

Auditory Perception and the Electrode-Neuron Interface in Children and Adults  
with Cochlear Implants

Mishaela DiNino

A dissertation

submitted in partial fulfillment of the  
requirements for the degree of

Doctor of Philosophy

University of Washington

2018

Reading Committee:

Julie Arenberg, Chair

David Perkel

Matthew Winn

Program Authorized to Offer Degree:

Neuroscience

© Copyright 2018

Mishaela DiNino

University of Washington

**Abstract**

Auditory Perception and the Electrode-Neuron Interface in Children and Adults with Cochlear Implants

Mishaela DiNino

Chair of the Supervisory Committee:  
Julie G. Arenberg  
Speech and Hearing Sciences

Cochlear implants (CIs) restore auditory perception to individuals with severe to profound hearing loss, yet identification of speech sounds remains limited due to degraded or distorted representations of frequency information. Spectral resolution, the ability to resolve the frequency components of an auditory signal, is essential for accurate speech identification but is limited in CI users of all ages. Spectral resolution of individuals with CIs can be further impaired by suboptimal interfaces between CI electrodes and auditory neurons in the cochlea. These interfaces could result from neuronal degeneration that occurred during the period of deafness prior to implantation, CI electrodes located distant to auditory neurons, or bone and tissue growth in the cochlea after CI surgery. The aim of this dissertation was to investigate the relation between the electrode-neuron interface and auditory perception in children and adults with CIs. These experiments respectively examined 1) the patterns of phoneme identification errors resulting from

specific spectral degradations that mimicked the effects of poor electrode-neuron interfaces, 2) the developmental time course of spectral resolution and vowel identification in pediatric CI users, 3) auditory perception related to the electrode-neuron interface in two groups with divergent hearing histories: early-implanted children and late-implanted adults, and 4) speech-based spectral resolution in children and adults with normal hearing and with CIs. The results from these experiments provide an improved understanding of the effects of suboptimal CI electrode-neuron interfaces on spectral resolution and phoneme identification performance of CI users. Findings from these studies also indicated that cochlear implantation in children enhances auditory system development and drives maturation of auditory perception. Finally, these experiments demonstrated that prelingually-deafened, early-implanted children differ from postlingually-deafened, later implanted adults in both cochlear physiology and integration of acoustic cues for speech sounds. The knowledge gained from this dissertation laid the foundation for the creation of experimental CI program processing strategies to enhance spectral transmission for children and adults with CIs.

# TABLE OF CONTENTS

List of Figures .....	viii
List of Tables .....	x
Chapter 1 : Introduction .....	1
Chapter 2 : Vowel and Vovsonant Confusions from Spectrally Manipulated Stimuli Designed to Simulate Poor Cochlear Implant Electrode-Neuron Interfaces .....	5
Abstract .....	5
Introduction .....	6
Methods .....	12
Subjects .....	12
Stimuli .....	12
Vocoder Processing .....	12
Procedure .....	18
Analysis .....	18
Results .....	21
Vowel Identification Performance .....	21
Phoneme Confusions .....	23
Discussion .....	31
Conclusion .....	37
Chapter 3 : Age-Related Performance on Vowel Identification and the Spectral-temporally Modulated Ripple Test in Children with Normal Hearing and with Cochlear Implants .....	39

Abstract.....	39
Introduction.....	40
Methods .....	43
Participants.....	43
Assessments .....	45
Demographic Information.....	49
Statistical Analyses .....	49
Sequential Information Analysis (SINFA) .....	51
Results.....	52
Children with Normal Hearing .....	52
Children with Cochlear Implants .....	55
Comparison of SINFA Results between Children with NH and with CIs.....	62
Discussion.....	62
Age-Related Factors.....	62
Relation between Spectral Discrimination and Vowel Identification Performance .....	66
SINFA Results .....	69
Comparisons Between CIs of Bilaterally-Implanted Children .....	71
Conclusion .....	74
 Chapter 4 : Hearing History Affects the Electrode-Neuron Interface in Cochlear Implant	
Listeners: A Comparison between Early-implanted Children and Late-implanted Adults .....	76
Abstract.....	76
Introduction.....	77
Methods .....	79

Participants.....	79
Assessments .....	81
Results.....	85
Discussion.....	90
Conclusion .....	94
Chapter 5 : Speech-based Spectral Resolution by Age and Hearing Modality .....	95
Abstract.....	95
Introduction.....	97
Methods .....	101
Results.....	109
Discussion.....	115
Conclusion .....	118
Chapter 6 : Conclusion.....	120
Bibliography .....	124

## LIST OF FIGURES

Figure 2-1. Vowel space of stimuli.....	14
Figure 2-2. Vocoder manipulations. ....	15
Figure 2-3. Spectrograms of vocoded vowels for the zero manipulation. ....	17
Figure 2-4. Vowel and consonant identification performance.....	23
Figure 2-5. Average confusion matrices for the all channels control condition. ....	24
Figure 2-6. Confusion matrices for the zero vocoder manipulation for each frequency region. .....	25
Figure 2-7. Results of SINFA for male and female talker vowels and consonants. ....	26
Figure 2-8. Perceptual distance results. ....	27
Figure 2-9. Acoustic vs. perceived vowel space.....	29
Figure 2-10. Comparison between NH and CI listener vowel confusions. ....	30
Figure 3-1. Vocoded vowel identification performance of NH children as a function of age. .....	53
Figure 3-2. Relationships between vowel feature information transmission and SMRT thresholds in NH children.....	54
Figure 3-3. Vowel identification performance of children with CIs as a function of age-related variables. ....	56
Figure 3-4. SMRT thresholds of children with NH and with CIs as a function of age-related variables. ....	57
Figure 3-5. Vowel identification performance of children with NH and with CIs compared to SMRT thresholds. ....	58
Figure 3-6. Comparison of task performance between first- and second-implanted ears of bilateral CI users. ....	59
Figure 3-7. Relationships between vowel feature information transmission and SMRT thresholds in children with CIs.....	61
Figure 4-1. Auditory detection thresholds for early-implanted children and late-implanted adults across the electrode array.....	86

Figure 4-2. Average channel-to-channel variability for early-implanted children and late-implanted adults. ....	87
Figure 4-3. Dynamic range for early-implanted children and late-implanted adults across the electrode array. ....	88
Figure 4-4. Most Comfortable Levels for early-implanted children and late-implanted adults across the electrode array. ....	89
Figure 4-5. Average intracochlear resistance values for early-implanted children and late-implanted adults. ....	90
Figure 5-1. Formant transition manipulation for /ba/-/da/ stimuli, from Winn and Litovsky (2015). ....	104
Figure 5-2. Spectral tilt manipulation for /ba/-/da/ stimuli with formant structure held constant, from Winn and Litovsky (2015). ....	105
Figure 5-3. Psychometric functions of children and adults with NH. ....	110
Figure 5-4. Psychometric functions of children and adults with CIs. ....	111
Figure 5-5. Relation between perceptual weighting of formant and spectral tilt cues for children and adults with CIs. ....	112
Figure 5-6. Relation between SMRT thresholds and phonemic cue coefficients. ....	114
Figure 5-7. Relation between vowel identification scores and phonemic cue coefficients. ....	115

## LIST OF TABLES

Table 2-1. The $p$ -values resulting from comparisons between each vocoder condition and the all channels control condition. ....	22
Table 2-2. Perceptual distance between CI subjects S43 and S47 and NH listeners' vowel identification confusion matrices for each vocoder condition. ....	31
Table 3-1. CI user participant demographics. ....	44
Table 3-2. CI user Sequential Information Analysis (SINFA) results. ....	60
Table 4-1. Demographic information for all child and adult participants. ....	80
Table 4-2. Means and standard deviations (SD) of values from each assessment. ....	85
Table 5-1. Demographics of participants with cochlear implants. ....	102

## **ACKNOWLEDGEMENTS**

I would first like to acknowledge my Supervisory Committee members, who provided me with mentorship and guidance throughout my time in the Ph.D. program: Julie Arenberg, Ione Fine, KC Lee, David Perkel, Ric Robinson, Matthew Winn, and Richard Wright. I am especially appreciative of my Ph.D. advisor, Julie Arenberg, who provided me with the training and resources necessary to conduct experiments in research field that was completely new to me.

I wish to express my gratitude to the many individuals who provided valuable insight or assistance with study design, data collection, analysis, or interpretation of results for these experiments: David Horn, Leonid Litvak, Torsten Dau, Johannes Zaar, Matthew Winn, Richard Wright, Elle O'Brien, Kelly Jahn, Lindsay DeVries, Connor McCabe, and Steven Bierer. I would also like to thank the individuals who participated in this research, as well as Susan Norton and Wendy Parkinson, who were instrumental in participant recruitment. Lastly, I am incredibly grateful for the support of my husband, parents, labmates, and friends. Thank you for your continued encouragement.

This research was supported by NIH T32DC005361 (DiNino, PI: Perkel), NIH NIDCD R03 DC014309 (Winn), and NIH NIDCD DC012142 (Arenberg).

## **DEDICATION**

This dissertation is dedicated to my husband, Rye Johnson. Thank you for your complete confidence in me and your steadfast support of my ambitions.

## CHAPTER 1 : INTRODUCTION

Cochlear implants (CIs) are auditory prostheses that restore some degree of auditory perception to individuals with severe to profound sensorineural hearing loss. While CIs have been very successful, speech and language outcomes of both child and adult CI users are highly variable. The overall goal of the experiments in this dissertation was to better understand factors that contribute to speech perception of children and adults with CIs. These experiments focused on two variables in particular: the ability to resolve frequency information in an auditory signal, and the effectiveness of CI electrodes to convey these signals to their target auditory neurons.

A CI consists of both external and internal elements that which process acoustic signals in the environment and transmit auditory information to the central nervous system. One or more external microphones on the CI detect and deliver sound to a speech processor behind the ear. This processor consists of a bank of bandpass analysis filters which split the signal into several frequency bands and converts them to digital signals. These are then sent to a transmitting coil housed inside an external headpiece. The signals are transferred via radiofrequencies through the scalp to an internal receiver/stimulator, where they are decoded, converted to electrical pulse trains, and relayed to an array of 12-22 electrodes implanted inside the scala tympani of the cochlea. Each electrode in the array receives and transmits the frequency component output from one bandpass filter in the processor, exploiting the tonotopic organization of the cochlea: high frequency components of a signal are delivered through electrode channels located in the basal region, whereas lower frequency components are sent through those placed further apically (Loizou 1999; Zeng et al., 2008). The electrodes stimulate adjacent spiral ganglion neurons with current pulses, causing net changes in cell membrane voltage. Depolarization of spiral ganglion neurons result in generation of action potentials which carry the auditory signal to the ascending neural pathway (Ranck 1975).

Although CIs can provide listeners with general access to the frequency components of a signal, CI users have fewer electrode channels with which to transmit frequency information than the thousands of receptor inner hair cells in a healthy human cochlea. As a result, spectral resolution, the ability to discriminate between spectral features, is limited in CI users compared to NH individuals (Litvak, et al., 2007). Spectral resolution is critical for identification of complex

auditory signals such as speech (e.g., ter Keurs, Festen, and Plomp, 1992) and consequently speech recognition is generally poorer in CI users than in normal hearing (NH) listeners (Henry, Turner, and Behrens, 2005).

Despite limitations in the spectral resolution of all CI users, some exhibit better speech and language outcomes than others. Investigations of speech perception in CI listeners have found considerable variation among speech identification scores. For example, Dorman et al. (1989) and more recently, Gifford, Shallop, and Peterson (2008) tested adults with CIs on a monosyllabic word test and observed scores ranging from 0-60% and 0-95% correct, respectively. Similar variability in identification scores has been found on tests of sentence recognition in both quiet and noise in the aforementioned studies and others (e.g., Koch et al., 2004 and Firszt et al., 2004). Many demographic factors have been associated with CI users' speech recognition abilities (see Holden et al., 2013 for review), but individual variation in spectral transmission likely also contributes to the observed variability in speech perception scores.

Suboptimal interfaces between electrodes and spiral ganglion neurons can decrease the quality of spectral information transmitted to the auditory nerve (Bierer, 2007). A poor electrode-neuron interface can occur when electrodes are positioned faraway from target neurons, when those neurons have degenerated, and/or when ossification and tissue growth occur around the electrode array. Spectral information transmitted through channels with poor electrode-neuron interfaces thus either does not reach the auditory nerve or is smeared due to interaction of current between channels. Yet, the effect of suboptimal electrode-neuron interfaces on perception of speech sounds was previously unknown. Experiment 1 of this dissertation, "Vowel and consonant confusions from spectrally manipulated stimuli designed to simulate poor cochlear implant electrode-neuron interfaces," was a vocoder stimulation in NH listeners. Vowel and consonant stimuli were manipulated in distinct ways to simulate auditory transmission through poor CI electrode-neuron interfaces in normal hearing (NH) adults. This investigation examined the patterns of phoneme identification errors that occurred when specific spectral components were degraded, and compared these to confusion patterns of CI listeners for whom the locations of suboptimal electrode-neuron interfaces had been predicted.

Studies of the relationship between spectral resolution, the electrode-neuron interface, and speech perception performance are particularly important for children with CIs. A child's ability to perceive speech sounds in his or her environment is critical for development of speech and

language skills. Children with any degree of hearing loss, from mild to profound, can have poorer performance on tests of language proficiency, academic skills, and adaptability than their NH peers. Cochlear implantation at an early age can result in development of age-appropriate speech comprehension for a proportion of implanted children (e.g., Geers et al., 2003). Still, many school age children who received CIs as young as 18 months perform poorly on tests of speech recognition. Thus, speech perception performance among pediatric CI users is still highly variable and difficult to predict. Investigation of factors that contribute to speech and language abilities of pediatric CI users could improve outcomes for these children.

The maturation of spectral resolution among children with CIs likely varies from child to child and could contribute to the overall variance observed in speech perception scores of this group. The ability to detect and discriminate sounds improves with age in NH children, corresponding to anatomical development of the central auditory system (Moore & Linthicum 2007). Children's capacities to resolve spectral information from a degraded auditory signal are poor, but improve with maturation: previous studies have found that young NH children require higher signal-to-noise ratios to perceive speech in noise (Fallon et al., 2000) and perform more poorly on tests of spectrally-degraded speech recognition (Eisenberg et al., 2000) than older children and adults. Experiment 2 of this dissertation, "Age-related performance on vowel identification and the Spectral-temporally Modulated Ripple Test in children with normal hearing and with cochlear implants" investigated whether spectral discrimination matures more slowly in children with CIs compared to NH children due to the limited number of channels in the implant. This experiment also addressed the relation of spectral discrimination abilities to pediatric CI users' vowel identification scores, as well as vowel feature perception in children with CIs compared to NH children.

Many studies have examined the electrode-neuron interface of adults with CIs (e.g., Bierer, 2007, Bierer & Faulkner, 2010), but such investigations had not yet been conducted in children. The electrode-neuron interface of early-implanted children likely differs from that of adults who lost their hearing and received an implant later in life due to the discrepant hearing histories of these groups. Comparison of factors related to the electrode-neuron interface between children and adults were thus necessary to determine whether differences in these variables existed between the two populations. This investigation was particularly important because early-implanted children and late-implanted adults receive the same CI speech processing strategies, despite their differing

hearing histories which likely impact auditory signal transmission through the CI. Experiment 3 of this dissertation, “Hearing history affects the electrode-neuron interface in cochlear implant listeners: A comparison of early-implanted children and late-implanted adults,” examined the electrode-neuron interface in these adults and children and provided evidence for discrepancies in auditory perception and physical aspects of the cochlea between these groups.

In addition to hearing history, another contrast between early-implanted children and late-implanted adults with CIs is the modality with which each acquired language. Children who are congenitally deaf or lose their hearing shortly after birth, and then receive an implant early in life, learn language with the CI as their sole auditory input. In contrast, late-implanted adults learned language with acoustic hearing. Late-implanted adults thus have an existing template of acoustic properties of speech sounds, whereas early-implanted children do not. These children may therefore exhibit different patterns of perceptual weighting of phonetic cues compared to late-implanted adults. Experiment 4 of this dissertation, “Speech-based spectral resolution by age and hearing modality,” tested early-implanted children and late-implanted adults on a test of speech-based spectral resolution that involved phonetic cue weighting. The relative weights that these individuals placed on specific cues for categorizing speech sounds were determined. NH children and adults also participated in this experiment so that a comparison of spectral cue usage could be made between populations with different hearing modalities.

The experiments within this dissertation progressively built on each other to provide increased understanding of the relation between the electrode-neuron interface, spectral resolution, and speech identification abilities of children and adults with CIs. The results from these investigations may lead to improvements in CI speech processing strategies, including advancements in spectral transmission for channels with varying quality in their interface to target auditory neurons.

## CHAPTER 2 : VOWEL AND VONSONANT CONFUSIONS FROM SPECTRALLY MANIPULATED STIMULI DESIGNED TO SIMULATE POOR COCHLEAR IMPLANT ELECTRODE-NEURON INTERFACES

Mishaela DiNino, Richard A. Wright, Matthew B. Winn, and Julie A. Bierer

Published in the *Journal of the Acoustical Society of America*, (140) 4404 – 4418 (2016).

### ABSTRACT

Suboptimal interfaces between cochlear implant (CI) electrodes and auditory neurons result in a loss or distortion of spectral information in specific frequency regions, which likely decreases CI users' speech identification performance. This study exploited speech acoustics to model regions of distorted CI frequency transmission to determine the perceptual consequences of suboptimal electrode-neuron interfaces. Normal hearing adults identified naturally spoken vowels and consonants after spectral information was manipulated through a noiseband vocoder: either (1) low-, middle-, or high-frequency regions of information were removed by zeroing the corresponding channel outputs, or (2) the same regions were distorted by splitting filter outputs to neighboring filters. These conditions simulated the detrimental effects of suboptimal CI electrode-neuron interfaces on spectral transmission. Vowel and consonant confusion patterns were analyzed with sequential information transmission, perceptual distance, and perceptual vowel space analyses. Results indicated that both types of spectral manipulation were equally destructive. Loss or distortion of frequency information produced similar effects on phoneme identification performance and confusion patterns. Consonant error patterns were consistently based on place of articulation. Vowel confusions showed that perceptions gravitated away from the degraded frequency region in a predictable manner, indicating that vowels can probe frequency-specific regions of spectral degradations.

## INTRODUCTION

Despite the success of cochlear implants (CIs) in restoring auditory perception, CI users exhibit a wide range of performance on speech perception tasks. Accurate speech recognition depends on adequate spectral resolution, or the ability to resolve frequency components in a speech signal. Differences among CI users' spectral resolving capabilities may thus contribute to the variability in CI user speech perception scores. Poor interfaces between CI electrodes and auditory neurons, which result from degeneration of adjacent neurons or suboptimal placement of the electrode array, could result in loss or distortion of the spectral information transmitted through those particular channels. Importantly, that spectral distortion may be localized to a specific range of frequencies. The present study implemented a vocoder simulation in normal hearing (NH) listeners to determine the influence of specific regions of frequency distortion on vowel and consonant recognition performance and confusions. Spectral degradations mimicked the negative effects of suboptimal electrode-neuron interfaces on transmission of particular frequencies within a CI. These spectral manipulations should decrease overall phoneme identification performance and alter confusion patterns. Vowels will be particularly affected because their identification relies heavily on the resolvability of distinct frequency regions.

The goal of this study was to understand and predict the patterns of phoneme perception errors resulting from degradation of specific frequencies, which may contribute to decreased speech identification scores of CI users with suboptimal electrode-neuron interfaces. The present study utilized a channel vocoder, which is a system of sound coding that transmits auditory signals via their simplified amplitude envelopes across frequency bands (Dudley, 1939); this is an essential part of modern cochlear implant processing (Loizou, 2006). In vocoder processing, an incoming signal is analyzed by a series of bandpass filters, from which the time-varying amplitude envelope is extracted through half-wave rectification and low pass filtering. The resulting signal is applied to a carrier with corresponding frequency bandwidth. Vocoder processing is widely used in experiments investigating spectral degradation of speech signals (e.g., Shannon et al., 1995; Dorman et al., 1997) and many recent studies have used vocoder processing to model conditions of poor spectral resolution specifically resulting from spread of excitation in the cochlea (e.g., Litvak et al., 2007; Bingabr et al., 2008; Winn et al., 2015; Won et al., 2015). However, while most vocoder studies have applied spectral distortion broadly across the spectrum, CI users with

poor electrode-neuron interfaces may have particular regions of frequency degradation. The present study thus degraded specific frequency regions to simulate the effects of region-specific, suboptimal electrode-neuron interfaces on vowel and consonant perception in NH listeners.

In previous vocoder studies and others that have manipulated speech stimuli, identification performance decreases dramatically when the frequency components of speech sounds, particularly of vowels, are distorted (ter Keurs et al., 1992). Spectral and spectrotemporal characteristics are critical to differentiating many classes of speech sounds (ter Keurs et al., 1992). Accordingly, some common measures of spectral resolution (e.g., the ability to discriminate spectral ripple signals with distinct phases) are related to speech recognition scores of NH and hearing-impaired listeners (e.g., Henry et al., 2005; Litvak et al., 2007). Frequency resolution is especially difficult for individuals with CIs. Compared to the continuous receptive area on the basilar membrane creating excitation among thousands of receptor inner hair cells in a healthy human cochlea, CIs contain only 12 to 22 electrodes with which to transmit frequency information, limiting the spectral resolution for listeners with these devices. Consequently, speech recognition in quiet and noise is generally poorer in CI users than in NH listeners (cf. Friesen et al., 2001).

For CI users, spectral-resolving capabilities are limited not only because of the small number of frequency processing channels, but also because of the interaction of current between channels (White et al., 1984). However, some perform better than others on tests of recognition of words (Dorman et al., 1989; Gifford et al., 2008) and sentences (Koch et al., 2004; Firszt et al., 2004) in both quiet and noise. This occurs even within groups of CI users with the same implant type and signal processing strategy. While many factors such as duration of deafness, age at implantation, and cognitive abilities have been associated with variability in CI users' speech recognition performance (see Holden et al., 2013, for review), variation in quality of spectral transmission among the channels in an individual's implant is also a likely contributor to the observed across-listener variance in speech perception scores. Each electrode in the array transmits a specific frequency band from the input signal, mimicking the tonotopic organization of a healthy cochlea. Because of the importance of spectral cues for vowel and consonant distinctions, loss or distortion of information from even one channel can potentially decrease speech recognition scores. Neighboring channels whose spread of current stimulates overlapping neural populations also distorts the perceived frequency spectrum. The degree of channel interaction within the

implant has been found to account for a large amount of variance in CI user speech identification performance (Stickney et al., 2006).

Within a single implanted array, certain electrodes may stimulate the auditory nerve less effectively than others due to suboptimal interfaces between the electrodes and the auditory neurons, resulting in degraded transfer of spectral information. A poor interface can occur as a result of electrodes placed relatively more distant from auditory neurons (Finley et al., 2008; Holden et al., 2013) or degeneration or death of those neurons (Miura et al., 2002). Spectral information transmitted through channels with such poor interfaces either does not reach the auditory nerve, or, if the current for that electrode is increased to elicit an auditory percept, is likely distorted due to channel interaction (for review, see Bierer, 2010).

