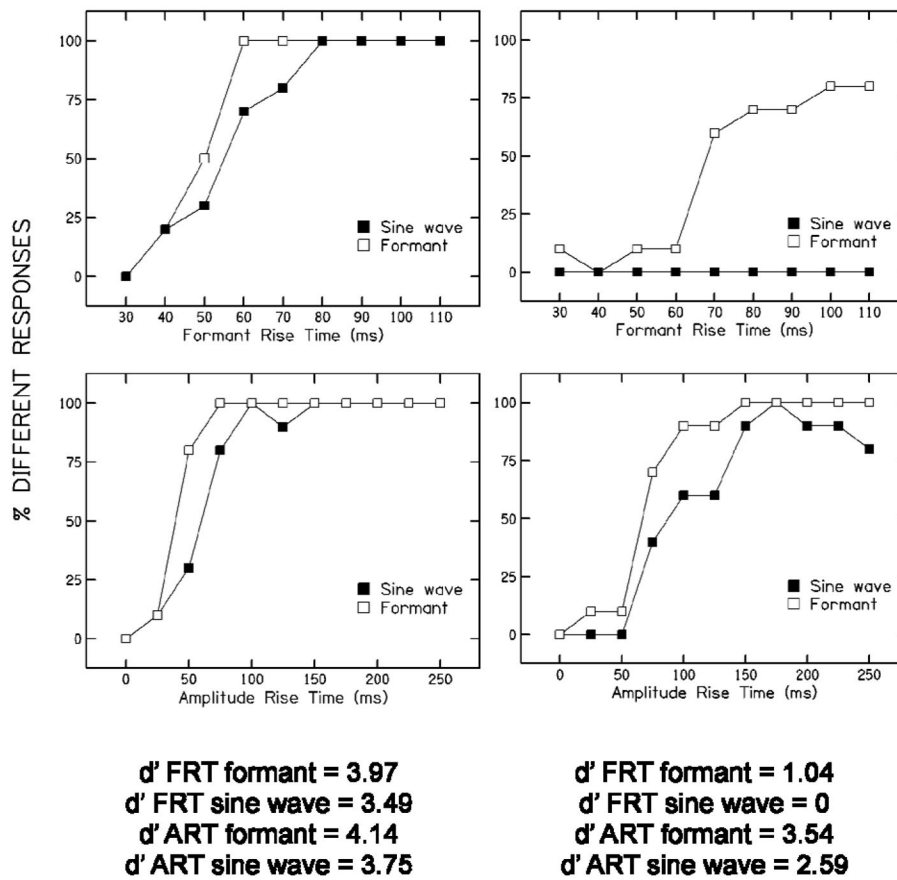
Figure 8.
Results for discrimination tasks and d′ values for two participants with different sensitivity to spectral and amplitude structure.
Participants from left to right: CI01, CI20. FRT: Formant rise time; ART: Amplitude rise time

Table 1

Means and standard deviations of formant rise time (FRT) and amplitude rise time (ART) weighting factors, FRT and ART d′ values, word recognition scores, and vocabulary scores

Measure	Mean	S.D.
FRT weighting factor	3.00	2.33
ART weighting factor	2.85	1.74
d′ FRT formant	2.34	1.18
d′ FRT sine wave	1.54	1.32
d′ ART formant	2.59	1.26
d′ ART sine wave	2.54	1.12
Word recognition score (% correct)	53	24
Standard vocabulary score	98	27

Note: d′ values are average d′ across each FRT and ART continuum.

Appendix

Participant demographics

Participant	Age (years)	Gender	Implant	Processor	Contralateral Aid	Speech PTA Contralateral Ear (dB HL)	Age of Onset of Hearing Loss (years)	Implantation Age (years)	Cause of Hearing Loss	SES	Income level
1	37	M	Freedom	Freedom	Yes	No response	20	30	Progressive as adult	30	4
2	40	M	Nucleus 24	CP810	No	No response	20	32	Menieres	30	4
3	30	F	Nucleus 24	Freedom	No	No response	0	21	Congenital	30	4
4	31	M	Freedom	CP810	Bilateral CI	No response	3	23	Progressive as child	42	4
5	29	M	Freedom	CP810	No	No response	3	25	Progressive as child	9	4
6	29	M	Advanced Bionics	AB Harmony	No	78.3	1	22	Ototoxicity as infant	12	1
7	32	F	Nucleus 22	Freedom	No	No response	2.5	9.5	Meningitis & ototoxicity	30	4
8	29	F	CI512	CP810	No	No response	1.5	27	Meningitis & progressive as child	36	4
9	34	F	Nucleus 24	CP810	No	No response	0	25	Congenital progressive	9	4
10	18	F	Cochlear	Unknown	Yes	No response	3	16	Progressive as child	12	4
11	37	M	Freedom	Freedom	Bilateral CI	No response	33	33	Meningitis as adult	6	4
12	54	M	Nucleus 24	CP810	Bilateral CI	No response	0	48	Congenital progressive	18	3
13	47	F	Nucleus 24	Freedom	No	No response	0	37	Congenital progressive	9	3
14	60	F	Freedom	Freedom	Bilateral CI	No response	40	54	Progressive as adult	18	4
15	46	M	Nucleus 24	Freedom	Bilateral CI	No response	18	38	Progressive as adult	10	4
16	62	M	CI512	CP810	No	No response	14	60	Menieres	24	4
17	52	F	Freedom	Freedom	Yes	68.3	3	48	Progressive as child	20	4
18	40	F	Freedom	Freedom	No	No response	0	33	Congenital progressive	6	3
19	57	F	CI512	CP810	Bilateral CI	No response	0	54	Congenital progressive	35	3
20	62	F	Freedom	CP810	Yes	76.7	13	62	Progressive as adult	35	4
21	62	F	Freedom	CP810	Yes	90	13	56	Progressive as adult	15	4

Note: PTA: Pure tone average at 0.5, 1, and 2 Hz, with "no response" indicating a PTA　90 dB HL; SES: Socioeconomic status SES scores (between 1 and 64): product of scores on two eight-point scales; a score of "8" represents the highest occupational status and the highest educational level. Income level for family: (1) less than $20,000 per year; (2) between $20,000 and $30,000; (3) between $30,000 and $45,000; or (4) greater than $45,000