To characterize CI channels with suboptimal interfaces, Bierer (2007) utilized focused electrical stimulation to target specific neural populations, and found that higher levels of focused electrical current were required to reach auditory perception thresholds in some channels of a user's implant than for other channels. Because voltage decreases with distance from the electrode (Jolly et al., 1996), high thresholds suggest some electrodes are distant from auditory neurons either because the electrodes are positioned near the lateral wall of the cochlea or the neurons near them had degenerated (Bierer, 2010; Long et al., 2014). Subsequently, Bierer and Faulkner (2010) observed that channels with elevated focused thresholds had broad psychophysical tuning curves, indicating poor frequency selectivity and providing further evidence of the connection between elevated focused thresholds and poor electrode-neuron interfaces. However, even after identification of such channels, it is not clear how poorly situated electrodes lead to specific errors in speech perception. Each CI electrode channel transmits a particular part of the frequency spectrum of a sound to the auditory nerve, and thus suboptimal electrode-neuron interfaces should result in a loss or distortion of predictable frequency components. This study utilized a vocoder simulation of poor electrode-neuron interfaces with NH listeners to examine the perceptual consequences of such degraded frequency regions on vowel and consonant identification.

Of the numerous studies that have used vocoder processing, only a limited set have investigated situations that reflect specifically poor regions of stimulation. Shannon et al. (2002) and Kasturi et al. (2002) simulated regions of CI user spiral ganglion neuron loss in NH listeners by setting the output of particular bandpass filters to zero to produce spectral "holes" in speech signals. The size and region of the spectral holes were varied. In addition, Shannon et al. (2002)

reallocated frequency information to other channels in some vocoder conditions, and also turned off or reallocated stimulation of specific electrodes in CI listeners' implants to correspond to NH listening conditions. In both studies, and for CI and NH listeners, recognition of speech sounds was found to decrease with increasing size of the spectral hole. Shannon et al. found that recognition of vowels, consonants, and sentences did not differ between conditions of dropping or reallocation, suggesting that reallocation of the dropped frequency information did not improve speech intelligibility, and may have instead distorted the frequency spectrum by creating a frequency-to-place mismatch. The perception of speech sounds requires accurate representation of spectral peak spacing and their relative intensities, and frequency reallocation likely distorted that information within the speech signal. Kasturi et al. (2002) also performed an analysis of the perceptual weight, or relative importance for phoneme identification, of each manipulated channel. They found equal perceptual weighting across all channels for consonants but varying weights across channels for vowels, indicating that certain frequency regions are more important for accurate vowel recognition than others.

Similarly, Throckmorton and Collins (2002) created low-, mid-, and high-frequency spectral holes as three of several vocoder manipulations with NH listeners and observed that the low-frequency loss was more detrimental to vowel and consonant identification scores than were the mid-frequency and high-frequency losses, respectively. Further, the low-frequency manipulation decreased vowel identification performance much more than consonant identification. Results from these vocoder studies demonstrate that a loss or distortion of spectral information through electrode channels can greatly impact overall speech perception scores. Additionally, certain channels may affect recognition of vowel sounds to a greater extent than other channels. Importantly, it is not yet clear whether frequency-specific spectral distortions result simply in increased errors, or in predictable errors. This study will test the hypothesis that phoneme error patterns are predictable based on the region of distortion. If this hypothesis is supported, it might be possible to use these error patterns in conjunction with other measures to learn about the contributions of putatively poor electrode-neuron interfaces to vowel and consonant perception.

The current set of experiments expands on previous work by examining speech confusion patterns to determine how degradation or loss of spectral information through particular channels affects the perception of specific speech sounds. Vowels were chosen as a focus because well-characterized spectral peaks called formants contrast these basic units of speech. The acoustics of

vowels can thus be easily exploited to better understand the auditory system. Vowels, then, are the most appropriate stimuli to determine how specific frequency distortions influence phoneme perception. This study will also test the hypotheses that vowel confusion patterns can corroborate the location of CI suboptimal electrode-neuron interfaces that are predicted by an individual's focused threshold patterns. This knowledge can lead to better understanding of why CI user speech errors occur, what kind of errors occur, and the perceptual consequences of spectral distortion.

This study refers to vowel contrasts in terms of vowel quality, so as to describe the spectral vowel contrasts of North American English (as opposed to vowel duration contrasts in languages such as Japanese; for example, Hirata and Tsukada, 2009). While duration can aid in English vowel identification (Ainsworth, 1972), its contribution is negligible when spectral cues are available. For example, Hillenbrand et al. (1995) performed a discriminant analysis to classify vowel tokens: they found that accurate classification of vowels depended highly on the inclusion of steady-state formant measurements, and that the addition of vowel duration measures only slightly improved classification. Accurate vowel recognition depends strongly on one's ability to resolve formants, particularly the first, second, and third formants (abbreviated as F1, F2, and F3) and to perceive the relationships between these formants within a vowel. Formants are observed as steady-state resonant frequencies (e.g., Peterson & Barney, 1952; Hillenbrand et al., 1995), and time-varying changes in formants within a vowel signal (Nearey & Assmann, 1986; Hillenbrand & Nearey, 1999; Assmann & Katz, 2005), both of which are relatively more difficult to perceive by listeners with CIs, leading to relatively greater use of durational cues by this population (Winn et al., 2012). While prior studies have found that degrading or shifting frequency information decreases vowel identification scores (e.g., Fu & Shannon, 1999), vowel perception confusion patterns resulting from frequency-specific spectral manipulation had yet to be examined.

This study also included consonant stimuli for comparison to vowels because consonant recognition is more robust to spectral distortion (ter Keurs et al., 1992; Xu et al., 2005) on account of their accompanying distinctive cues in the temporal domain. Consonants contrast among place of articulation (where they are produced in the oral cavity), manner of articulation (how the airflow is constricted in the vocal tract), and voicing (ostensibly whether or not the vocal cords vibrate during production). Perception of the place feature depends primarily on spectral information, whereas manner and voicing contrasts can be identified primarily by temporal cues. Thus, both spectral and temporal cues can be used for accurate consonant identification. Because of the

importance of spectral cues for perception of place of articulation, prior studies have found that consonants were most likely to be confused with others of the same manner and voicing but different place of articulation under conditions of frequency distortion (Dorman et al., 1997) and in noise (Miller & Nicely, 1955).

While the very basic tasks of vowel and consonant perception are not the same as tests of more global speech perception abilities, spectral manipulation of phonemes can be well controlled, which is of particular interest to this study. Vowel and consonants can be degraded in more specific ways than more complex speech sounds (e.g., entire words or sentences), allowing for a better understanding of speech processing and categorization.

The present study used confusion matrices to examine the patterns of vowel and consonant identification errors resulting from particular manipulations of the frequency spectrum. These matrices indicate the number of times each phoneme was correctly identified and which phoneme(s) it was confused with when incorrectly identified. CI users' patterns of vowel and consonant confusions indicate the distinguishing features of these speech sounds that are not adequately transmitted through the implant (Remus et al., 2007). For example, Sagi et al. (2010) utilized a model of vowel perception to predict individual CI user's vowel confusion patterns based on the quality of transmission of steady-state formant cues through the implant. Therefore, to use phoneme error patterns to identify suboptimal electrode channels, the relationship between frequency-channel allocation and phoneme acoustics must be explored. In this study, systematic examination of vowel and consonant error patterns resulting from loss or distortion of spectral information was conducted using three methods: sequential information analysis (SINFA; Wang & Bilger, 1973), perceptual distance analysis (e.g., Shepard, 1972), and perceptual vowel space analysis. These analyses progressively provide a more detailed understanding of the effects of spectral distortion. Together, these methods exploit the well-established history of linguistic feature analysis, psychoacoustic exploration, and visualization. Comparison of CI users' phoneme confusions to those made by NH listeners in this experiment support the hypothesis that subject-specific phoneme perception errors can help to interpret the perceptual consequences of suboptimal CI electrode-neuron interfaces.

## METHODS

### *Subjects*

Twelve adult NH listeners (6 male) between the ages of 21 and 29 years (mean age = 25.2 years) were recruited from the University of Washington campus and surrounding community to participate in this study. Subjects were native speakers of American English and underwent a screening to verify hearing at 20 dB hearing level across frequencies from 250 to 8000 Hz. All subjects gave written informed consent and were compensated for their participation. Experimental procedures were approved by the University of Washington Human Subjects Division.

### *Stimuli*

Ten vowels in /hVd/ context (/i/, “heed”; /ɪ/, “hid”; /eɪ/, “hayed”; /ɛ/, “head”; /æ/, “had”; /ɑ/, “hod”; /u/, “who’d”; /ʊ/, “hood”; /oʊ/, “hoed”; /ʌ/, “hud”) were presented under various conditions. Vowel stimuli were recorded from one male and one female talker from the Pacific Northwest to match the region that the subjects were recruited from, as regional dialect has been found to influence recognition of vowel sounds (Wright & Souza, 2012). A head-mounted close talking microphone was used to record vowel sounds in a double-walled sound-treated booth. Recordings were digitized at 44 100 Hz using a 16-bit quantization rate and were resampled to 22,050 Hz. Original stimuli were filtered from 60 to 10,000 Hz using a Hanning filter with a slope of 100 Hz to eliminate proximity effects. This study also used sixteen consonants in /aCa/ context (/p/, “aPa”; /t/, “aTa”; /k/, “aKa”; /b/, “aBa”; /d/, “aDa”; /g/, “aGa”; /f/, “aFa”; /h/, “aTHa”; /s/, “aSa”; /ʃ/, “aSHa”; /v/, “aVa”; /z/, “aZa”; /dʒ/, “aJa”; /m/, “aMa”; /n/, “aNa”; /l/, “aLa”) naturally spoken by a male talker (stimulus materials were the same as those used by Shannon et al., 1995, which were created by Tyler, Preece, and Lowder at the University of Iowa Department of Otolaryngology, 1989).

### *Vocoder Processing*

The processing of speech stimuli was designed to simulate that of CI Fidelity 120 processing with the same frequency band allocations as those used in most Advanced Bionics

devices (Advanced Bionics Corp., Valencia, CA). Speech stimuli were digitally sampled at 17,400 Hz and divided into 15 contiguous pseudo-logarithmically spaced frequency bands from 250 to 8700 Hz. The square root of the total energy in each channel was calculated to compute the envelope and the resulting signal was low-pass filtered at 68 Hz. The envelope from each channel was used to modulate a noise band with a center frequency equal to that of the corresponding channel. Filter output slopes were set at 30 dB/octave. This filter slope was selected to bring the performance of NH listeners to the range of better-performing CI listeners (Litvak et al., 2007). The modulated noise bands were summed and presented to subjects through speakers in a sound-attenuating booth.

The channels chosen for vocoder manipulation were based on analysis of the resonance frequencies of the vowel stimuli performed in Praat (Boersma & Weenink, 2013). The first, second, and third formants at the 25%, 50%, and 75% points of the duration of each vowel were identified from the estimated spectral peaks in each signal. This information was used to select low, middle, and high frequency channels (corresponding to apical, middle, and basal cochlear locations), so that the vocoder manipulations would affect vowel identification in predictable ways. Figure 2.1 shows the acoustic vowel spaces of the male and female talker vowels. Shaded areas indicate the low-, mid-, and high-frequency regions that were manipulated. Because the vowel space is different for female and male talkers and they are differentially affected by the manipulations, data from female and male talkers were analyzed separately.

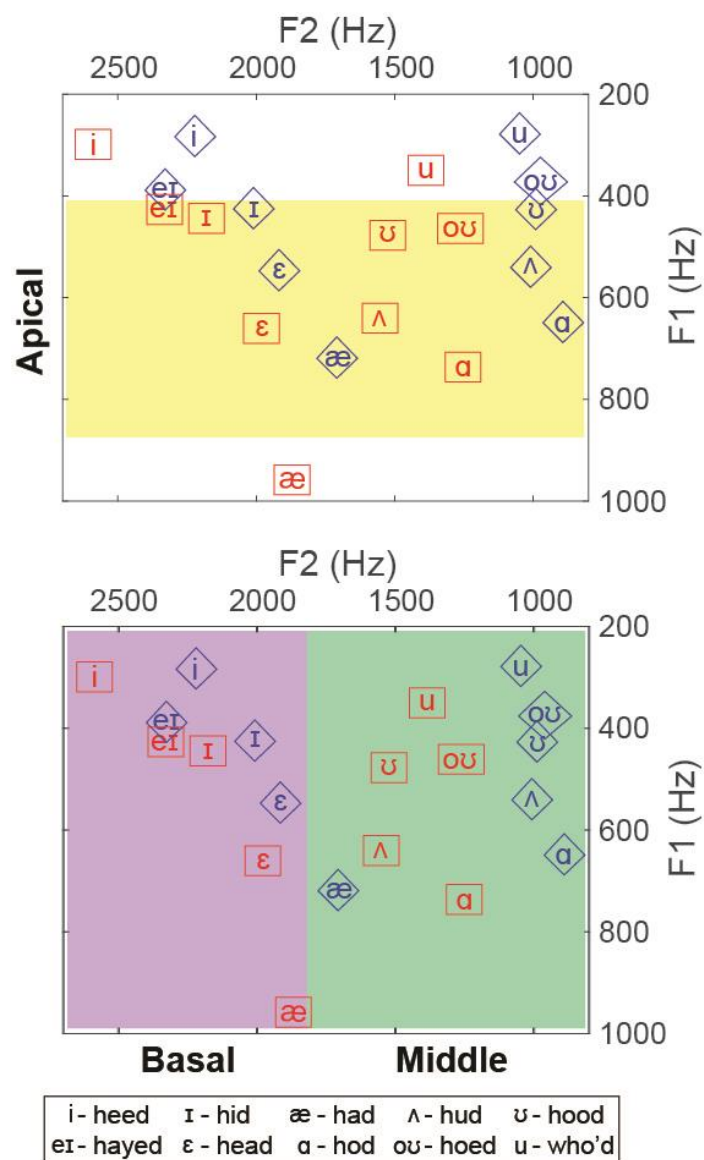


Figure 2-1. Vowel space of stimuli.

First formant frequencies are shown on the ordinate and second formant frequencies are on the abscissa. Formants shown were measured at the midpoint of each vowel. Shaded areas indicate manipulated frequency regions. Male talker vowels are enclosed in diamonds and female talker vowels are in rectangles. Plots show that the formant frequencies for each vowel are different between the male and female speaker and vocoder conditions will therefore affect male and female vowels in diverse ways. Note that the apical frequency region does not cover the first formants of all vowels, and the middle and basal regions each cover the second formants of approximately half the vowels. The basal region also covers the third formant of all vowels (not shown).

Spectral information was degraded via vocoder processing in one of two ways: (1) “zero”—the output of the filters of specified channels was set to zero, simulating complete loss of information through the channel, and (2) “split”—the output of specified filters was set to zero and half of the envelope energy from those filters were added to those of neighboring filters instead, simulating absence of information in the target channel as well as interaction with adjacent channels. Figure 2.2 illustrates these manipulations and channel frequency allocations.

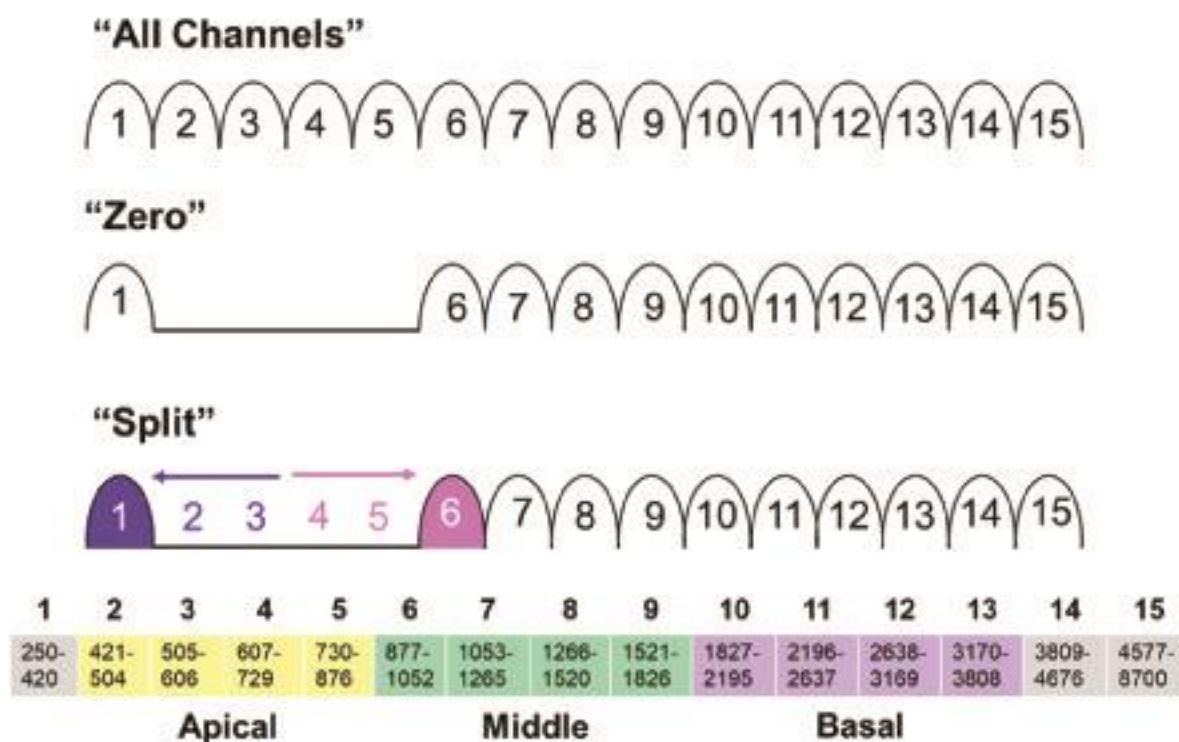


Figure 2-2. Vocoder manipulations.

The all channels control condition consisted of a 15-channel bandpass filter with no manipulations. In the zero condition, the output of particular channels was set to zero. In the split condition, particular channel output was also zero, but the frequency information from channels 2 and 3 (in this example), was sent through channel 1 in addition to the frequency components normally transmitted through channel 1. Similarly, the information from channels 4 and 5 is sent through channel 6, in addition to the normal frequency information transmitted through channel 6. Vocoder channel frequency allocations, corresponding to those of Advanced Bionics CIs, are shown at the bottom.

The locations of the spectral manipulations were shifted through channels corresponding to apical (channels 2, 3, 4, and 5), middle (channels 6, 7, 8, and 9), and basal (channels 10, 11, 12, and 13) cochlear locations to vary the frequencies missing or altered from the speech spectrum. In addition, an “all channels” condition, in which stimuli underwent vocoder processing but were not further manipulated, served as a control condition. While the vocoder manipulations in the present study are similar in principle to those used by Shannon et al. (2002), this study used filter frequencies corresponding to those used by the clinical speech processors developed by Advanced Bionics Corp. (Valencia, CA). Thus, the simulation of CI listening in the present study was more comparable to the clinical experience of Advanced Bionics CI users.

Figure 2.3 includes the spectrograms for female talker vowel stimuli “hid” (left column) and “hood” (right column) for the all channels condition (top), and each frequency region that was removed for the zero vocoder manipulation. “Hid” and “hood” have similar F1 but distinct F2 frequencies, evident in the all channels spectrogram. Frequencies corresponding to the F2 of “hood” were removed in the middle frequency region manipulation (third panel from the top) and the F2 frequencies of “hid” were removed in the basal frequency region manipulation (second panel from the top). In the absence of F2 cues, listeners relied more on F1 cues, and thus “hid” and “hood” were confused due to similarity of their F1 frequencies.

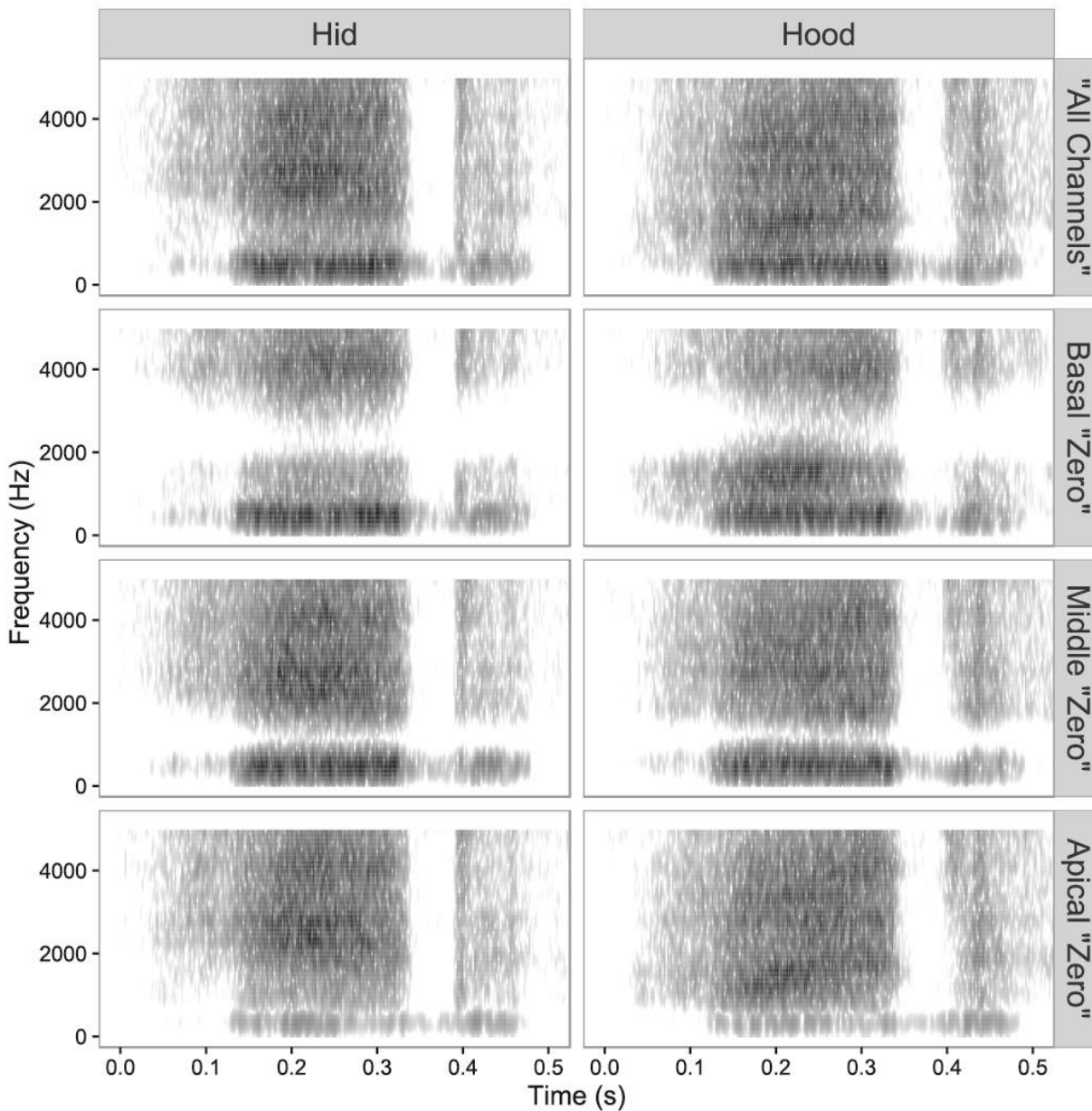


Figure 2-3. Spectrograms of vocoded vowels for the zero manipulation.

Plotted as frequency (ordinate) over time (abscissa), with darker coloration indicating greater energy. Left column: Spectrograms for the vowel sound “hid” for the different regions of frequency removal. Right column: Spectrograms for the vowel sound “hood” for different regions of frequency removal.

### *Procedure*

Testing was performed in a double-walled sound-treated booth (IAC RE-243). Stimuli were played through an external A/D device (SIIF USB SoundWave 7.1) and a Crown D75 amplifier and were presented at 60 dB-A through a Bose 161 speaker placed at 0° azimuth 1m from the subject in the booth. Custom software (ListPlayer2 version 2.2.11.52, Advanced Bionics, Valencia, CA) was used to present the stimuli and record subject responses.

Male talker vowels, female talker vowels, and male talker consonants were presented in separate blocks. For each block, subjects completed two runs with three repetitions of each vowel or consonant with the vocoder manipulations pseudo-randomly interleaved, for a total of six data points for each speech token within each vocoder condition. After presentation of each sound, a list of possible choices was displayed on the computer screen and subjects used a computer mouse to select which one they thought they heard. Test runs were scored as percent correct and the two runs for each list were averaged for each subject. Confusion matrices for each list were also averaged across runs for each subject. Prior to testing a particular list, subjects completed a practice run consisting of one presentation of each vowel or consonant sound in the all channels condition only. Subjects could repeat the sound as many times as desired, and feedback was given after each response to familiarize participants with the vocoded stimuli and the task. Practice data were not included in the average performance scores or confusion matrices.

### *Analysis*

For each vocoder condition, phoneme identification scores were averaged across subjects for each list. Average male and female talker vowel and consonant recognition performance of each vocoder condition were compared to that of the control condition. One-way analysis of variance (ANOVA) with planned multiple comparisons was conducted to determine which conditions of degraded spectral information significantly decreased phoneme identification performance relative to the all channels condition.

In addition to statistical comparison of performance across conditions, the vowel and consonant sounds that were confused as a result of each vocoder condition were examined. Confusion matrix responses were pooled across subjects for each condition in order to conduct these analyses.

## SINFA

The amount of information related to vowel and consonant phonetic features transmitted by each vocoder condition was quantified using SINFA (Wang & Bilger, 1973). This analysis is based on that of Miller and Nicely (1955) and utilizes subjects' phoneme confusions to determine how different conditions of spectral degradation affect perception of phoneme features. However, since the contributions of each feature for phoneme recognition are not independent (e.g., as the nasal feature is recovered, the voicing feature is redundant, since all nasals are voiced), SINFA eliminates this internal redundancy. In the first iteration of the SINFA, the feature with the most perceptual importance for identification is selected and the percent of information transmitted for that feature is calculated. The effect of that feature is then held constant in the second iteration of the analysis, in which the feature with the second highest amount of perceptual information transmitted is identified and calculated, and so on. The analysis concludes when the contributions of all specified features have been elucidated.

To conduct the SINFA, vowels were categorized as having either short (<250ms; lax vowels) or long (>250ms; tense vowels) duration, low (<420Hz), middle (420 to 520 Hz), or high (>520Hz) F1 values (which correspond to high, mid and low vowels, respectively), and low (<1330 Hz), middle (1330 to 2000Hz), or high (>2000 Hz) F2 values, consistent with the approach taken by Xu et al. (2005). Consonants were classified by their manner of articulation (stop, fricative, affricate, nasal, or liquid), place of articulation (bilabial, dental, alveolar, palatal, or velar) and voicing (voiced or unvoiced).

## Perceptual Distance Analysis

This analysis compares confusion matrices resulting from distinct conditions of spectral manipulation to evaluate the difference in phoneme perception between the two conditions (e.g., Zaar & Dau, 2015). Each cell of one confusion matrix is compared to the corresponding cell in the other confusion matrix to determine the overall difference in phoneme perception between the two matrices, scaled from 0% to 100%: a distance of 0% indicates that the matrices compared are exactly the same, and a distance of 100% indicates complete dissimilarity between the matrices.

Baseline perceptual distance for the analyzed matrices is determined by calculating the within-subject perceptual distance. For example, Zaar and Dau (2015) applied this process to consonant identification confusion matrices to determine the effects of different talkers, types of

noise used, and listeners on consonant perception. They calculated the perceptual distance within each subject from test and retest runs and used the resulting values as a baseline for other measurements, since these values represent listener uncertainty. Therefore, this study also calculated a baseline by examining the perceptual distance between each subject's confusion matrices from the first and second runs from each vocoder condition for each phoneme list. Results from perceptual distance calculations were compared to these baseline values, and meaningful results were determined to be those that were significantly higher than the calculated baseline values. In this study, perceptual distance was calculated between listeners for the "all channel" condition to determine the variability among subjects' responses in the control condition. Perceptual distance was also calculated between the male and female talker vowels for the all channels condition to determine the effect of speaker on vowel perception. In addition, this analysis was used to compare the zero and split manipulations for each degraded frequency region within each phoneme list to obtain the differences in phoneme perception due to vocoder manipulation. These calculations were performed within each subject and then averaged across subjects.

#### Perceptual Vowel Space Analysis

The goal of the perceptual vowel space analysis was to determine the direction of the errors resulting from each vocoder manipulation, in terms of physical articulator space and acoustic space. This analysis demonstrates the tendency of specific channel manipulations to warp the perceptual vowel space in a way that can be illustrated with a traditional two-dimensional vowel map. Each vowel was assigned a feature value between 1 and 5 for height (1 = low, 5 = high) and for advancement (1 = back, 5 = front). These numbers reflect the general phonological feature distribution of the vowel space rather than exact formant frequencies, which cannot be easily resolved down to single numbers on account of their dynamically changing state. The difference in feature values between each target and responded vowel was calculated in order to translate confusion matrices into summaries of vowel feature perception and to aid in the illustration of modified perceptual vowel spaces.

#### Comparisons with Cochlear Implant Listeners

The present study was a simulation of suboptimal CI electrode-neuron interfaces and accordingly, the confusion patterns made by NH listeners were compared to those made by CI

users. Two CI listeners in particular were selected because they were judged in a previous study (DeVries et al., 2016) to have poor electrode-neuron interfaces in only the middle frequency region (S43) or the middle and basal frequency regions (S47), somewhat matching those used in this study.

DeVries et al. (2016) measured auditory perception thresholds with focused electrical stimulation as a method to identify the potential locations of suboptimal electrode-neuron interfaces. These thresholds were obtained using a steered quadrupolar electrode configuration, which consists of four intracochlear electrodes: two middle electrodes serve as active electrodes, and two outer electrodes serve as return electrodes. Stimuli consisted of biphasic, charge-balanced pulse trains with a 102 microsecond phase duration and pulse rate of 997.9 presented directly to the CI through the Bionic Ear Data Collection System version 1.18.315 (Advanced Bionics, Valencia, CA). Participants performed an adaptive two-up one-down, two-interval forced choice procedure in which they identified the interval that contained the sound. Each run contained six reversals. For the first two reversals, step size increased by 2 dB for correct responses or decreased 2 dB for incorrect responses. The step size decreased to 0.5 dB for the remaining four reversals. Each reversal converged at 70.7% correct on the psychometric function (Levitt, 1971). Thresholds were determined based on the average of the last four reversals. Four runs were collected and averaged for each electrode. The procedure was repeated until focused auditory perception thresholds were obtained for electrodes 2 to 15.

## RESULTS

### *Vowel Identification Performance*

Significant differences were found in average phoneme identification performance between the seven vocoder conditions for male talker vowels [ $F(6, 77) = 8.64, p < 0.001$ ] and female talker vowels [ $F(6, 77) = 8.03, p < 0.001$ ], but not for consonants [ $F(6, 77) = 2.07, p = 0.067$ ]. For both vowel lists, planned multiple comparisons indicated that all combinations of vocoder manipulation and location resulted in significantly lowered identification compared to the all channels condition at an alpha level of 0.05. However, when a Bonferroni correction was applied to correct for multiple comparisons ( $0.05/18, \alpha=0.002$ ), the vowel identification scores for the female talker Apical zero condition and the male talker Apical split condition were no longer

significantly lower than the all channels condition. The results for all other conditions for vowels remained significant. For consonants, neither the zero nor the split manipulations significantly lowered consonant identification ( $\alpha=0.002$ ) relative to the control condition for any frequency region, although the split manipulation resulted in slightly better identification performance compared to the zero manipulation. Table 2.1 depicts the  $p$ -values for each planned comparison from the ANOVA performed for each phoneme list.

Table 2-1. The  $p$ -values resulting from comparisons between each vocoder condition and the all channels control condition.

Shaded cells indicate a significant result at the  $p < 0.002$  level.

	<b>Apical “Zero”</b>	<b>Middle “Zero”</b>	<b>Basal “Zero”</b>	<b>Apical “Split”</b>	<b>Middle “Split”</b>	<b>Basal “Split”</b>
Female talker vowels	0.006	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Male talker vowels	< 0.001	< 0.001	< 0.001	0.038	< 0.001	< 0.001
Consonants	0.008	0.013	0.005	0.144	0.092	0.191

Figure 2.4 illustrates the median and average phoneme identification performance for each vocoder condition within a list.

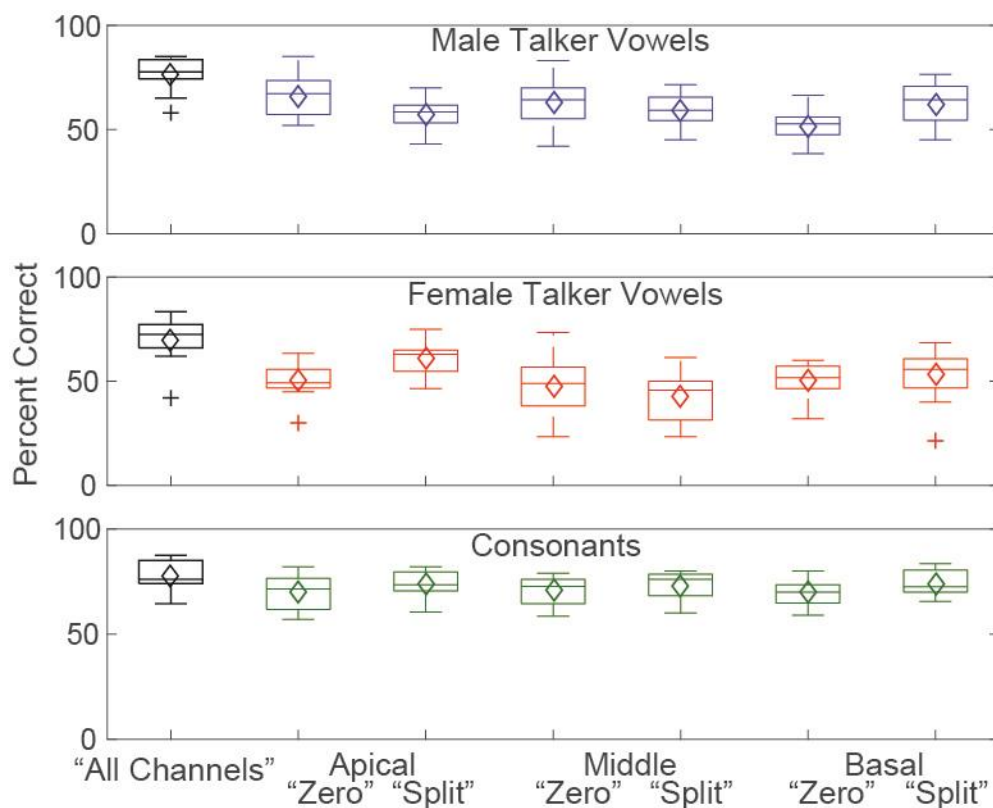


Figure 2-4. Vowel and consonant identification performance.

Box plots depict the average performance across subjects for each condition for each phoneme list. Lower and upper end of the boxes indicate the 25th and 75th percentiles, respectively. Whiskers extend from the third quartile to the highest value that is  $+1.5 \times$  the interquartile range and from the first quartile to the lowest value that is  $-1.5 \times$  the interquartile range. Plus signs indicate outliers. The middle line of each boxplot is the median and the diamond symbols represent the mean.

### *Phoneme Confusions*

Responses from the all channels condition indicated that vowel and consonant confusions were made even when spectral channels were not dropped or reallocated, probably because the spectral resolution of the all channels condition was still limited compared to natural speech

(Figure 2.5). For example, listeners made identification errors in this condition for the male talker vowels in the words “hid,” “had,” “hud,” and “hood,” and for the female talker vowels in the words “who’d,” “hood,” “hoed,” and “hud,” and the consonants in “aGa,” “aKa,” and “aTHa.” However, between-subject perceptual distance for each phoneme in the all channels condition (29.6% for male talker vowels, 32.8% for female talker vowels, and 17.7% for consonants) were comparable to baseline within-subject perceptual distance values for this condition, indicating that overall limited spectral resolution resulted in similar phoneme error patterns for all subjects.

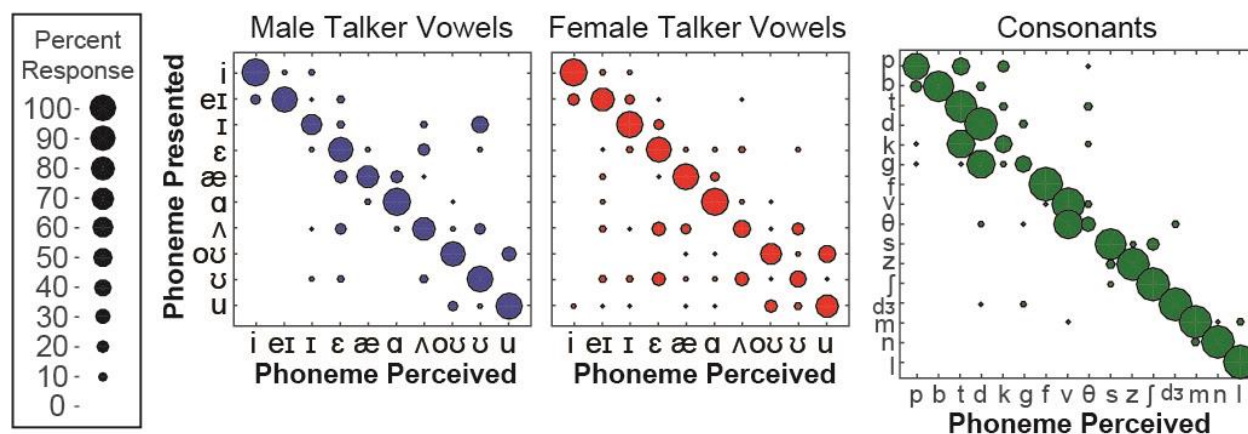


Figure 2-5. Average confusion matrices for the all channels control condition.

Responses are plotted as phoneme presented (ordinate) vs phoneme perceived (abscissa). The size of the circle indicates the percentage of response.

Patterns of vowel confusions from the vocoder conditions reflected the region of degraded information. Figure 2.6 shows the male and female talker vowel and consonant confusion matrices for the zero vocoder manipulation. Confusion patterns were very similar between the zero and split conditions. In these plots, adjacent vowels have similar second formant values and vowels on opposite ends of the matrix have similar first formant values. Manipulation of frequency regions corresponding to first formants (apical regions) resulted in confusions of vowels with similar second formants (i.e., similar vowel advancement). Manipulation of frequency regions corresponding to second formants (middle regions for back vowels and basal regions for front vowels) resulted in confusions between vowels with similar first formants (i.e., similar vowel height). In Figure 2.6, consonants are ordered by manner of articulation. Consonant confusions

occurred between those most similar in manner, exhibited by confusions clustered around the diagonal, and are generally consistent for all conditions of spectral manipulation.

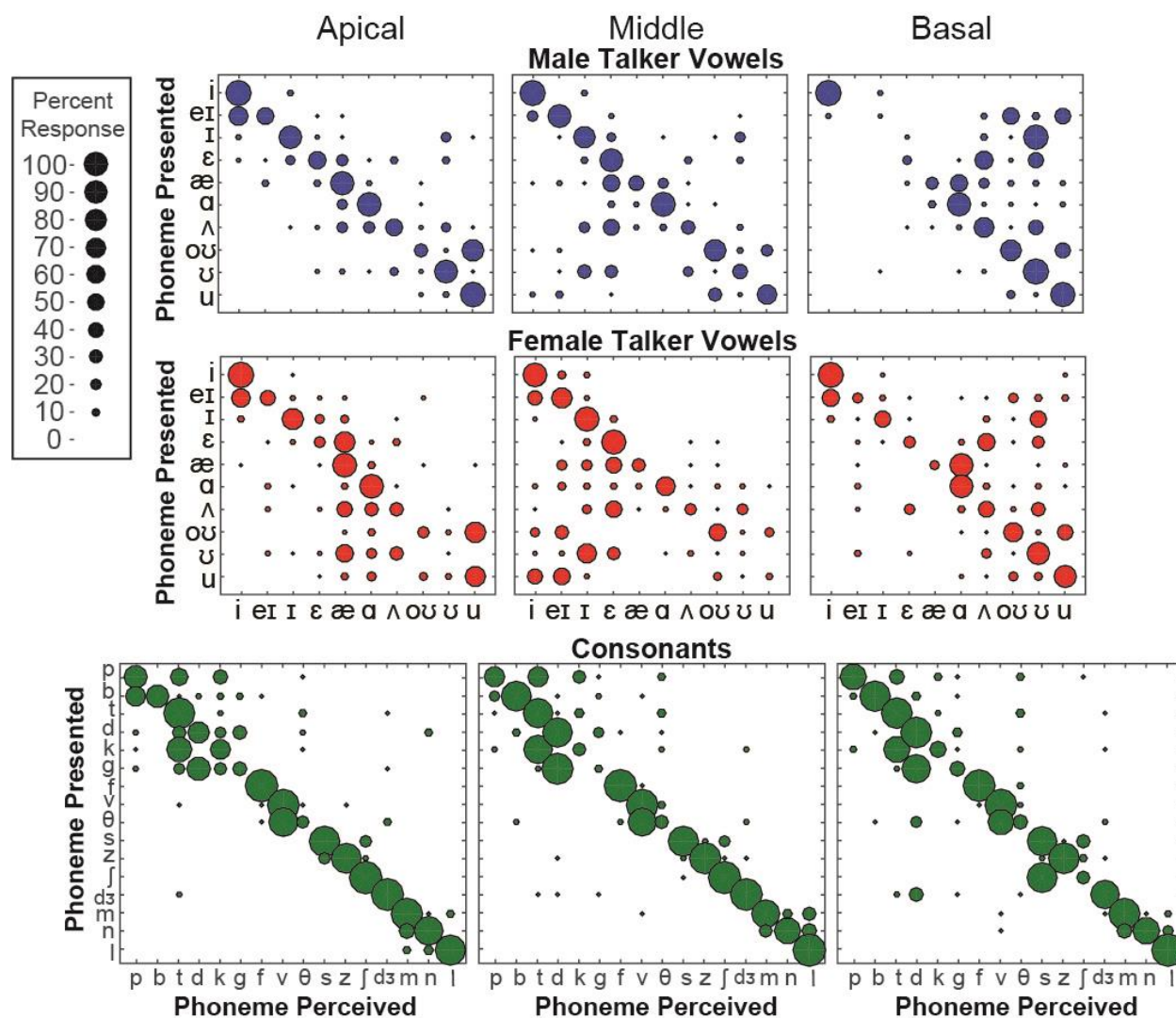


Figure 2-6. Confusion matrices for the zero vocoder manipulation for each frequency region.

Responses are plotted as phoneme presented (ordinate) vs phoneme perceived (abscissa). Percent responded is indicated by circle size, such that larger circles indicate higher percentage of response. Vowels are ordered according to vowel space (see Fig. 2.1), starting with “heed” and moving counterclockwise. Adjacent vowels have similar second formants and vowels opposite each other have similar first formants. Consonants are ordered by manner of articulation (stops, fricatives, affricates, nasals, and liquid).

Figure 2.7 shows the results of the SINFA for each condition for male and female talker vowels and consonants. Percent of information transmitted in the all channels condition for all phoneme lists was greater or not significantly different from information transmitted in other vocoder conditions. For vowels, no clear patterns were found for the amount of information transmitted between vocoder manipulation types or locations, perhaps because of the lack of independence between acoustic attributes of vowels in English. However, SINFA results for consonants revealed a high amount of information transmitted for manner in all conditions, consistent with the relatively intact temporal envelope transmitted through a vocoder, which would yield cues for manner. The split condition transmitted slightly more information related to manner than did the zero conditions, but consonant feature transmission overall was much less affected by particular vocoder manipulations than vowels were. These results are consistent with identification performance results and in agreement with the acoustic cues available for each type of sound category.

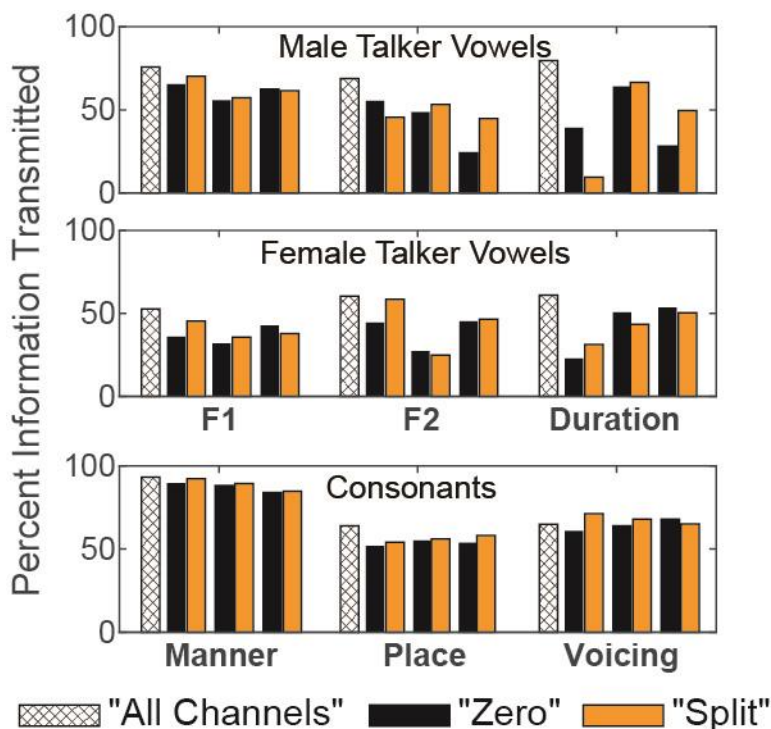


Figure 2-7. Results of SINFA for male and female talker vowels and consonants.

Bars are grouped for each investigated phonetic feature by the apical, middle, and basal regions. The height of each bar depicts the percent of information transmitted for that feature within the specified condition.

The perceptual distance values between male and female talker vowels for the all channels condition depended on the phoneme being examined [see Fig. 2.8(A)]. Responses to the cardinal vowels (/i/, /A/, and /u/) were similar between male and female talkers resulting in perceptual distance values (7.9%, 9.6%, and 19.4%, respectively) that were smaller than the within-subject baseline perceptual distance (24.1%) that was calculated between male and female speakers. Responses to other, more centralized vowels were quite variable between the two speakers resulting in larger perceptual distance values (24.3%-59.9%) and greater magnitude of perceptual distance than the baseline value of 24.1%.

Figure 2.8(B) shows the perceptual distance calculations between zero and split for the apical, middle, and basal frequency regions. These values are either lower than or not significantly higher than the baseline perceptual distance for each frequency region within a phoneme list, indicating nonsignificant differences in perception due to vocoder manipulation type.

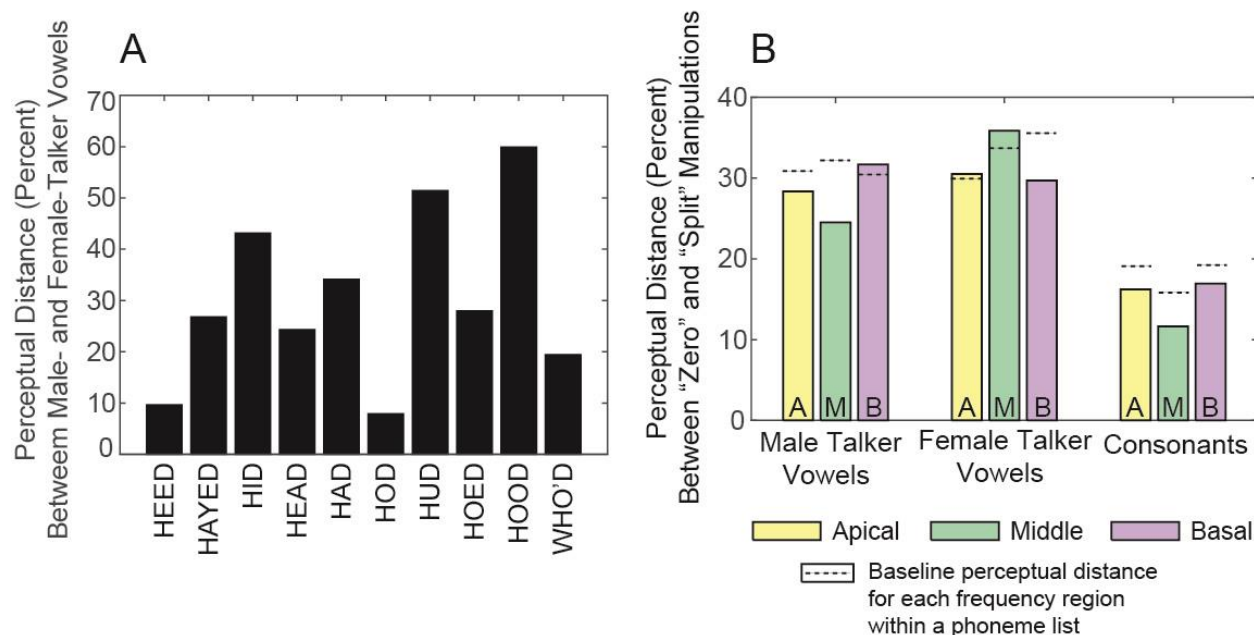


Figure 2-8. Perceptual distance results.

(A) Male and female talker vowels. Results shown are from the all channels control condition and were averaged across subjects. Each bar represents one vowel and the height of the bar indicates the perceptual distance for that vowel between talkers, in percent. (B) Zero and split conditions. Results were averaged across phonemes within a list and subsequently across

subjects. Bars are grouped by phoneme list. Bar height indicates the perceptual distance in percent between the two vocoder manipulation types. The dashed line within or above each bar indicates the baseline within-subject perceptual distance between the zero and split manipulations for that particular frequency location and phoneme list.

The reliable arrangement of vowels in a two-dimensional acoustic space enables visualization of how perceptions can drift from one acoustic region to another. For most vowels, perceptual vowel space was shifted when spectral information was missing or distorted. Figure 2.9 shows the shifts in perceptual vowel space resulting from specific vocoder manipulations. The apical vocoder conditions in which the low-frequency (apical) region, typically containing the first formant, was manipulated resulted in some shifts in perceived vowel height (indicated by arrows pointing upward or downward). Errors in perceived vowel advancement typically indicated that perceptions shifted away from the area of spectral distortion. Similarly, for the vocoder conditions that manipulated the middle frequencies corresponding to back vowels' low second formants, more "front" vowels were perceived (arrows pointing leftward). For manipulations of the higher-frequency basal regions, front vowels with high second formants were perceived as being more "back" (indicated by arrows pointing rightward). It is notable that these errors are not symmetrical: vowel pairs were not confusable; perception of specific vowels shifted toward other specific vowels.

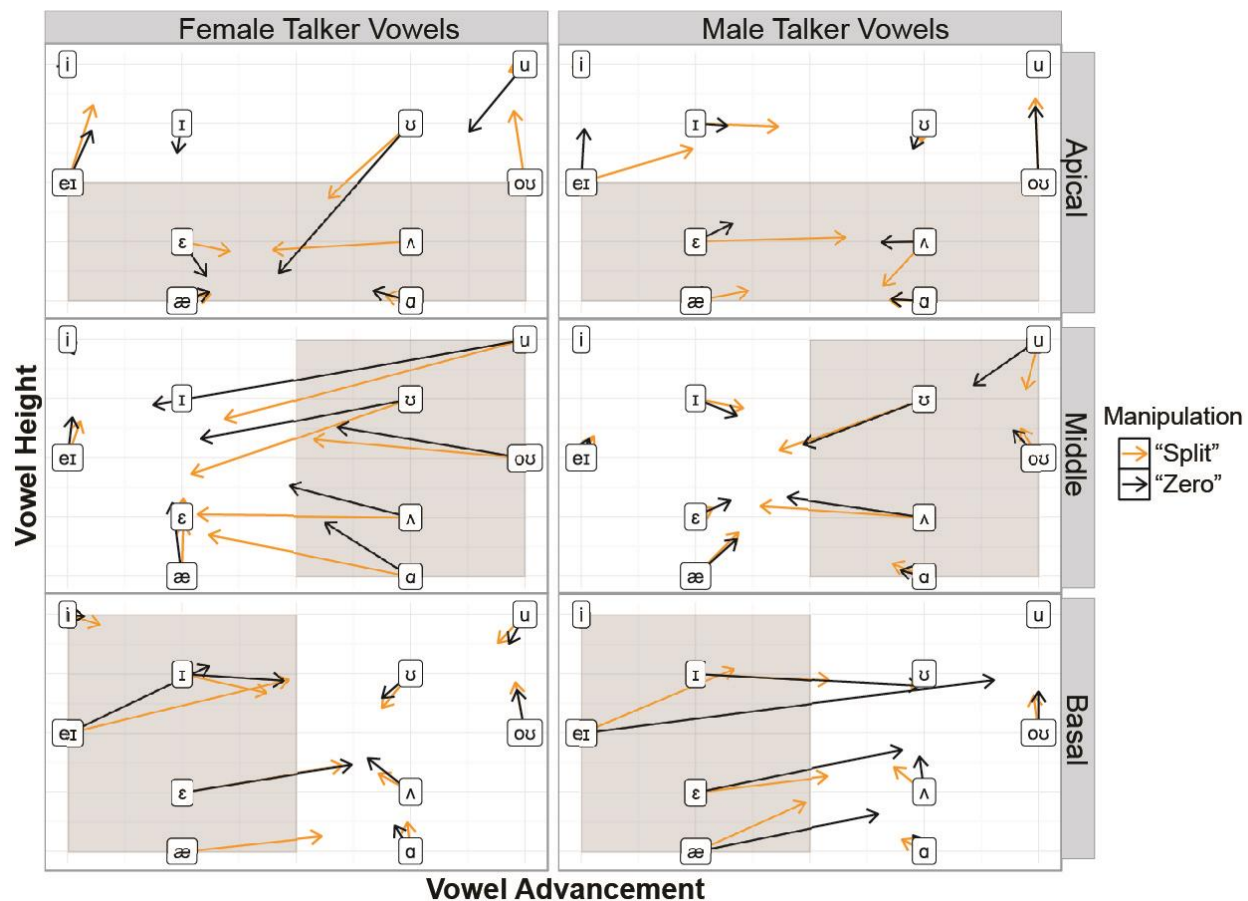


Figure 2-9. Acoustic vs. perceived vowel space.

Placement of vowel symbols correspond to their height and advancement in acoustic vowel space. Gray shaded areas indicate the region of vowel space that was manipulated for the apical, middle, and basal vocoder locations. Arrows depict the shift in vowel perception due to vocoder manipulations in particular frequency regions.

Comparison of consonant and vowel confusions between NH listeners in the vocoder simulation and the CI users tested revealed similar confusion patterns between the groups. For NH and CI listeners, /p/ was confused with /t/ and /k/, /g/ was confused with /d/, /n/ was confused with /m/, and /k/ was confused with /t/ and /p/. These confusions were unidirectional, analogous to patterns of vowel confusions. These are also classic consonant confusions resulting from other kinds of signal degradation (Miller & Nicely, 1955). Perceptual distance analysis results indicated that vowel confusion patterns made by each CI listener were somewhat analogous to the confusions made by NH listeners in the vocoder condition(s) that best matched the CI users' region

of elevated focused thresholds (which likely indicate locations of suboptimal electrode-neuron interfaces). Figure 2.10(A) shows the focused threshold profiles of the two example CI users. These listeners exhibit elevated focused thresholds in the channels corresponding to the middle (S43) and middle and basal (S47) frequency regions used in the vocoder experiments. Figure 2.10(B) shows that S43 made vowel confusions comparable to those made by NH listeners in the middle zero vocoder condition. S47 had two regions of elevated thresholds that were manipulated separately in the present study, and Fig. 2.10(B) shows that this subject also made confusions somewhat similar to NH listeners in both the middle and basal zero vocoder conditions.

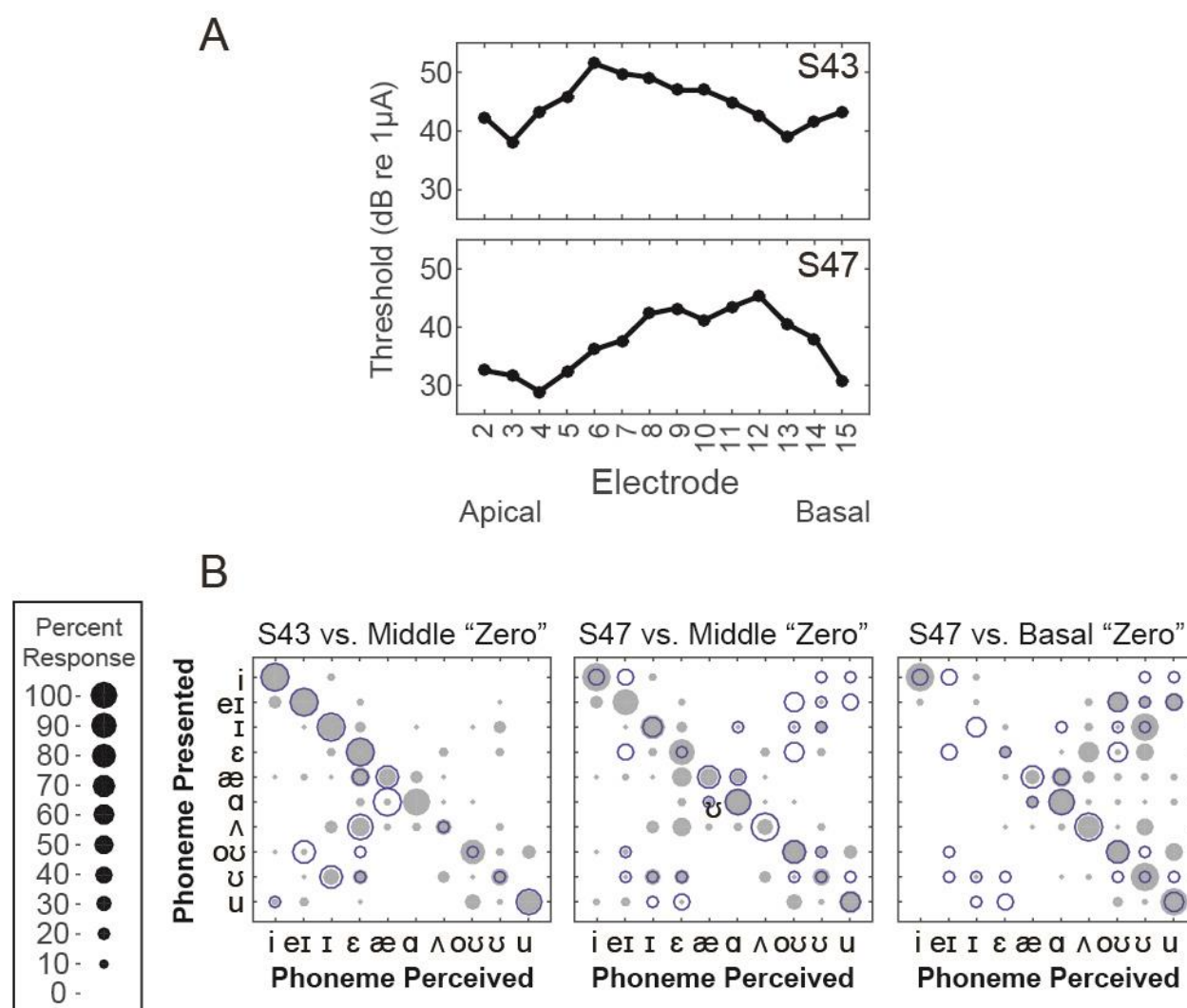


Figure 2-10. Comparison between NH and CI listener vowel confusions.

(A) Focused threshold profiles for two CI subjects, S43 and S47. S43 exhibits elevated thresholds, indicating suboptimal electrode-neuron interfaces, in the channels corresponding to

the middle vocoder frequency region. S47 has elevated focused thresholds in channels corresponding to both the middle and basal vocoder frequency regions. (B) Vowel confusion matrices from CI subjects S43 and S47 (colored circles) overlaid on average NH listener data (gray filled circles).

Table 2.2 contains the perceptual distance values that compare these CI listeners' male and female talker vowel identification confusion matrices to those of NH listeners from all vocoder conditions. The lowest perceptual distance values, indicating the most similar confusion patterns, were observed between S43's confusion matrices and the middle zero and split vocoder conditions. The perceptual distances between S47's confusion matrix and middle and basal vocoder conditions are greater than they might have been if the present study had used a larger span of manipulated channels to match this listener.

Table 2-2. Perceptual distance between CI subjects S43 and S47 and NH listeners' vowel identification confusion matrices for each vocoder condition.

The smallest perceptual distance values, which signify the most similar confusion patterns, are between S43's confusion matrices and the vocoder conditions that manipulated the middle frequency regions.

SUBJECT #	versus:	Apical		Middle		Basal	
		"Zero"	"Split"	"Zero"	"Split"	"Zero"	"Split"
S43	Male talker	49.83	58.13	36.77	34.11	73.14	57.92
	Female talker	56.01	49.09	39.83	36.44	66.02	62.52
S47	Male talker	52.55	55.12	51.91	55.86	48.32	49.24
	Female talker	48.20	37.10	45.58	54.52	46.16	40.93

## DISCUSSION

In this study, NH listeners participated in a simulation of CI listening in which vowels and consonants were spectrally degraded to mimic the effects of localized suboptimal electrode-neuron

interfaces on phoneme perception. Systematic analyses were performed on the resulting phoneme identification scores and confusion patterns.

Results of phoneme identification performance from the present study indicated that both the zero and split manipulations significantly lowered vowel identification relative to the all channels condition. Only slight differences, however, were found between the effects of the zero and split manipulations on vowel recognition performance and confusion patterns in the present study. These findings indicate that distortion of the frequency spectrum is as detrimental to vowel identification as complete loss of particular frequency ranges. This pattern did not hold true for consonant identification, which was not significantly reduced in the zero or split manipulations compared to the all channels condition. These results suggest that consonant perception is robust to frequency distortion.

A comparison of the current results with those of Shannon et al. (2002) shows consistency in the finding that the zero and split conditions deteriorate performance for vowels. However, Shannon et al. (2002) found that both manipulations significantly decreased consonant recognition scores, while the current study showed that consonant intelligibility was not significantly affected by the zero or split conditions. This discrepancy may be the result of differences in stimuli, and/or in frequency band allocation. Shannon et al. (2002) based their vocoder processing on the Cochlear Corporation SPEAK CI processor and used 20 filter bands with analysis filters from 150 to 10,823 Hz. In contrast, the present study modeled vocoder processing after Advanced Bionics Fidelity 120 CI processing and used both a smaller number of filter bands (15) and a smaller range of frequency analysis filters (250 to 8700 Hz) and a different frequency allocation table. Accordingly, the spectral “holes” used by Shannon et al. (2002) ranged from 1.1mm (two channels in the basal location) to 10.7mm (eight channels in the apical location) of cochlear space, while those used in the present study were 4.3, 4.8, and 5.0mm for the four channels whose output was manipulated in the apical, middle, and basal frequency regions, respectively (Greenwood, 1990). The different spatial manipulations of frequencies utilized by Shannon et al. (2002) may have affected consonant recognition to a greater extent than those used in the present study.

Performance for both vowels and consonants was below ceiling even for the all channels condition, demonstrating the difficulty of phoneme perception when spectral resolution is degraded. In a previous vocoder experiment, Litvak et al. (2007) tested NH listeners on vowel and consonant identification with 5, 10, 20, or 40 dB/octave output filter slopes (with higher numbers

corresponding to progressively better resolution) and found that shallower filter slopes, indicated by lower dB/octave values, resulted in decreased phoneme identification scores compared to narrower filter slopes. The present study used the same type of vocoder processing and a 30 dB/octave output filter slope, which in pilot testing was the slope that resulted in comparable phoneme identification performance between NH listeners participating in this experiment and the better performing CI listeners who identified natural stimuli in a previous study (DeVries et al., 2016). Results from the present study are consistent with findings from Litvak et al., in that NH listeners' average vowel and consonant identification scores from the all channels condition with the 30 dB/octave filter slope fall between the range of performance that Litvak et al. observed with their 20 dB/octave and 40 dB/octave filter slopes.

In this study, the hypothesis was tested that manipulations of particular formant frequencies would lead to predictable vowel identification errors, i.e., those vowels whose formants are removed or distorted for each condition will have the most errors in that condition. Examination of vowel confusion patterns indicated that vowel errors were indeed predicted based on the frequency information that was degraded. Furthermore, errors were asymmetrical, tending to gravitate away from the area of manipulation and into the area of relative spectral preservation. In conditions that manipulated the frequencies corresponding to the first formant of a vowel, confusions of vowel height were made, and vowels were confused with other vowels that have similar second formants, indicating the preservation and reliance on second formant information in the absence of first formant cues. While the same pattern was observed in the conditions in which second formant frequency information was degraded, errors on perception of vowel advancement were made to a higher degree than corresponding degradations of vowel height. Again, confusions of vowels according to F2 resulted in perceptions of vowels with similar F1, indicating correct perception of at least one feature rather than complete misperception. The first formant was more robust to degradation than the second formant, perhaps because the Advanced Bionics electrode array, which this vocoder study simulated, has more space devoted to transmitting F1 frequencies (six channels) than F2 frequencies (four channels).

Despite the different vocoder manipulations and locations of manipulated frequency regions, some vowels seemed to be inherently easier to identify than others. For example, for both male and female talker lists, "heed" was identified from 84% to 98% correct for all vocoder conditions. Conversely, "hud" was often misidentified independent of vocoder manipulation or

location. It could be the case that cardinal vowels, having fewer acoustical neighbors, are more robust to degradation, while lax vowels have a higher number of potential confusable pairs.

Results of consonant error pattern analyses indicated that confusions occurred between consonants with the most similar manner of articulation, regardless of the vocoder condition. These results corroborate the idea of the vocoder as preserving temporal envelope structure (a key determiner of manner of articulation) and coincide with results from the vocoder stimulation in NH listeners conducted by Kasturi et al. (2002). They performed perceptual weighting of the vocoder channels to determine the region(s) of frequency information most critical for vowel and consonant identification. While they found that some channels were weighted higher than others for vowel recognition, the weighting function for consonant identification was flat. This indicated that all frequency regions contributed equally to consonant recognition, or that there was greater interdependence of frequency channels for consonants. In the present study, as indicated by consonant confusion patterns, NH listeners were able to use temporal cues reliably to accurately identify consonant manner, even under conditions of spectral degradation. This resulted in patterns of consonant errors that preserved some feature(s) of the target sound, rather than complete misperceptions. These findings support the prediction that vowel stimuli are more useful than consonants in studies that intend to measure effects of frequency-specific degradations. Since adequate spectral resolution is more important for vowel recognition than consonant recognition, use of vowel stimuli can provide valuable information about the perceptual effects of distorted frequency regions.

The SINFA results from the current study demonstrated that the all channels vocoder condition transmitted the largest amount of information compared to the other conditions for both vowels and consonants, validating the spirit of the measurement and corroborating phoneme identification performance. For consonants, SINFA findings were consistent with confusion patterns, indicating that manner of articulation was the feature with the highest percentage of information transmitted. These results correspond to the preservation of temporal aspects of the manner feature during vocoder processing. Place of articulation exhibited the lowest information transmitted, due to the dominance of the spectral information that was degraded. This finding is in line with results from previous studies that tested consonant recognition under conditions of spectral degradation (for example, Dorman et al., 1997; Xu et al., 2005).

While the SINFA yielded clear patterns of results relating to feature perception of spectrally manipulated consonants, this analysis failed to elucidate changes in vowel feature perception due to manipulated frequency regions or type of vocoder manipulation. Results of the SINFA should have coincided with patterns of errors on vowel perception (for example, vowel height errors when apical regions were distorted). This shortcoming is likely because of the relative dependence of features for vowels. A fundamental assumption of SINFA (or any information-theoretic approach to confusion patterns) is that phoneme features are independent bits of information. This independence can be argued for consonants: e.g., alveolar consonants should have prototypical formant trajectories regardless of manner of articulation, and stop sounds should have a particular temporal pattern regardless of place of articulation, although consonants do have unequal numbers of connections between features which violate the information-theoretic assumptions of SINFA. On the other hand, acoustic properties of vowels, such as F1 and F2, are decidedly interdependent (Syrdal & Gopal, 1986). For example, front vowels with lower F1 will also have higher F2, and longer vowels will generally have higher F1, though this is affected considerably by tense/lax status. In short, there is no example in the English vowel system where only one acoustic parameter is changed independently of the others. Therefore, while SINFA may be satisfactory for analysis of consonant identification confusions, other analysis methods, such as the perceptual distance or vowel space analysis, can be more appropriate for examination of vowel perception confusion patterns.

The perceptual vowel space analysis was the most informative for the purposes of this study. These results showed how each condition of spectral degradation warped perception of vowels in articulatory-acoustic space. Shifts in perception of vowel height and advancement were clearly related to the spectral information manipulated in each condition. In the conditions in which frequencies corresponding to vowels' first formants were manipulated (apical region), some vowels shifted in perceived height. In conditions where low values of F2 were manipulated (the "middle" frequency region), perception drifted toward more front vowels, and vice versa for front vowels whose F2 was manipulated with basal frequency degradation.

The findings from these experiments and other vocoder studies (for example, Shannon et al., 2002; Litvak et al., 2007) show similar decreases in speech identification performance between CI users and NH listeners under certain conditions of spectral degradation. Still, a study of how CI users' phoneme confusion patterns relate to those of NH listeners identifying spectrally

degraded stimuli had not yet been conducted. Comparison of confusion patterns from NH listeners in this study and two example CI listeners revealed similar patterns of consonant errors. Even more interesting, however, was that the spectral manipulations in the present study resulted in NH listener vowel confusions that were akin to those made by CI users with poor electrode-neuron interfaces in similar frequency regions as those in the present simulation study. While S43's region of elevated focused thresholds matched one region of frequency manipulation used in the present study (middle), S47's elevated focused thresholds spanned two of the frequency regions manipulated in this experiment (middle and basal). Thus, S43's vowel confusion patterns were more similar to those made by NH listeners in the middle vocoder conditions than were S47's vowel confusions compared to the middle and basal vocoder conditions. If the present study had included a vocoder condition with a larger frequency range that matched the entire region of S47's elevated focused thresholds, it is likely that NH listeners' vowel confusion patterns would be more comparable to those of S47. These findings suggest that simulations of spectral holes are reasonable for mimicking the poor interfaces between CI electrodes and spiral ganglion neurons.

Because of the significance of spectral cues for accurate vowel identification exemplified by this study and others, analysis of the vowel confusions of CI users may allow for the prediction of CI subject-specific causes of poor spectral cue transmission. For instance, Harnsberger et al. (2001) mapped the perceptual vowel spaces of CI listeners resulting from basalward frequency shifts, and while they found that most CI users tested were able to adapt to the frequency shift, they observed differences in perceptual vowel space between NH listeners and individual CI users in terms of acoustic space and compactness of vowel categories. Therefore, future work could involve perceptual vowel space analysis of the vowel identification confusion matrices of CI listeners for whom the locations of poor electrode-neuron interfaces have been predicted. Such analyses could identify the particular spectral cues that are missing due to the effects of suboptimal electrode-neuron interfaces. CI users' speech processor settings or stimulation modes could then be changed in an attempt to restore the missing spectral information.

Previous studies have shown that experimental CI processor programs can improve CI user speech identification performance. Zwolan et al. (1997) found that some CI users exhibited higher scores on sentences, monosyllabic words, and phonemes within monosyllabic words while using experimental maps that contained only the electrodes they could discriminate on a previous task. Similar investigations have demonstrated that CI user speech perception can be improved by

deactivating channels determined to be suboptimal based on particular criteria, and reallocating those frequencies to active electrodes. For example, deactivating channels with a high level of forward masking (Box et al., 2003), poor temporal sensitivity (Garadat et al., 2013), or channels for which a computational model predicted a high degree of interaction with other electrodes (Noble et al., 2013) and reallocating frequencies to active electrodes have been found to result in better CI user performance on tests of consonant identification in quiet (Boëx et al., 2003; Garadat et al., 2013), sentence recognition in noise (Garadat et al., 2013), and performance on the Bamford-Kowal-Bench Speech-In-Noise (BKB-SIN) test (Noble et al., 2013). Similarly, Bierer and Litvak (2016) found that deactivating channels with high focused thresholds, thereby eliminating channels with suboptimal electrode-neuron interfaces, and reallocating frequencies to remaining electrodes increased consonant and vowel identification scores for some individuals. Larger improvements for a greater number of CI users may be observed if channels selected for deactivation were more targeted: such channels would be determined to have poor-electrode neuron interfaces but also be affecting perception of particular vowels, based on perception of stimuli with frequency-specific cues. In this study, we showed that vowels, in account of their arguably well-defined acoustic structure, could provide a good probe for frequency-specific deficits in a way that word or sentence stimuli cannot target as specifically.

## CONCLUSION

This study simulated the spectrally degrading effects of suboptimal CI electrode-neuron interfaces on vowel and consonant recognition of NH listeners. As predicted, vowel identification performance significantly decreased with a loss or distortion of frequency information, but consonant recognition was less affected by the type of distortion, consistent with previous studies. This study utilized relatively novel techniques, perceptual distance and vowel space analyses, to examine phoneme confusion patterns. Results indicated that vowel confusions occurred between those most similar in residual frequency space and consonant confusions occurred between those with the same manner of articulation. Perception of vowels specifically drifted away from areas of frequency distortion, rather than simply causing uncertainty and greater number of errors. Vocoder spectral degradation in NH listeners resulted in patterns of vowel and consonant errors that were somewhat similar to those made by example CI users with matching region(s) of suboptimal electrode-neuron interfaces. This is the first experiment to demonstrate that perception will drift

away from areas of frequency-specific distortion in a predictable manner. These results provide insight into the perceptual consequences of spectral distortion on speech confusions. These findings may be useful for interpreting speech confusions of individual CI users and translating those data into diagnostic markers of the electrode-neuron interface.

## CHAPTER 3 : AGE-RELATED PERFORMANCE ON VOWEL IDENTIFICATION AND THE SPECTRAL-TEMPORALLY MODULATED RIPPLE TEST IN CHILDREN WITH NORMAL HEARING AND WITH COCHLEAR IMPLANTS

Mishaela DiNino and Julie G. Arenberg

Published in *Trends in Hearing*, (22) 1 – 20 (2018).

### ABSTRACT

Children's performance on psychoacoustic tasks improves with age, but inadequate auditory input may delay this maturation. Cochlear implant (CI) users receive a degraded auditory signal with reduced frequency resolution compared to normal, acoustic hearing; thus, immature auditory abilities may contribute to the variation among pediatric CI users' speech recognition scores. This study investigated relationships between age-related variables, spectral resolution, and vowel identification scores in prelingually deafened, early-implanted children with CIs compared to normal hearing (NH) children. All participants performed vowel identification and the Spectral-temporally Modulated Ripple Test (SMRT). Vowel stimuli for NH children were vocoded to simulate the reduced spectral resolution of CI hearing. Age positively predicted NH children's vocoded vowel identification scores, but time with the CI was a stronger predictor of vowel recognition and SMRT performance of children with CIs. For both groups, SMRT thresholds were related to vowel identification performance, analogous to previous findings in adults. Sequential information analysis of vowel feature perception indicated greater transmission of duration-related information compared to formant features in both groups of children. In addition, the amount of F2 information transmitted predicted SMRT thresholds in children with NH and with CIs. Comparisons between the two CIs of bilaterally-implanted children revealed disparate task performance levels and information transmission values within the same child. These findings indicate that adequate auditory experience contributes to auditory perceptual abilities of pediatric CI users. Further, factors related to individual CIs may be more relevant to psychoacoustic task performance than are the overall capabilities of the child.

## INTRODUCTION

Cochlear implants (CIs) are highly successful in restoring auditory perception to individuals with severe to profound hearing loss. Still, some adults and children with CIs perform more poorly than others on tests of speech identification. This is a particularly significant problem for children with CIs, because a child's ability to perceive speech sounds is critical for development of verbal speech and language skills (e.g., Eisenberg et al., 2016; Niparko et al., 2010). Early-implanted children who receive a CI prior to 5 years of age often exhibit the most favorable outcomes in verbal speech and language development; nevertheless, speech perception performance remains highly variable even among these children (Horn et al., 2017; Jung et al., 2012; Tyler et al., 1997; Wang et al., 2008). Delayed or deficient spoken language abilities can affect a child's ability to learn effectively and to develop social skills (Wake et al., 2004). Therefore, identifying and attempting to ameliorate factors that can decrease speech perception abilities is imperative for optimizing verbal language outcomes for children with CIs.

Older children with normal hearing (NH) perform better than younger children on many psychoacoustic tasks, including spectrally-degraded speech identification (e.g. Eisenberg et al., 2000), and the timeline of auditory cortical development corresponds well with these improvements in NH children's speech perception abilities (e.g., Eggermont & Ponton, 2003). Despite copious evidence of the relationship between development and auditory task performance in NH children, chronological age has not been found to predict pediatric CI users' speech recognition scores (e.g., Eisenberg et al., 2002) or performance on other psychoacoustic tests (e.g., Horn et al., 2017; Jung et al., 2012; Landsberger et al., 2017). These null findings may have resulted from the tremendous variability in outcomes of children with CIs. Still, auditory experience has been demonstrated to be critical for both the structural (e.g., Moore & Linthicum, 2007), and functional (Eggermont et al., 1997; Kral et al., 2005) development of the central auditory system; therefore, the present study investigated the role of auditory experience in children with CIs as a mediator of the improvements in psychoacoustic task performance that normally occur with chronological age.

The central auditory system undergoes many physiological changes during the process of normal development, including changes in neuronal and axonal structure and density, which reach adult levels in adolescence (Moore & Guan, 2001). In addition, the morphology of auditory evoked

potentials become adult-like around age 12 in NH children (e.g., Ponton et al., 2000; Sharma et al., 1997), corresponding to the structural maturation of the generators of these potentials. However, the development of central auditory structures depends on auditory experience (for review, see Moore 2002). Auditory brainstem and cortex are stunted with a lack of auditory stimulation in the perinatal period and early childhood (e.g., Moore & Linthicum, 2007). Most prelingually deafened children do not receive a CI until they are at least one year old, and often later, while audition begins as early as 19 weeks gestation in humans (Hepper & Shahidullah, 1994). Auditory system development as discussed above is thus likely impeded in prelingually deafened CI users, even when implanted early. Both the auditory deprivation prior to implantation and the degraded auditory signal (compared to normal, acoustic hearing) that the CI provides could impair development.

In NH children, developmental changes in the central auditory system coincide with the time course of improvements in psychoacoustic abilities (e.g., Eggermont & Ponton, 2003; Moore & Linthicum, 2007) such that older children achieve higher performance levels than younger children on a number of auditory tasks (e.g., Dawes & Bishop, 2008; Elliott, 1979; Hall et al., 2004; Hartley et al., 2000; Maxon & Hochberg 1982; Peter et al., 2014; Talarico, et al., 2006). In particular, spectral resolution, or the ability to resolve the frequencies in a complex auditory signal, has been found to be poor at young ages and is enhanced with normal development (Dorman et al., 2000; Eisenberg et al., 2000; Eisenberg et al., 2002; Vongpaisal et al., 2012). Age-related effects have been observed in studies utilizing a non-linguistic test of spectral resolution called spectral ripple discrimination (SRD), in which an individual discriminates between spectrally-modulated rippled noise stimuli (Supin et al., 1994). For example, Kirby et al. (2015) observed that young NH children's SRD performance improved with age. Several additional investigations have found poorer SRD performance in groups of younger NH children compared to groups of older NH children and adults (Allen & Wightman, 1992; Landsberger et al., 2017; Peter et al., 2014; Rayes et al., 2014). NH children (Allen & Wightman, 1992) as well as adults with NH, hearing loss, and CIs (e.g., Henry et al., 2005; Won et al., 2007) who perform well on SRD tasks also tend to demonstrate high speech identification scores. Presumably, this relationship exists because spectral smearing or distortion will reduce an individual's speech recognition abilities (e.g., Shannon et al., 1995). The relation between spectral resolution and speech recognition scores in children with CIs, however, is not as well-understood, as results have varied depending on the

speech materials used (e.g., Jung et al., 2012). These mixed findings may be at least partly due to variance in the maturity of spectral discrimination abilities among children with CIs.

Behavioral findings support this hypothesis: for example, Kirby et al. (2015) tested NH children and children with hearing impairment on a recently-developed SRD task, the Spectrally Modulated Ripple Test (SMRT; Aronoff & Landsberger, 2013) and found that performance in both groups was better for older children. However, performance began to asymptote around age 9 for NH children but seemed to continue to improve beyond this age in the hearing-impaired group. While the authors did not specifically discuss the age at which SMRT performance peaked in hearing impaired children, a possible interpretation of this observation is that decreased access to auditory information, and/or the degraded auditory signal that the central auditory system receives due to hearing loss, prolongs the maturation of spectral resolving capabilities. Other prior studies with pediatric CI users have demonstrated immature temporal sensitivity (Jung et al., 2012) and intensity resolution (Park et al., 2015) in school-age children and adolescents compared to adults with CIs.

While these findings, as well as those from previous studies in NH children, indicate that pediatric CI users' performance on psychoacoustic tasks should improve with age (but perhaps on a delayed time course), previous studies have found no relationship between chronological age and SRD performance (Horn et al., 2017; Jung et al., 2012) or speech identification scores (Eisenberg et al., 2002) in early-implanted, school-age children with CIs. Evidence from structural examinations of auditory system development with auditory deprivation (e.g. Moore & Linthicum, 2007) suggest that aided hearing age or amount of time with the CI may instead be better predictors of auditory capabilities in these children.

The present study explored the link between chronologic and/or hearing age and auditory perception, as well as the relationship between spectral discrimination and speech identification performance, in school-age children. Both NH children and prelingually deafened, early-implanted children with CIs were tested with the goal of relating auditory perceptual development with acoustic hearing to that with a CI. Further, immature speech identification and spectral discrimination abilities were investigated as a potential contributor to the variance in speech identification outcomes among pediatric CI users. In addition, bilateral, sequentially-implanted children were assessed with each CI individually, providing the ability to compare performance with CIs implanted at different time periods within the same child. This also allowed for the

evaluation of age-related effects while minimizing potential confounding variables such as linguistic knowledge, cognitive abilities, and intelligence quotient (IQ).

## METHODS

### *Participants*

Twelve children with CIs between the ages of 11 and 17 (mean age at first lab visit = 14.2 years) were recruited from the Seattle area to participate in this study. No pediatric CI users in this study were younger than 11 years old due to the limited number of children in the Seattle area who met eligibility criteria for the study. Half of the participants with CIs failed a newborn hearing screening in both ears (P01, P02, P03, P10, P11, and P12). The remaining participants either passed their newborn hearing screening (P04, P08) or did not have one (P05, P06, P07, P09). However, all child CI users in this study were diagnosed with unilateral ( $n = 1$ ) or bilateral ( $n = 11$ ) severe to profound hearing loss prior to 4 years of age and were thus considered to be prelingually deafened. These children were either unilaterally- ( $n = 1$ ) or sequentially bilaterally-implanted ( $n = 11$ ) and received their first implant before their 5th birthday (mean age at first implantation = 2.18, mean age at second implantation = 8.0; see Table 3.1 for demographic details).

All children with CIs had Advanced Bionics HiRes90K devices and used oral communication. Two participants were fraternal twins (P11 and P12). One bilateral CI user had been diagnosed with severe to profound hearing loss at age 4 in their second-implanted ear (P05L), so data from that perilingually deafened ear were not included in this study. The one unilaterally-implanted participant (P08) wore a hearing aid in their contralateral ear that was turned off during testing. Bilaterally-implanted participants completed the study tasks with each CI individually in separate testing sessions. A total of 22 CI ears were included in the analyses for this investigation.

Table 3-1. CI user participant demographics.

EVA: enlarged vestibular aqueduct. DFNB1: genetic nonsyndromic hearing loss. The reported preoperative unaided thresholds are the average or range of pure-tone hearing thresholds at 500, 1000, and 2000 Hertz in decibels relative to hearing level (dB HL). The “>” symbol indicates no response to sound at equipment level limits. N/A: information not available. The CI in the gray shaded column (P05L) was not included in the analysis for this study, as hearing loss in this ear was not diagnosed at a prelingual age.

<b>Subject:</b>	<b>P01</b>	<b>P02</b>	<b>P03</b>	<b>P04</b>	<b>P05</b>	<b>P06</b>	<b>P07</b>	<b>P08</b>	<b>P09</b>	<b>P10</b>	<b>P11</b>	<b>P12</b>
<b>Gender:</b>	M	M	M	F	M	F	F	M	F	M	F	M
<b>Etiology:</b>	Unknown	EVA	Unknown	Unknown	DFNB1	Unknown	Unknown	EVA	Unknown	DFNB1	DFNB1	DFNB1
<b>First-Implanted Ear</b>												
<b>Ear:</b>	R	R	R	R	R	R	R	L	L	L	R	R
<b>Age at Implantation:</b>	2.3	1.1	1.4	1.5	4.1	4.3	1.9	2.9	2.6	1.1	1.4	1.7
<b>Chronological Age:</b>	15.7	11.8	12.9	13.2	17.7	17.2	13.3	15.3	13.5	13.3	13.3	13.3
<b>Aided Hearing Age:</b>	14.7	11.8	12.6	12.1	16.6	15.3	11.8	13.0	13.0	13.1	13.0	13.0
<b>CI Age:</b>	13.4	10.8	11.5	11.7	13.7	12.8	11.4	12.3	10.9	12.2	11.9	11.6
<b>Preoperative unaided thresholds:</b>	>105	102.5	N/A	N/A	108.3	85	>115	90-105	103.3	90->115	108	>108
<b>Second-Implanted Ear</b>												
<b>Ear:</b>	L	L	L	L	L	L	L	N/A	R	R	L	L
<b>Age at Implantation:</b>	12.1	3.1	5.6	4.5	13.9	11.0	4.9		3.9	5.1	10.2	10.2
<b>Chronological Age:</b>	16.1	12.1	13.2	13.2	17.9	17.2	13.7		13.7	13.4	13.4	13.4
<b>Aided Hearing Age:</b>	15.1	12.1	13.0	12.1	16.8	15.3	12.2		13.3	13.1	13.1	13.1
<b>CI Age:</b>	4.0	9.1	7.6	8.7	4.1	6.2	8.8		9.8	8.2	3.2	3.2
<b>Preoperative unaided thresholds:</b>	101.7	80-110	96.6	N/A	100	98.3	>115		106.7	92	108	108

Thirty-seven NH children between the ages of 8 and 17 performed a test battery that was comparable to that performed by the group of children with CIs to obtain a metric of performance levels on the psychophysical tasks resulting from normal auditory system development. Three NH children could not fully complete the tasks and thus the final sample included thirty-four NH children (mean age = 12.97 years; fourteen males). The age distribution of NH child participants was approximately uniform, with at least three children falling within each one-year age group. One NH child (NHP09) was a fraternal twin of the CI participant P03 and another (NHP14) was a fraternal triplet of the CI participant P04. Several other NH participants were non-twin siblings. NH children were recruited from the Seattle community and the University of Washington Speech and Hearing Sciences Communication Studies Participant Pool (NIH P30 DC004661). These participants did not have any prior hearing problems or ear surgeries and completed a screening to verify hearing at 20 dB HL from 250 to 8000 Hz.

Children with NH and with CIs were all native speakers of American English and were born and raised in the Pacific Northwest. Most participants had no impairments in speech, language, vision, motor skills, or cognition. One pediatric CI user, P01, had a diagnosis of mild Asperger Syndrome but was able to perform the tasks. Children gave written informed assent and a parent or guardian gave written informed consent. Families were compensated for their participation. Experimental procedures were approved by the University of Washington Human Subjects Division (IRB #28778).

### *Assessments*

Participants performed testing in a double-walled sound-treated booth (IAC RE-243). Stimuli were played through a Crown D75 amplifier and an external A/D device (SIIF USB SoundWave 7.1) and were presented through a Bose 161 speaker placed at 0° azimuth one meter from the subject. Children with CIs performed the tests using their clinical CI processor with the contralateral hearing aid or CI turned off. Bilaterally-implanted children completed testing in this manner with each CI separately on different dates. NH children performed the tests with both ears simultaneously.

## Vowel Identification

Speech stimuli consisted of ten vowels (/i/, /ɪ/, /eɪ/, /ɛ/, /æ/, /ɑ/, /u/, /ʊ/, /oʊ/, /ʌ/) in /hVd context presented at 60 dB SPL. Vowels were naturally-spoken by a female talker from the Pacific Northwest because regional dialect can influence speech identification performance (Wright and Souza, 2012). Vowels were chosen as the stimuli for this study for several reasons: 1. vowel identification performance in quiet matures at a younger age than for other speech stimuli (Johnson et al., 2000), 2. vowels are simple units of speech and thus minimize the effects of children's linguistic knowledge on recognition scores, and 3. while non-spectral cues can be utilized for vowel identification, one's ability to resolve the formants, or spectral peaks, in the vowel sound is important for accurate vowel recognition (e.g. Boothroyd et al., 1996; DiNino et al., 2016; Shannon et al., 2002). Accordingly, vowels are appropriate speech stimuli for an investigation of spectral resolution and speech identification in children.

Participants used a computer mouse to select the presented vowel from the closed set of possible responses listed on the computer screen. Custom software (ListPlayer version 2.2.11.52, Advanced Bionics, Valencia, CA) was used to present the stimuli and record subject responses. Children with CIs first identified the vowels in quiet. Participants who received a score of 80% or higher in quiet with the CI being tested also performed the task in the presence of Auditech 4-talker babble at a +10 dB signal-to-noise ratio.

NH children identified vowel stimuli that were processed through a 15-channel noiseband vocoder with a 30 dB/octave filter slope to mimic the reduced spectral resolution of listening through a CI. This vocoder processing utilized the same frequency band allocations as the Fidelity F120 or Optima speech processing strategies used in Advanced Bionics implants. These settings were chosen because a previous study demonstrated similarities between adult CI users' unprocessed vowel identification scores and NH adults' performance on vowel identification with these settings. (DiNino et al., 2016). No floor or ceiling effects were observed for the NH children in this experiment.

Both groups of children completed two runs, each consisting of three repetitions of each vowel. Runs were scored in percent correct. If the difference in performance between the two runs was greater than 10%, the child performed a third run. Scores from all test runs were averaged for each subject.

Prior to the test runs, subjects completed a practice run consisting of three presentations of each vowel in which they could repeat the sound as many times as desired. Feedback was also given after each response. The practice runs consisted of the stimuli (vowels in quiet, in noise, or vocoded) that the child was going to be tested on following the practice. Additionally, NH children completed a practice run with one repetition of the unvocoded vowel stimuli in quiet at the beginning of the test. This familiarized the children with the vowel list and ensured that they could accurately identify the vowels prior to vocoder processing. Practice data were not included in the average performance scores.

#### Spectral-temporally Modulated Ripple Test (SMRT)

A potential issue with the conventional spectral ripple tasks used in many previous studies is that subjects may rely on local loudness cues to discriminate between stimuli (Aronoff & Landsberger, 2013). Intensity resolution continues to mature through adolescence in NH children (Horn et al., 2017; Maxon & Hochberg, 1982) and in children with CIs (Park et al. 2015) and therefore cues in the intensity domain may confound results on this test. The SMRT, which uses spectral ripples with drifting modulation phases to diminish potential within-channel loudness cues, may be more appropriate than traditional spectral ripple discrimination tasks for use with pediatric populations.

The SMRT assesses the ability to discriminate between rippled noise stimuli at increasing densities of spectral peaks and valleys. The stimuli consisted of the sum of 202 amplitude-modulated pure tones from 100 to 6400 Hz with a modulation depth of 20 dB and a drifting phase rate of 5 Hz (Aronoff and Landsberger, 2013). Stimuli were presented at 65 dB SPL in a 3-interval forced-choice 1-down/1-up adaptive procedure with ten reversals. Each trial consisted of two reference stimuli and one target stimulus. The ripple density, or density of spectral peaks and valleys, of the reference stimuli were set at 20 ripples per octave (RPO), while the target stimulus was set initially at 0.5 RPO and was altered in step sizes of 0.2 RPO. The starting phase of each target and reference stimulus for each trial was randomly selected between values of 0,  $\pi/2$ ,  $\pi$ , and  $3\pi/2$ . The SMRT becomes more difficult as the RPO value of the target stimulus, and thus the density of the ripples, increases. The threshold for each run was calculated based on the average of the last six reversals and indicated the highest RPO of the target stimulus at which 50%

discrimination was achieved. Hence, larger threshold values indicate better spectral discrimination abilities.

Each participant completed one practice run and two test runs. The practice run was identical to the test run, but these data were not included in the calculation of participants' average SMRT thresholds. If a participant's thresholds from two test runs were greater than one RPO apart, he or she completed a third run. This occurred for testing from one ear of three children with CIs (P06, P09, and P10) and for sixteen of the thirty-four NH children. Results from all test runs were averaged to determine the mean SMRT threshold for each subject.

Despite the consistent finding of a strong association between performance on SRD tasks and speech identification scores in adults with CIs (e.g., Henry et al., 2005; Won et al., 2007), potential issues exist in the interpretation of such results at medium and high ripple density rates. In particular, the CI processor cannot accurately represent spectral ripple stimuli at high RPO values, which results in a non-monotonic relationship between presented and perceived stimuli with increasing ripple density. However, SRD tasks have been a very popular method for assessing spectral discrimination abilities and have been utilized extensively in populations of adults with NH and hearing impairments; thus, goals of the current study were to examine performance on such assessments by children and compare those to results found consistently in adults with NH and with CIs. Precautions were taken in the present study to interpret the results of the SMRT appropriately (see Discussion).

Efforts were made to minimize the influences of decreased attention and practice effects on auditory task performance. All children were given several breaks and active participation was incentivized with snacks and small toys. The order in which a participant performed vowel identification, the SMRT, and one other task not reported here was randomized. Bilaterally-implanted pediatric CI users completed the tasks over two sessions to reduce testing fatigue. These individuals performed vowel identification and the SMRT with one CI during the first visit, and with their other CI during their second visit. However, previous research has found that psychophysical test scores, particularly those on the SMRT, improve between sessions due to learning effects (e.g., de Jong, Briaire, and Frijns, 2017); testing each ear separately over two sessions could thus result in a bias of better performance with the second-implanted ear if that CI was always tested last. Therefore, two of our participants performed the tasks with their second-

implanted ear first. Full randomization of implant testing order was not completed because of a concern that participants with CIs may not be interested in or available to return for a second visit, and data with the earlier-implanted ear were most valuable for this study.

### *Demographic Information*

Birthdates were obtained for each child participant in the study to determine the potential influence of age on vowel identification and SMRT performance. The dates that each child with a CI received a hearing aid in each ear and the date(s) of the first stimulation of their implant(s) were obtained via parent report and were corroborated by examination of medical records from Seattle Children's Hospital (see Table 3.1).

### *Statistical Analyses*

Age at testing was calculated for each participant. Aided hearing age (time between receiving a hearing aid and receiving a CI) and CI age (time between CI activation and the date of testing) were also calculated for participants with CIs. These variables were determined separately for each ear of bilaterally-implanted children because they 1) had been sequentially-implanted and 2) had completed testing with each of their CIs during separate sessions, which were typically months apart.

The vowel identification scores in quiet and in noise for children with CIs and vocoded vowel identification scores for NH children were converted to rationalized arcsine units (RAU) to normalize error variance and transform the data into a more appropriate form for analysis (Studebaker, 1985).

All statistical analyses were performed using IBM SPSS Statistics (IBM Corp., 2010). For data obtained from NH children, a multiple linear regression analysis was conducted to determine whether SMRT thresholds and/or age predicted vocoded vowel identification performance. A second linear regression testing an additional hypothesis was performed to determine whether chronological age, representing maturity of auditory perception, significantly predicted spectrally-degraded vowel identification.

A series of planned mixed-model, repeated-measure linear regressions testing separate hypotheses were conducted with data obtained from pediatric CI participants to identify the relationships between auditory perceptual development, SMRT thresholds, and vowel

identification performance in quiet and noise. For all models, data from each CI of bilaterally-implanted children were analyzed as separate data points while including “subject” as a random intercept and ear tested (first- or second-implanted ear) as the repeated measure to address the lack of independence in the dataset. An unstructured covariance matrix was specified for each model. The first set of analyses examined whether chronological age, aided hearing age, and/or CI age significantly predicted vowel identification in quiet (Model 1) or in noise (Model 2) or SMRT thresholds (Model 3). The next set of analyses determined whether SMRT performance was a significant predictor of vowel identification in quiet (Model 4) and/or in noise (Model 5). All predictor variables were set as fixed factors in each model.

To examine the effects of ear implanted (first- or second-implanted CI) on spectral discrimination and vowel identification of bilaterally-implanted children, a multivariate analysis of variance (MANOVA) was performed with Ear Tested as the predictor variable for the independent variables of vowel identification in quiet, vowel identification in noise, and SMRT performance.

Vowel recognition scores and SMRT thresholds were not directly compared between children with CIs and with NH; vowel stimuli differed between groups, and device limitations of CI users restrict global performance on the SMRT. In addition, a positive relationship between SMRT performance and vowel identification scores would be interpreted differently for each group of children. Identifying vocoded vowel stimuli mimics CI listening, but NH children have an intact peripheral auditory system to process spectrally-degraded speech. Accordingly, NH children would not utilize the same frequency-resolving mechanisms to perform both vocoded vowel identification and spectral discrimination; a positive relationship between performance on these tasks would more likely result from general development of the auditory system. In children with CIs, however, performance on both tests would be affected by the limited spectral resolution of the implant as well as potentially delayed auditory system maturation. The relationship between spectral discrimination abilities and vowel identification performance were examined separately for children with CIs and with NH, while the effects of presumed auditory perceptual development on psychophysical tasks were compared between the two groups of children.

*Sequential Information Analysis (SINFA)*

SINFA was performed on each participant's pattern of vowel confusions to determine the degree of vowel phonetic feature perception of children with CIs and with NH. This feature analysis is based on that of Miller and Nicely (1955) and utilizes perceptual responses to a closed-set of presented phonemes to calculate the percent of information transmitted related to a specific phonetic feature. Yet, the contribution of phonetic features to their identification are not completely independent; SINFA was created to account for at least some of this redundancy. The SINFA performs iterations in which the feature with the highest amount of information transmitted is first identified, its percent information transmitted calculated, and then held constant while the next iteration is performed to determine the next most important feature (Wang & Bilger, 1973). While SINFA does not control for the entirety of vowel feature interdependence, and thus may be better suited for use with stimuli with less internal redundancy (e.g., DiNino et al., 2016), this analysis provides a means of at least approximating the amount of vowel feature information available to each group of children for perceiving these sounds.

Three features integral for accurate vowel discrimination were chosen for the SINFA: the first formant (F1; the lowest-frequency spectral peak), the second formant (F2; the next lowest-frequency spectral peak) and duration (as in Xu et al., 2005; DiNino et al., 2016). Vowels were classified as having low (<420 Hz), middle (420 to 520 Hz), or high (>520 Hz) F1 values, low (<1330 Hz), middle (1330 to 2000 Hz), or high (>2000 Hz) F2 values, and short (<250 ms) or long (>250 ms) duration.

The matrices of vowel responses to vowel stimuli presentations, or confusion matrices, from test runs within a vowel identification condition (vocoded, in quiet, or in noise) were averaged for each NH participant and for each ear tested of the children with CIs. SINFA was utilized to calculate the percent, from 0 to 100%, of information transmitted for F1, F2, and duration from each averaged confusion matrix. Higher percentages signify a larger amount of transmitted information for that feature. Although different vowel stimuli were presented to each group, vocoder processing degrades spectral feature information in a manner which simulates CI listening. Thus, an independent t-test was performed to compare vowel feature information transmission values resulting from vowel identification in quiet of children with CIs and spectrally-degraded vowel identification of NH children. The ranking of F1, F2, and duration feature

transmission (from most to least information transmitted) was also examined between the groups, as well as between the quiet and noise conditions for pediatric CI users' vowel recognition.

In addition, as spectral discrimination capabilities are important for accurate identification of speech sounds, this study sought to determine whether SMRT performance was related to perception of specific vowel features. The relationship between vowel feature information transmission values and SMRT thresholds was examined for each group to determine if one's ability to perceive a certain vowel feature may have contributed to SMRT performance. Multiple linear regression analyses were performed for NH children's data and mixed-models repeated-measures regression analyses were conducted for data from children with CIs, with SMRT threshold as the independent variable and F1, F2, and duration information transmission values as predictors. Diagnostic tests of collinearity revealed that the information transmission values of all vowel features from children with CIs' vowel identification in quiet were highly correlated with each other (Variance Inflation Factor (VIF) values ranging from 9.3 to 16.0). High VIF values were not observed among these variables for vocoded vowel identification of children with NH (VIF values between 1.4 and 2.5), but follow-up correlation analyses revealed strong, statistically significant correlations between information transmission values of all vowel features (R values between 0.51 and 0.75). For these reasons, separate regression models with corrections for multiple comparisons were run to determine the relation between F1, F2, and duration information transmission values on SMRT thresholds. Neither multicollinearity (VIF values: 1.1, 1.2, and 1.1) nor significant correlations were observed among F1, F2, and duration-related information transmission for vowel identification in noise of children with CIs, and therefore all three independent variables were included in the same repeated-measures regression model for this condition.

Means (m) and standard deviations (SD) of all data are reported below.

## RESULTS

### *Children with Normal Hearing*

Vocoded vowel identification performance varied greatly among NH children, ranging from 20.0 to 88.5 (m = 63.1, SD = 17.3) percent correct. SMRT thresholds (with larger numbers

indicating better performance) of NH children varied between 5.2 to 10.7 ( $m = 8.6$ ,  $SD = 1.34$ ) RPO. A multiple linear regression analysis to predict vocoded vowel identification performance based on chronological age and SMRT thresholds resulted in a significant regression equation [ $F(2,31) = 12.4$ ,  $p < 0.001$ ,  $R^2 = 0.44$ ]. Both age ( $p = 0.001$ ) and SMRT thresholds ( $p = 0.03$ ) in this analysis were significant predictors of vocoded vowel identification scores. Figure 3.1 shows the relationship between chronological age and vocoded vowel identification performance of NH children. On average, the regression model revealed that vocoded vowel recognition performance improved 2.9 percentage points for each one-year increase in age, and 4.0 percentage points for each 1 RPO increase in SMRT thresholds.

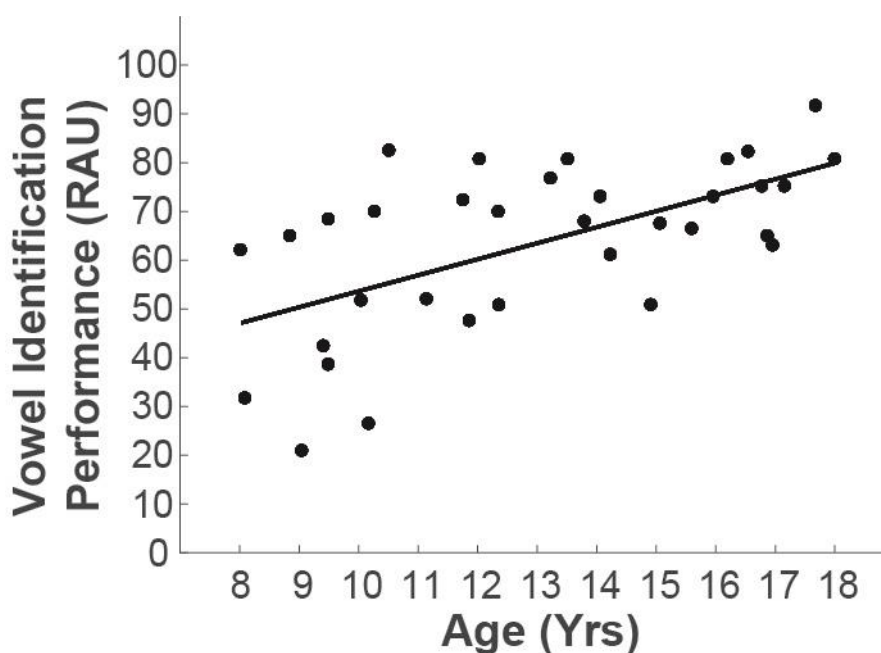


Figure 3-1. Vocoded vowel identification performance of NH children as a function of age. Vowel identification scores are in rationalized arcsine units (RAU). Each circle represents data from one child. Solid line represents the line of best fit.

While age was found to significantly predict vocoded vowel identification, a simple linear regression analysis revealed no significant relationship between age and SMRT thresholds of NH children [ $F(1,32) = 1.78$ ,  $p = 0.19$ ,  $R^2 = 0.05$ ]. However, the analysis revealed that, on average, SMRT thresholds did increase slightly (0.1 RPO) for each year increase in age [see Figure 3.4(A)].

Results of the SINFA indicated a large amount of variation in the percent of vowel feature information transmitted among NH children. These values ranged between 8% - 100% for F1, 17% - 94% for F2, and 1% -100% for duration. On average, duration exhibited the highest amount of information transmitted ( $m = 68.0\%$ ,  $SD = 27.8$ ), followed by F1 ( $m = 61.1\%$ ,  $SD = 24.3$ ) and F2 ( $m = 57.4\%$ ,  $SD = 19.4$ ). The analyses to determine the relationship between vowel feature information transmission and SMRT thresholds revealed a significant relation between F2 transmission on SMRT performance [ $F(1,32) = 9.0$ ,  $p = 0.005$ ,  $B = 3.3$ ,  $R^2 = 0.22$ ]. The relationships between SMRT thresholds and transmission of F1-related information [ $F(1,32) = 4.2$ ,  $p = 0.05$ ,  $B = 1.9$ ,  $R^2 = 0.12$ ] and transmission of duration feature information [ $F(1,32) = 4.4$ ,  $p = 0.045$ ,  $B = 1.7$ ,  $R^2 = 0.12$ ] did not survive correction for multiple comparisons ( $\alpha = 0.016$ ). These analyses also revealed that F2 information transmission had larger effects on SMRT thresholds than did F1 and duration information transmission. These results therefore indicated that NH children's perception of the F2 feature was a predictor of SMRT performance (see Figure 3.2).

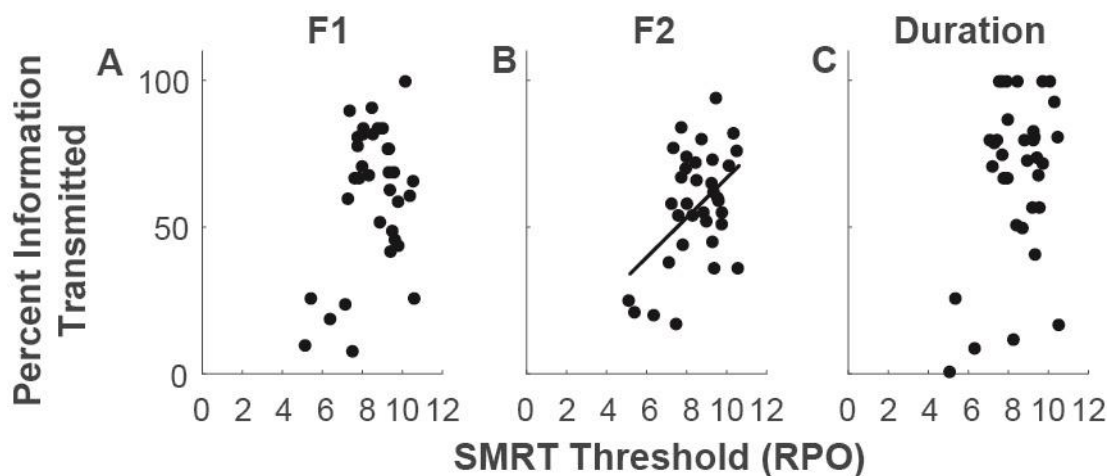


Figure 3-2. Relationships between vowel feature information transmission and SMRT thresholds in NH children.

Percent of information transmitted for (A) F1, (B) F2, and (C), duration plotted against SMRT thresholds. Each circle represents data from one child. Solid line represents the line of best fit for the significant relationship.

*Children with Cochlear Implants*

Analogous to speech testing results observed in previous studies, a large amount of variability was observed in vowel identification performance among children with CIs. This variance was found also between the bilateral CIs of a child. Vowel recognition scores in quiet ranged from 13.5 to 100 ( $m = 79.0$ ,  $SD = 25.5$ ) percent correct. Only children who received 80% correct or higher on vowel identification in quiet performed the test in noise; among those individuals, scores varied between 40 and 86.5 ( $m = 71.7$ ,  $SD = 12.7$ ) percent correct. In addition, as expected, on average pediatric CI user's SMRT thresholds were much lower (indicating poorer performance) than those observed in NH children, ranging between 0.6 to 7.7 ( $m = 2.9$ ,  $SD = 1.9$ ) RPO.

Unlike the observed results in NH children, mixed-model regression tests of fixed effects revealed that chronological age did not significantly predict vowel identification performance in quiet ( $p = 0.46$ ) or in noise ( $p = 0.26$ ), nor did aided hearing age (quiet:  $p = 0.70$ ; noise:  $p = 0.57$ ), likely because most children were fit with amplification soon after birth. Instead, CI age significantly predicted vowel identification performance in quiet [ $F(1, 9.4) = 10.1$ ,  $p = 0.011$ ], suggesting that the timeline of speech perception development may coincide with the amount of time a child has had their CI. Further examination of this relationship revealed that performance with the second-implanted CIs were driving these results: when the data were separated by first- or second-implanted ears, a significant result between vowel identification scores in quiet and CI age was found for second- [ $F(1,6) = 22.2$ ,  $p = 0.003$ ], but not for first-implanted ears [ $F(1,8) = 0.02$ ,  $p = 0.90$ ]. No significant relationship between CI age and vowel identification performance in noise was observed ( $p = 0.27$ ), although only the better-performing children (who most often had their implants for longer periods of time) were tested in noise (see Figure 3.3).

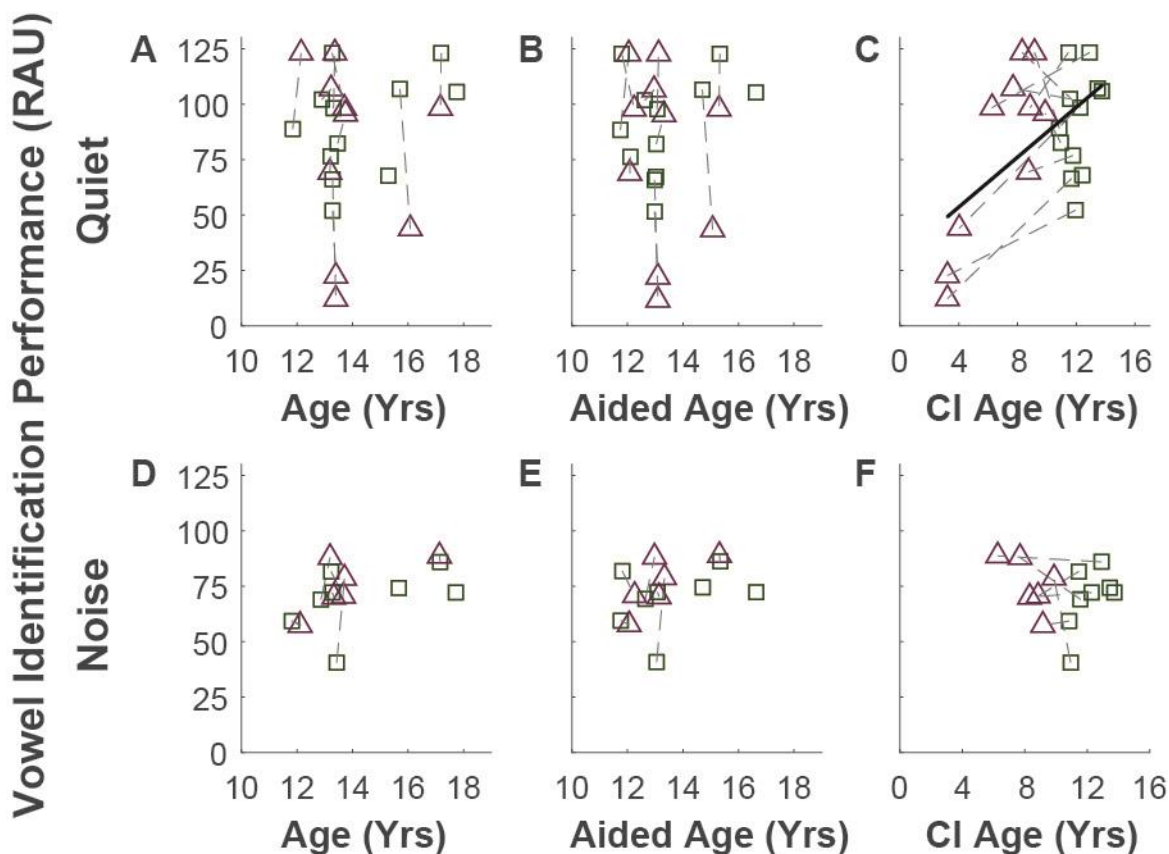


Figure 3-3. Vowel identification performance of children with CIs as a function of age-related variables.

(A) – (C), vowel identification in quiet versus (A) chronological age, (B) aided hearing age, and (C) CI age. (D) – (F), vowel identification in noise versus (D) chronological age, (E) aided hearing age, and (F) CI age. Vowel identification scores are in rationalized arcsine units (RAU).

Each symbol represents data from one CI. Squares indicate first-implanted CIs and triangles indicate second-implanted CIs. Dashed lines connect data from the two CIs of each bilaterally-implanted child. Solid lines represent the lines of best fit for significant relationships.

A mixed-model regression analysis to investigate the relationship between age variables and spectral discrimination abilities revealed that CI age also predicted SMRT performance [ $F(1,9.8) = 9.4, p = 0.012$ ], while chronological age ( $p = 0.32$ ) and aided hearing age ( $p = 0.67$ ) did not (see Figure 3.4). However, unlike the relationship between CI age and vowel identification scores, separate-ear analyses revealed no significant relationship between CI age and SMRT thresholds for either first- or second-implanted ears alone.

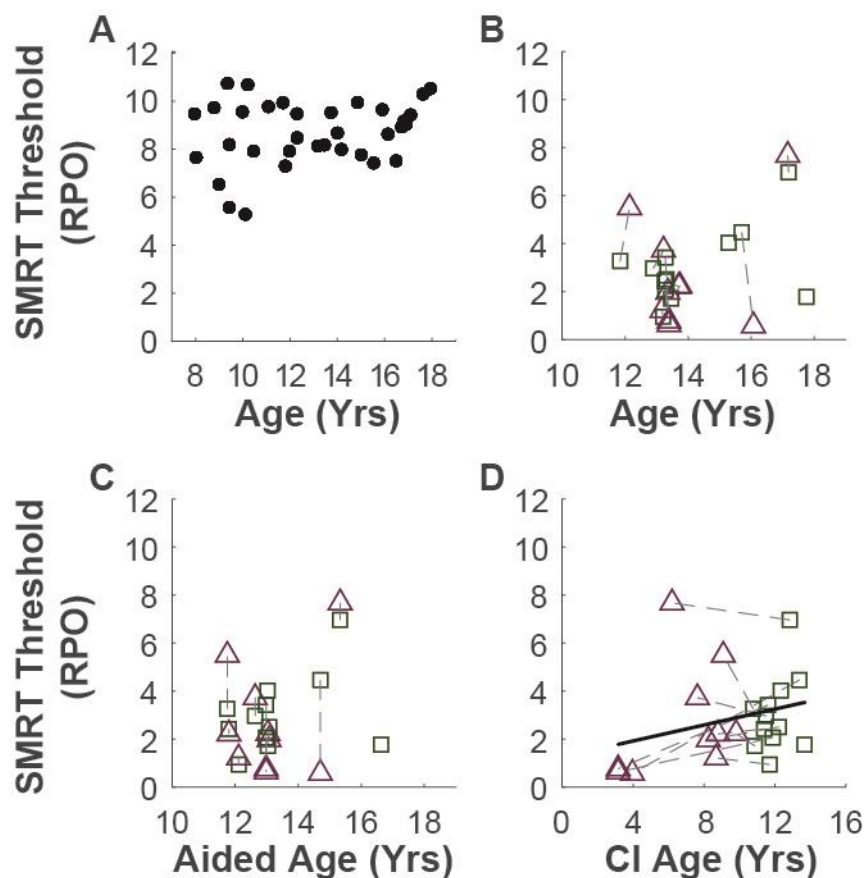


Figure 3-4. SMRT thresholds of children with NH and with CIs as a function of age-related variables.

(A) SMRT thresholds of NH children by age. Each circle represents data from one child. (B-D) SMRT thresholds of children with CIs by (B) chronological age, (C) aided hearing age, and (D) CI age. Each symbol represents data from one CI. Squares indicate first-implanted CIs and triangles indicate second-implanted CIs. Dashed lines connect data from the two CIs of each bilaterally-implanted child. Solid lines represent the lines of best fit for significant relationships.

Tests of multicollinearity revealed that the chronological and aided hearing age variables in models testing data from pediatric CI users were highly correlated with each other ( $r = 0.93$ ,  $p < 0.001$ ). Both variables also exhibited high VIF values (between 7.52 and 12.2) in all models. This likely occurred because most of the children received their hearing aid(s) shortly after birth, and thus aided hearing age values were very similar to those of chronological age. As chronological age was more relevant to the goals of the study than was aided hearing age, the repeated-measures models were run again after removing aided hearing age as a factor. Multicollinearity was no

longer present (VIF values between 1.01 and 1.07), and the pattern of results was the same as those found in the prior repeated-measures analyses: chronological age was not a significant predictor of any independent variable, and CI age was significantly related to vowel identification in quiet [ $F(1,9.3) = 10.0, p = 0.011$ ] and SMRT performance [ $F(1,10.2) = 10.8, p = 0.008$ ].

Consistent with previous studies in adults (e.g., Henry et al., 2005) and one study in NH children (Allen and Wightman, 1992), the next set of analyses revealed that spectral ripple discrimination performance was a significant predictor of vowel identification scores in quiet [ $F(1,15.4) = 20.0, p < 0.001$ ] of children with CIs. The relationship between SMRT threshold and vowel identification in noise was not significant ( $p = 0.16$ ; see Figure 3.5). However, again, only the better-performing children performed vowel identification in noise, which reduced the sample size for analysis.

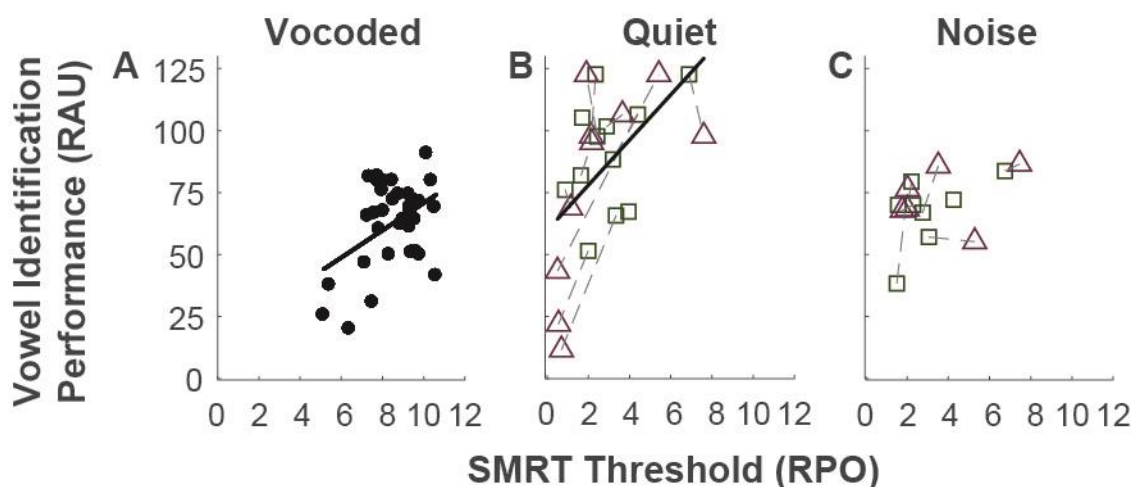


Figure 3-5. Vowel identification performance of children with NH and with CIs compared to SMRT thresholds.

(A) Vocoded vowel identification performance in rationalized arcsine units (RAU) of NH children plotted against SMRT thresholds. Each symbol represents data from one child. (B)

Vowel identification in quiet scores in RAU of pediatric CI users plotted against SMRT thresholds. (C) Vowel identification in noise scores in RAU of pediatric CI users plotted against SMRT thresholds. Each symbol represents data from one CI. Squares indicate first-implanted CIs and triangles indicate second-implanted CIs. Dashed lines connect data from the two CIs of each bilaterally-implanted child. Solid lines represent the lines of best fit for significant relationships.

No significant differences in performance were observed between first and second CIs of bilaterally-implanted children for vowel identification in quiet ( $p = 0.62$ ) or in noise ( $p = 0.51$ ) or on the SMRT ( $p = 0.57$ ). Some children performed better on these assessments with their first-implanted ear, while others exhibited better performance with their second (see Figure 3.6).

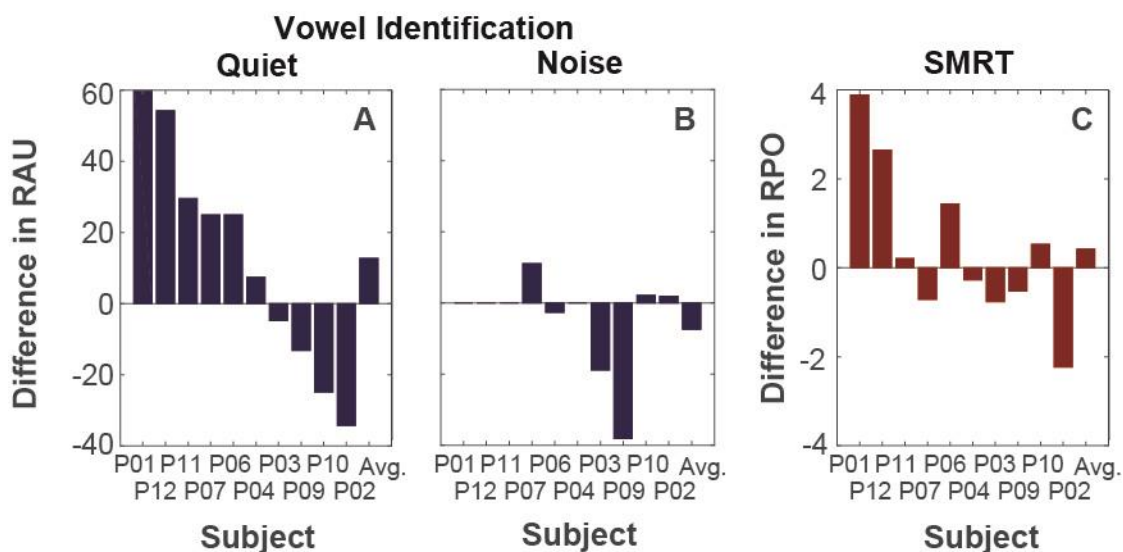


Figure 3-6. Comparison of task performance between first- and second-implanted ears of bilateral CI users.

Bars indicate differences in performance (first implant – second implant) on (A) vowel identification in quiet, (B) vowel identification in noise, and (C) the SMRT. Scores above 0 indicate better performance when tested with the first implanted ear. Subjects are ordered by amount of enhancement in vowel identification in quiet scores with their first-implanted ear compared to their second. Only subjects who scored >80% on vowel recognition in quiet were tested in noise (P02, P03, P06, P07, P09, P10).

Similar to the SINFA results observed in NH children, vowel feature information transmission values varied greatly among children with CIs. These values also differed between the two CIs of the same child (see Table 3.2). For vowel identification in quiet, information transmission ranged between 9% -100% for F1, 7% -100% for F2, and 4% - 100% for duration. For vowel identification in noise, they ranged between 21% - 94% for F1, 27% - 86% for F2, and 35% to 100% for duration. Results of the SINFA for individual ears of children with CIs are shown in Table 3.2

Table 3-2. CI user Sequential Information Analysis (SINFA) results.

N/A: information not available, as these children were not tested on vowel identification in noise.

Participant	Ear	Percent Information Transmitted					
		Quiet			Noise		
		F1	F2	Duration	F1	F2	Duration
P01	R	100	91	100	91	42	87
	L	35	35	18	N/A	N/A	N/A
P02	R	91	86	100	44	59	100
	L	100	100	100	44	43	61
P03	R	100	87.7	82.2	38	80	81
	L	100	89	100	87	83	55
P04	R	86	62	100	N/A	N/A	N/A
	L	79	50	85	N/A	N/A	N/A
P05	R	100	89.2	100	49	68	74
P06	R	100	100	100	80	71	85
	L	100	90	100	71	86	100
P07	R	100	100	100	89	69	85
	L	100	90	100	63	58	88
P08	L	58	73.4	74	N/A	N/A	N/A
P09	L	70	73	96	21	27	35
	R	85	85	100	72	70	65
P10	L	100	90	100	94	38	84
	R	100	100	100	50	76	67
P11	R	46	39	40	N/A	N/A	N/A
	L	9	23	4	N/A	N/A	N/A
P12	R	51	50	54	N/A	N/A	N/A
	L	3	7	20	N/A	N/A	N/A

Pediatric CI users demonstrated the same pattern of vowel feature ranking as children with NH: on average, the feature with the largest amount of information transmitted for vowel identification in quiet and in noise was duration (quiet:  $m = 81.6\%$ ,  $SD = 31.6$ ; noise:  $m = 76.2\%$ ,  $SD = 18.0$ ). The second-highest amount of information transmitted was found for F1 (quiet:  $77.9\%$ ,  $SD = 30.9$ ; noise:  $63.8\%$ ,  $SD = 23.0$ ), and the least amount for F2 (quiet:  $m = 73.2\%$ ,  $SD = 27.4$ ; noise:  $62.1\%$ ,  $SD = 18.3$ ). The amount of information transmitted for each feature was lower for vowel identification in noise compared to identification in quiet.

Examinations of the relationship between the amount of vowel feature information transmitted and SMRT thresholds revealed significant results for F1 [ $F(1,10.7) = 40.5$ ,  $p < 0.001$ ],

F2 [ $F(1,12.8) = 32.0, p < 0.001$ ] and duration [ $F(1,11.1) = 23.4, p = 0.001$ ] of vowel identification in quiet. All results were significant after correcting for multiple comparisons ( $\alpha = 0.016$ ), indicating that high transmission of any of the three vowel features when identifying vowels in quiet predicted better SMRT performance. As expected, no significant relationships were observed between SMRT thresholds and information transmission values of any feature from vowel identification in noise (F1:  $p = 0.16$ , F2:  $p = 0.71$ , duration:  $p = 0.29$ ), as performance on this test had not been found to relate to performance on the SMRT (see Figure 3.7).

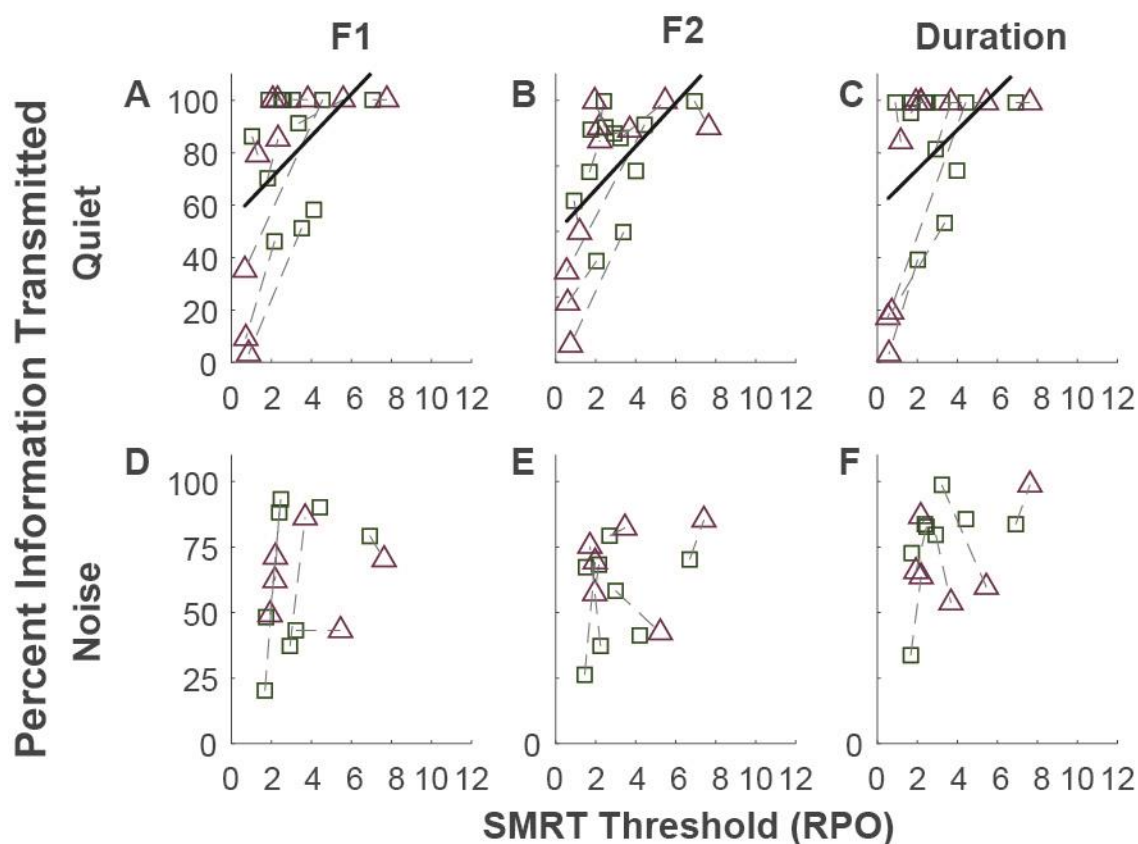


Figure 3-7. Relationships between vowel feature information transmission and SMRT thresholds in children with CIs.

Percent of information transmitted for F1 (A,D), F2 (B,E), and duration (C,F) from vowel identification in quiet (A – C) and vowel identification in noise (D – E) plotted against SMRT thresholds. Each symbol represents data from one CI. Squares indicate first-implanted CIs and triangles indicate second-implanted CIs. Dashed lines connect data from the two CIs of each bilaterally-implanted child. Solid lines represent the lines of best fit for significant relationships.

### *Comparison of SINFA Results between Children with NH and with CIs*

Statistical comparisons of the mean vowel feature information transmission values between groups indicated that NH children had significantly lower transmission of F1- [ $t(54) = 2.3, p = 0.027$ ] and F2-related information [ $t(54) = 2.4, p = 0.025$ ] for vocoded vowel recognition compared to that from pediatric CI users' identification of vowels in quiet. No significant differences were found between groups for transmission of the duration feature [ $t(54) = 1.6, p = 0.12$ ].

## DISCUSSION

This study examined the influence of age-related variables on pediatric CI users' vowel recognition scores, as well as the relationship between spectral modulation detection abilities and vowel identification performance. NH children completed tasks analogous to those performed by children with CIs to compare the differences in psychoacoustic abilities with acoustic hearing across ages to that with a CI. Each CI of bilaterally-implanted children was tested separately to assess age-related factors while limiting other potential sources of variability. In addition, a SINFA was performed to investigate whether the pattern of vowel feature perception differed between children with CIs and with NH. Information transmission values were also compared to SMRT thresholds to determine the relationship between conveyance of specific vowel features and spectral discrimination abilities.

### *Age-Related Factors*

#### Vowel Identification Performance

Vocoded vowel identification scores were found to increase with age of NH children. These results are consistent with those from previous studies that have observed age effects in children's recognition of spectrally-degraded stimuli. For example, Dorman et al. (2000) found that very young children required more channels of spectral information than adults did to correctly identify words processed through a sine-wave vocoder. Similarly, several prior studies have observed better performance on noise-vocoded word and sentence recognition in older compared to younger school-age children (Eisenberg et al., 2000; Eisenberg et al., 2002; Roman et al., 2016). The present study extended these findings to vowel stimuli, whose identification is more dependent on spectral cue resolution than are words or sentences. Data from the present study demonstrate

that vocoded vowel recognition performance may begin to plateau around age 14, the age of the oldest children tested in Eisenberg et al. (2000) and (2002) and Roman et al. (2016). However, several 16- and 17-year-olds performed at higher levels than children 14 years old and younger (see Figure 3.1). These results suggest that spectrally-degraded vowel identification performance might continue to improve through the late teenage years, but other sixteen- and seventeen-year-olds performed more poorly on vocoded vowel identification than did younger children. Thus, the number of children tested from each one-year age bin was too small, and the variability among these children's scores too large, to conclusively determine the age at which vocoded vowel recognition abilities are fully mature.

A number of factors may underlie the large amount of variance in vocoded vowel identification scores observed among the NH children. As unprocessed vowel identification matures at an age earlier than the youngest NH children tested (e.g., Allen & Wightman, 1992; Johnson et al., 2000), and since phonemes are the simplest speech stimuli, linguistic or vocabulary knowledge should not have significantly contributed to this variability. However, Roman et al. (2016) found correlations between children's performance on vocoded word and sentence identification and both auditory attention and short-term memory capacity. Further, because the vowel identification task required children to read and select the vowel sound they perceived from a visual display, differences in children's reading skills could have affected the difficulty of this test. These variables may explain the range of vocoded vowel identification scores that were observed even in NH children of the same age, although cognitive factors or reading skills were not tested in this study.

In children with CIs, vowel identification performance was significantly related to time with the implant, demonstrating the importance of auditory experience for development of speech recognition abilities. Previous research has shown that the speech perception performance of children with CIs indeed improves between implantation and 2-3 years later. Parkinson, el-Kholy, and Tyler (1998) tested prelingually deafened children (who received a CI between the ages of 4 and 14) on vowel identification at 12, 24, and 36 months after receiving their implant. They found that most children's vowel recognition scores improved significantly between 12 and 24 months post-implantation and continued to improve between the 24- and 36-month time points. Similarly, Wang et al. (2008) tested prelingually deafened, early-implanted children with CIs on a hierarchy of speech identification tests at 6, 12, 18, and 24 months after implantation. Children were required

to obtain a criterion score on each test before moving on to a progressively more difficult test. They found that at 6 months post-implantation, most children did not perform well enough to move past the first test. However, by 24 months after receiving their CI, many children were able to advance to the most difficult assessment. Results from the current study demonstrate that children may continue to improve on speech identification longer than 3 years post-implantation.

Separate analyses of first- and second-implanted CIs showed that the relationship between vowel identification scores in quiet and CI age was significant for the second- but not first-implanted ears. A t-test revealed that the average CI age of first implants ( $m = 12.02$ ,  $SD = 0.92$  years) was significantly greater than that of second implants [ $m = 6.87$ ,  $SD = 2.86$ ;  $t(10.9) = 6.02$ ,  $p < 0.001$ ]. The findings from the separate-ear analyses may be indicative of an upper limit on vowel identification improvement with time. However, a much broader range of CI ages as well as vowel identification scores existed for second- compared to first-implanted ears, potentially explaining why such a relationship was detected for second- but not first-implanted ears.

#### SMRT Thresholds

In the present study, no relationship was found between age and spectral discrimination performance of NH children aged 8 years and older. However, previous studies which have found improvements in SRD abilities beyond age 9 utilized traditional SRD tasks (e.g., Peter et al., 2014; Rayes et al., 2014). The test used in the present study, the SMRT, attempts to control for the potential confound of within-channel intensity cues that may exist in traditional SRD tests (Aronoff & Landsberger, 2013). Intensity resolution continues to develop through at least early adolescence in NH children (e.g., Maxon & Hochberg, 1982), and this maturation is likely further delayed in children with CIs (Park et al., 2015). Older children may perform better than younger children on traditional SRD tests because they are better able to discriminate between within-channel intensity differences in the stimuli. It is therefore possible that performance on the SMRT, which may reduce these developmental effects of intensity resolution, could mature at a younger age than performance on a traditional SRD task. Results from studies conducted by Kirby et al. (2015) and Landsberger et al., (2017) provide evidence for this theory: both investigations found an effect of age on SMRT thresholds in children aged 6 to 12 years old, but observed that SMRT performance began to peak at age 8 or 9. These findings, in combination with those from the present study, suggest that NH children's performance on the SMRT (which diminishes the

potential confound of immature intensity resolution within a single channel) may mature at 8 or 9 years old. Comparison of SMRT and traditional SRD test performance within the same group of children is thus necessary to determine whether age-related performance differences indeed exist between traditional SRD tasks and the SMRT.

While no significant relationship was found between SMRT performance and age in the group of NH children tested in the present study, an age-related factor, CI age, was found to predict SMRT thresholds of pediatric CI users in the repeated-measures statistical model. Maturation of SMRT performance with a CI may be prolonged compared to NH children; in this study, NH children's spectral discrimination abilities seem to have matured by age 8 but pediatric CI users' performance seems to improve up to a "CI age" of 13.7. As delayed maturation of performance on other psychoacoustic tasks has been observed in children with CIs (e.g. Park et al., 2015), the period of auditory deprivation prior to receiving a CI as well as the degraded auditory signal through a CI compared to acoustic hearing may indeed impede maturation of spectral discrimination abilities.

Although the results from the present study suggest that CI age may play a role in SMRT performance of pediatric CI users, a recent study conducted by Landsberger et al. (2017) did not find an effect of time with the CI on SMRT thresholds of early-implanted children. However, these studies assessed participant populations that differed in two key ways. While both groups of participants received their first implant very early in life, the children tested in the current study were older (aged 11.94 to 17.92 years) than those tested in Landsberger et al. (aged 6.0 to 13.1 years), resulting in older "CI ages" as well. Perhaps the effect of time with the CI on spectral resolving capabilities does not appear until the child has had their CI for longer periods of time. Additionally, the interaction between development of the auditory system and auditory experience may be more discernable at later chronological ages.

Another possible explanation for the dissimilar findings between the two studies could be the difference in statistical methods utilized. Landsberger et al. (2017) performed separate correlation analyses between time with the implant and SMRT thresholds for first- and second-implanted CIs of bilaterally-implanted children; the current study combined data from both CIs of bilaterally-implanted children into a mixed-model repeated-measures design. This model inherently possessed greater power for detecting statistical effects because repeated-measures analyses control for between-subject error. Indeed, when the authors of the present study separated

the dataset by ear implanted, and also performed correlation analysis akin to those performed by Landsberger et al., the results were not significant ( $r = 0.35$ ,  $p = 0.27$  for first-implanted CIs;  $r = 0.35$ ,  $p = 0.33$  for second-implanted CIs). Still, because these children do have two CIs, not one, a repeated measures design is a more appropriate statistical analysis to determine the influence of CI age for bilaterally-implanted children: such an analysis simultaneously considers the effects of auditory experience on spectral ripple discrimination performance of each individual CI. Despite these differences in methods and subject populations, the discrepancy in results between the two studies could potentially be explained by the large degree of variability among pediatric CI users. Additional studies utilizing the SMRT in children with CIs are necessary to fully elucidate the relationship between time with the CI and spectral discrimination performance with this test.

#### *Relation between Spectral Discrimination and Vowel Identification Performance*

This is the first study to our knowledge to examine the relationship between performance on a modified version of traditional SRD tests, the SMRT, and speech identification performance of children. SMRT thresholds were found to positively relate to vocoded vowel identification scores of NH children and vowel identification performance in quiet of children with CIs. These relationships are consistent with those found previously between performance on traditional SRD tasks and scores on speech identification tasks in adults with NH (Henry & Turner, 2003) and with CIs (e.g., Henry et al., 2005; Won et al., 2007; Won et al., 2011), as well as those previously observed between SMRT and speech identification performance of adult CI users (Holden et al., 2016; Lawler, et al., 2017; Zhou 2017). Data from the present study demonstrate that the SMRT, while distinct from traditional SRD tasks, could be used to predict vowel identification performance of children.

Previous studies with pediatric CI users have observed varying relationships between SRD thresholds and speech identification scores. Specifically, Jung et al. (2012) found a strong correlation between SRD performance and identification of spondees in noise in prelingually deafened, early-implanted children, but did not observe this result when correlating SRD thresholds and monosyllabic word recognition scores. Similarly, Horn et al. (2017) found significant positive relationships between SRD performance and spondee identification in noise in early-implanted, school-age children, but only in two of five SRD ripple depths tested. In both studies, the correlations observed in children with CIs were weaker than those observed in adults.

The differences between these findings and those from the present study may be due to speech stimuli: formant frequencies are an important cue for accurate vowel identification (e.g., DiNino et al., 2016; Shannon et al., 2002), so vowels may be more related to spectral discrimination than are monosyllabic words. The vowels within spondees are indeed the main cue for their recognition, and this may explain why these prior investigations have found some positive relationships (although weaker than the relation observed between SMRT performance and vowel identification scores in the present study) between SRD thresholds and spondee identification in noise. Additionally, the speech stimuli for NH listeners in Jung et al. (2012) and Horn et al. (2017) were unprocessed and presented in quiet or in noise, whereas the vowel stimuli in the present study were vocoded. Perhaps SRD requires cognitive or developmental mechanisms that are more related to identifying spectrally-degraded speech than they are to recognizing speech in quiet or in noise. Further, the SMRT attempts to control for potential local loudness cues in traditional SRD stimuli (e.g. Aronoff & Landsberger, 2013). This test may thus eliminate the developmental confound of immature single-channel intensity resolution, which would delay maturation of children's performance on traditional SRD tasks compared to performance on the SMRT. If so, based on results from prior studies in children that utilized traditional SRD tests, frequency resolution as measured by the SMRT may be a better predictor of speech identification abilities of children than are other SRD tasks.

While the present study found a strong, positive relationship between SMRT thresholds and vowel identification performance in children with CIs, an issue with presenting SRD stimuli to CI users warrants caution when interpreting such results. The CI speech processor contains a small number of frequency analysis channels (15 for those tested in the present study), which limits the number of spectral peaks and valleys that can be transmitted. Recent evidence demonstrated that CI speech processor output is unpredictable for stimuli above about 2 RPO (O'Brien & Winn, 2017), and therefore, there is no monotonic relationship between ripple input and processor output that can be interpreted as a single spectral dimension. Above a critical ripple density, the spectral envelope is aliased (akin to aliasing of frequencies when a sound is under-sampled), so the moderate- and high-density ripple stimuli contain spectral cues not intended by the experimenter. It is likely that CI users who achieve high spectral ripple discrimination thresholds may be able to utilize different perceptual strategies that enable them to discern these non-linear and non-monotonic spectral distortions, which consequently improves their performance on the test.

The SMRT differs from traditional SRD tasks in that the stimuli contain a temporal cue in addition to the spectral signal; however, when the SMRT is presented through a CI processor, the amplitude modulation of this temporal component has also been found to distort at high RPO values. Lawler et al. (2017) recorded electrograms from Advanced Bionics processors (the device used by all children with CIs in this study) and observed temporal smearing of the signal as ripple density increased from one to three RPO. The amplitude modulation depth of the temporal signal became shallower with increasing RPO values. The temporal component of SMRT stimuli may be thought of as an additional cue for SMRT discrimination, and substantial distortion of this cue through a CI processor may not occur until a larger RPO value than that at which spectral aliasing occurs; yet, if the spectral signal is smeared, both cues are no longer monotonically changing with increasing RPO. The stimulus parameter is thus unpredictable beyond the RPO value at which spectral distortion occurs.

As many children in the present study achieved high SMRT thresholds, the relationship between SMRT performance and vowel identification was examined again after setting an SMRT threshold upper limit. A similar analysis was conducted by O'Brien and Winn (2017), in which they set a cutoff value of 2.56 RPO and recalculated the average thresholds for each participant. This cap on SRD thresholds resulted in stronger relationships between spectral ripple discrimination and scores on speech-based outcome measures than in the analyses which had included higher threshold values. The cutoff value of 2.56 RPO was based on the frequency allocation table of Cochlear devices and was determined to be the critical limit after which ripple stimulus aliasing occurs in the spectral domain. However, CI participants in the present study used Advanced Bionics devices, which contain a smaller number of frequency analysis bands than Cochlear devices. An approximation of the critical RPO limit for Advanced Bionics CIs was conducted and yielded a value of 1.46 RPO. All individual SMRT runs with thresholds greater than 1.46 RPO were changed to this value and the model relating vowel identification scores to SMRT performance was run again. The result was statistically significant with a higher certainty [ $F(1,11.4) = 35.4, p < 0.001$ ] than the original model that had included SMRT runs with higher thresholds.

The non-monotonic relationship between rippled stimuli input and CI processor output may also explain why the ability to “discriminate” stimuli with high spectral ripple rates, as indicated by high SRD thresholds, has been found to consistently relate to identification of speech, which

contain low densities of spectral peaks and valleys. The spectral modulation rates of vowels are less than 1-2 peaks per octave; this explains why setting a critical limit on SMRT thresholds from the present study resulted in a significant relation to vowel recognition performance with a larger *F* value than in the model that included RPO thresholds above 1.46 RPO. Discriminating between medium- to high-density ripple stimuli does not assess aspects of spectral resolution that relate to important features of speech (e.g., Saoji et al., 2009). An individual who demonstrates high SRD thresholds is likely instead using some beneficial perceptual process that they may also access to accurately identify speech sounds.

### *SINFA Results*

Statistical comparisons of SINFA results between groups indicated that children with CIs had significantly higher F1- and F2-related information transmission compared to NH children. The spectral information of vocoded vowel stimuli is degraded, but in a similar method as through a CI. This difference in results may be due to acclimatization (or, lack thereof). Perhaps, over time and with CI experience, children with CIs are able to resolve the frequency components of an auditory signal more effectively. NH children performed vocoded vowel identification without prior exposure to the stimuli. It is possible that allowing NH children more practice with the vocoded stimuli would result in better F1 and F2 feature perception of the spectrally-degraded vowels.

Interestingly, despite significantly higher transmission of F1- and F2-related information in the children with CIs compared to NH listeners, both groups demonstrated similar values for transmission of the duration cue. The SINFA also revealed greater amounts of duration-related feature transmission compared to F1 or F2 cue transmission in both groups of children. These results are consistent with findings from previous studies in adults: duration is an important cue for vowel recognition when the contrasts between spectral cues are diminished (e.g. Ainsworth, 1972). Due to the reduced spectral resolution of the implant, adults (e.g., Winn, Chatterjee, & Idsardi, 2012) and children (e.g., Nittrouer et al., 2014) with CIs have been found to rely less on spectral information and more on duration cues for speech identification compared to NH individuals. In addition, vocoder studies of NH adults have found greater transmission of duration-related information than F1- or F2-related information in vowel identification conditions with considerable spectral degradation (e.g. Xu, Thompson, & Pfingst, 2005). A previous study in NH

adults with the same vocoder processing and vowel stimuli as the present study also demonstrated this pattern of SINFA results (DiNino et al., 2016); the current investigation extends these findings to NH children. This study provides further evidence that under conditions of spectral degradation, children and adults utilize temporal information to a greater extent than spectral cues to discriminate between vowel sounds because the spectral information is reduced.

The amount of F2 information transmitted in NH children's vocoded vowel recognition was significantly related to their SMRT performance. These results may pertain to the redundancy of vowel features, as F1 and duration covary – the F1 of vowels corresponds to vowel height, and high vowels (those with low F1 frequencies) are generally shorter in duration than are low vowels (those with high F1 frequencies; Heffner, 1937). The F2 feature may simply be more salient in the SINFA, which assumes independence of phoneme features. Still, F2 corresponds to vowel advancement; this feature is akin to place of articulation, the primary spectral cue for consonant recognition, and perception of place of articulation is greatly affected by spectral smearing (e.g., Boothroyd et al., 1996). NH children's perception of the advancement feature of vocoded vowels may be positively related to SMRT performance because the ability to utilize F2 cues in the presence of spectral degradation is analogous to spectral discrimination abilities.

As in NH children, F2 information transmission values were found to significantly relate to SMRT thresholds of children with CIs; however, transmission levels of the F1 and duration features were also positively related to SMRT performance of children in this group. It seems that high transmission rates of any vowel feature, even the non-spectral feature of duration, was correlated with spectral discrimination abilities of pediatric CI users. Greater perception of the vowel duration feature may relate to better SMRT performance because SMRT stimuli include a temporal signal in addition to a spectral signal. It is possible that individuals with CIs utilize similar temporal mechanisms to identify vowels and perform the SMRT. Nevertheless, the relationship between duration transmission values and SMRT thresholds was weaker than those observed between the transmission of formant features and SMRT thresholds (see Figure 3.7).

The finding that spectral cue perception of speech sounds is related to spectral discrimination abilities of child CI users is similar to that observed by Winn et al. (2016) in a study with adults with CIs: they found that usage of a formant cue for categorizing speech sounds was significantly related to SRD thresholds of adult CI listeners. Evidence from that study also suggested that formant cue categorization relates more strongly to speech identification abilities

than does SRD performance. Evaluation of the phonetic cues that children with CIs rely on to categorize speech sounds may further elucidate the relationship between utilization of spectral and non-spectral speech-based cues, spectral discrimination performance, and speech recognition scores.

### *Comparisons Between CIs of Bilaterally-Implanted Children*

Examination of vowel identification scores, SMRT thresholds, and SINFA results from individual CIs of bilaterally-implanted children consistently revealed differences in performance and in information transmission values between the two ears of the same child. These results suggest that factors related to the individual CIs of bilaterally-implanted children can influence both spectral discrimination and perception of speech sounds. Peripheral variables (such as differences in individual device program processing and in the electrode-neuron interface) as well as variables that could affect function of the central auditory system (such as age at implantation and amount of time with each CI), likely underlie the discrepancies in performance between the two ears of bilaterally-implanted pediatric CI users. The current study provided evidence for the contribution of age-related variables on psychoacoustic task performance. Future investigations could examine the relationship between peripheral factors and bilaterally-implanted children's auditory abilities with each CI.

Previous studies of bilateral, sequentially-implanted individuals with CIs have observed better performance on speech identification in quiet and noise with the first implanted CI compared to the second in adults (Ramsden et al., 2005) and in children (Reeder et al., 2017). However, in the current study, no significant differences in average vowel identification performance in quiet or in noise were observed between first- and second-implanted CIs. Further, not all bilaterally-implanted children performed better with their first-implanted ear. Four of the ten children (P02, P03, P09, and P10) obtained higher scores on vowel identification with their second-implanted CI. Two of the children who exhibited large gains in performance with their second-implanted compared to their first-implanted ear also had the shortest time periods between first and second implantations: 1.34 (P09) and 1.99 (P02) years. Previous studies have observed poorer speech recognition scores with the second-implanted ear if the duration between implantation is long (e.g., Reeder et al., 2014); these results provide additional evidence that simultaneous or quick

successive bilateral-implantation can result in good or equivalent speech identification outcomes for the second-implanted ear.

Additionally, no significant differences in SMRT thresholds were found between first- and second-implanted CIs of the participants. These differences are in contrast with results found by Landsberger et al., (2017) who observed that SMRT thresholds were significantly better for first-implanted ears compared to second-implanted ears of children with CIs. Half of the bilaterally-implanted children tested (five of ten) in the current study performed better on the SMRT with their second-implanted ear compared to their first. While testing conditions in Landsberger et al. were completely randomized and those in the present study were not, this difference in results is unlikely to be explained by practice effects: two children in the current study (P04 and P06) completed the assessments with their second-implanted CI prior to testing with their first-implanted ear, and these were two of the five children who demonstrated higher SMRT thresholds with their second-implanted ear.

The finding that some children performed better on auditory tasks with their second-implanted ear compared to their first indicates that the relationship observed between CI age and these assessments was not simply due to the children performing better with their earlier-implanted CI; rather, the distribution of CI ages in this study engendered significant effects of this variable on auditory task performance. Still, numerous prior studies have demonstrated that earlier age at implantation is associated with higher scores on assessments of speech perception and language comprehension (e.g., Svirsky, Teoh, & Neuburger, 2004; Geers, et al., 2009; Niparko et al., 2010; Ruffin et al., 2013; Tobey et al., 2013). In these studies, children who were implanted before age five (and often earlier) demonstrated the best outcomes. All children in the present study received their first CI before age five, but most bilaterally-implanted children did not obtain their second implant until years later. Because most of the children in this study were around the same age, early receipt of a CI tended to be associated with a greater CI age in that ear; therefore, we performed follow-up analyses to disentangle the effects of time with the CI and age at implantation on the task performance observed in this study.

Repeated-measures mixed-model regression analyses were conducted with both age at implantation and CI age as predictor variables, as multicollinearity was low. Vowel identification scores in quiet (model 1) and SMRT thresholds (model 2) were the dependent variables. Results indicated that age at implantation did not significantly predict vowel identification performance in

quiet [ $F(1,10.8) = 1.1, p = 0.324$ ] or SMRT performance [ $F(1,9.0) = 4.1, p = 0.072$ ] of the pediatric CI users in this study. Further, these analyses revealed that CI age had larger effects on auditory task performance compared to those of age at implantation, as well as a significant relation to SMRT thresholds [ $F(1,7.8) = 9.7, p = 0.015$ ] and a nearly-significant relationship to vowel identification in quiet [ $F(1,11.0) = 4.5, p = 0.057$ ]. This suggests that the diminished spectral resolution of a CI could play a larger role in hindering maturation of vowel identification scores and SMRT thresholds than does the amount of time prior to receiving an implant. Further, these results indicate that age at implantation is not driving the relationships between CI age and auditory task performance observed in the population of CI users tested in this study.

At least some of the improvement over time with the CI is likely due to acclimatization, or adjusting to hearing with the implant. Adult CI users who lost their hearing later in life, and thus possess fully-developed auditory systems, also tend to improve on tests of speech identification over time after implantation. This increase in performance presumably occurs as they adapt to perceiving sound through the implant, which is very different than the acoustic hearing they had previously. However, such improvement in adults appears to plateau after only about one year (Ruffin et al., 2007; Zhang et al., 2015). A longitudinal investigation of speech identification performance of both adults and children with CIs demonstrated that adults' speech recognition scores peaked prior to two years post-implantation, while pediatric CI users' scores continued to improve throughout the four years of the study (Oh et al., 2003). In addition, performance on SRD tasks has been observed to be stable between one month and twelve months after implantation in adults with CIs (Drennan et al., 2016). Because the present study found that CI age was positively related to SMRT and vowel identification performance, and the difference in CI ages between the children tested was generally greater than one year, these results provide further evidence that children with CIs improve on auditory tasks for a longer duration than do adults with CIs. Further, the children tested in this study had little to no acoustic hearing at any time prior to receiving the implant. As a result, while acclimatization may have driven these children's improvement in psychoacoustic task performance for a short period of time, development of the auditory system is likely contributing to the improvements that occur over a longer duration.

The findings from this study suggest that immature auditory perceptual abilities may contribute to the variability among pediatric CI users' psychoacoustic task performance. However, children's improvements in auditory task performance over time with the CI is not expected to

continue indefinitely; rather, some upper limit likely exists on the development of auditory abilities. Based on the coincident developmental timelines of psychoacoustic task performance and central auditory system function (e.g. Eggermont & Ponton, 2003), this limit is presumably set by the development of the auditory cortex. The current study was not a longitudinal study and thus the age of vowel identification and SMRT maturation in children with CIs was not identified. Future research with a longitudinal design, larger sample size, and broader age range of child CI participants is warranted to form definitive conclusions about auditory development with a CI.

This investigation had several limitations. “Aided hearing age” and “CI age” represented auditory experience in this study, but the progression of hearing loss and effectiveness of aided hearing, which also define auditory experience, could not be quantified over time for the pediatric CI users who participated in this research. This study also did not assess cognitive or intellectual variables such as auditory attention, short-term memory capacity, and reading abilities. These factors could have contributed to the auditory task performance of children in this study. For example, the CI participant P06 performed very well with their second-implanted CI despite the moderate CI age of this ear; as P06 also performed well with their first-implanted ear, cognitive or intellectual factors could have played a role in this subject’s vowel identification scores and SMRT thresholds. An additional limitation to this study was the small sample size of pediatric CI users, which may have reduced the statistical power to find a significant relationship between predictor variables and vowel identification in noise, as well as between CI age and SMRT thresholds for the analyses conducted with first- or second-implanted CIs alone. Finally, this investigation was limited by the age range of children with CIs who met study eligibility criteria. Future studies could examine younger pediatric CI users.

## CONCLUSION

This study demonstrated that chronological age predicted spectrally-degraded vowel identification performance of NH children, but was not related to performance on any assessment in children with CIs. The amount of time with the implant was a better predictor of vowel recognition performance, as well as spectral discrimination abilities, than was chronological age. Further, spectral resolving capabilities significantly predicted vowel identification performance in both NH children and children with CIs. SINFA results revealed that transmission of F2-related information in vowel perception was positively related to SMRT thresholds of children with NH.

However, transmission of F1, F2, and duration were all significantly related to SMRT thresholds of children with CIs, indicating that perception of any vowel feature may predict spectral discrimination abilities of these children. Further, differences in performance and vowel feature information transmission were observed between individual CIs of the same child, indicating a need for additional investigation of device-specific factors related to the electrode-neuron interface that could contribute to auditory task performance of these children.

The findings from this study have implications for assessment of verbal language abilities of pediatric CI users at particular time periods after implantation; performance on such tests may continue to improve throughout childhood and/or adolescence. Results from this study suggest that normal auditory system development could interact with auditory experience to enhance speech perception and spectral discrimination abilities of children.

## CHAPTER 4 : HEARING HISTORY AFFECTS THE ELECTRODE-NEURON INTERFACE IN COCHLEAR IMPLANT LISTENERS: A COMPARISON BETWEEN EARLY-IMPLANTED CHILDREN AND LATE-IMPLANTED ADULTS

Mishaela DiNino, Gabrielle O'Brien, Steven M. Bierer, Kelly N. Jahn, and Julie G. Arenberg

### ABSTRACT

Cochlear implant (CI) programming is similar for adults and children despite limited understanding of the electrode-neuron interface (ENI) in children. The ENI is the effectiveness of each CI electrode to activate target auditory neurons and is primarily influenced by three factors: electrode position, neural health, and bone and tissue growth in the cochlea. Hearing history likely affects these patient-specific variables, suggesting that the efficacy of each channel of stimulation may differ between children who were implanted at young ages and adults who lost their hearing and received a CI much later in life. Therefore, this study examined whether hearing history differentially affected measures of ENI quality. Auditory detection thresholds and Most Comfortable Levels (MCLs) were obtained with monopolar and focused electrode configurations in early-implanted children and late-implanted adults. Channel-to-channel variability and dynamic range were calculated for both types of stimulation. Electrical field imaging data were also acquired to estimate levels of intracochlear resistance from voltage gradients of CI electrodes. Results indicated that children had lower average auditory perception thresholds compared to adults, particularly with focused stimulation. Dynamic range did not differ between groups but children demonstrated lower MCLs, indicating that the range of perceptible current was shifted downward in children. These findings suggest that early-implanted children have better neural health compared to late-implanted adults. Yet, channel-to-channel variability were not significantly different between groups, indicating that children exhibit the same variation in ENI quality as do adults. Finally, children had increased intracochlear resistance levels relative to the adult group, possibly reflecting higher levels of ossification or tissue growth after CI surgery. These results illustrate physical and perceptual differences related to the ENI of early-implanted children compared to late-implanted adults. Evidence from this study suggests that further

investigation of CI programming interventions tailored to the ENIs of individual CI users are warranted.

## INTRODUCTION

Newborn hearing screening has led to earlier detection of hearing loss in children, resulting in an increasing number of children who receive cochlear implants (CIs) at young ages (e.g., Boons, et al., 2013; Halpin et al., 2010; Lammers et al., 2015). In contrast, many current adult CI users had maintained acoustic hearing until later in life and were implanted at older ages. These two groups of CI listeners thus have distinct etiologies of deafness, ages at implantation, and often, durations of hearing loss. These divergent hearing histories may result in differences in the interface between CI electrodes and spiral ganglion neurons between these populations (Bierer, 2010).

The electrode-neuron interface (ENI) refers to the effectiveness with which each electrode activates the target spiral ganglion neurons to transmit an auditory signal. Suboptimal interfaces can result from electrodes positioned distant from spiral ganglion neurons (Finlay & Skinner, 2008), degeneration of those neurons (Miura et al., 2002), or bone and tissue growth near stimulating electrodes (Seyyedi & Nadol, 2014). Electrode position should not differ between children and adults, as the cochlea is adult-like at birth. However, age at implantation and duration of deafness differ substantially between these groups and likely impacts neural health and tissue growth in the cochlea. Degeneration of peripheral and central auditory system neurons occurs in the absence of auditory input (Otte et al., 1978) and with normal aging (Makary et al., 2011). Thus, early-implanted children may have healthier auditory neurons compared to adults who received a CI years after diagnosis of severe to profound hearing loss.

To assess the quality of the ENI in adult CI users, several studies have narrowed the spatial extent of electrical current spread by routing the return current path primarily within the cochlea (Bierer et al., 2010; Kral et al., 1998; Synder et al., 2004). This focused stimulation is believed to be more sensitive to the quality of the ENI than monopolar stimulation: several studies in adults have observed similar monopolar auditory perception thresholds across the CI array within individual subjects, but highly variable focused thresholds for the same individuals (e.g., Bierer, 2007). Variability in channel-to-channel thresholds likely reflects variation in the quality of the ENI, which may contribute to inconsistent transmission of spectral information. Accordingly, prior

studies in adults have shown that greater focused threshold variability is associated with lower speech recognition scores (Long et al., 2014). However, little is known about auditory perceptual thresholds across the electrode array in children with CIs. Therefore, one goal of the current study was to utilize thresholds to assess the quality of the ENI of early-implanted pediatric CI users.

In adults with CIs, channels with relatively elevated focused thresholds often have small dynamic ranges, defined as the difference between threshold and Most Comfortable Level (MCL; Bierer & Nye, 2014). Additional studies have observed that MCLs determine the dynamic range more than thresholds (e.g., Firszt et al., 2002), suggesting that suboptimal ENIs may also reduce MCLs. Small dynamic ranges in adults (Firszt et al., 2002) and children (Robinson et al., 2012) correlate with poorer speech identification performance, which could be particularly detrimental to children because language development depends on their capacity to perceive speech sounds. Yet, the relation between the measures assessing the ENI and dynamic range in children is unknown.

Electrical field imaging (EFI) is another technique that could provide information about the effect of hearing history on the cochlear environment. EFI uses the built-in telemetry systems of modern CIs to measure how intracochlear potential is distributed in the cochlea (for review, see Mens, 2007). Although EFI measurements cannot inform experimenters of the status of the auditory nerve, the profile of intracochlear potential depends critically on the anatomy of the cochlea and may be one tool for making inferences about the status of tissue surrounding the electrode array (Vanpoucke et al. 2004). To our knowledge, there has not been a systematic collection of EFI measurements in children. The results of EFI measurements may provide cursory evidence of anatomical and physiological differences between children and adults that are pertinent to clinical practice.

At present, investigations of the ENI of CI users have been conducted exclusively in late-implanted adults (e.g., Bierer et al., 2007; Finley et al., 2008; Long et al., 2014). This study tested the hypothesis that factors related to the ENI would differ between early-implanted children and late-implanted adults due to these groups' divergent hearing histories. Further, investigation of the ENI in early-implanted children would provide insight into the effects of electrical stimulation after early hearing loss.

## METHODS

### *Participants*

Eleven children aged 11 to 17 (mean age = 14.06, standard deviation = 1.9 years) and eleven adults aged 48 to 84 (mean age = 63.08, standard deviation = 11.6 years) with CIs participated in this study. Half of the children (P01, P02, P03, P10, P11, and P12) failed a newborn hearing screening. The other children were diagnosed with severe to profound hearing loss prior to age four (P04, P05, P06, P07, P08, and P09). All children received their first implant before age five (mean age at implantation = 2.17, standard deviation = 1.2 years). All adults were born with acoustic hearing, were diagnosed with severe to profound hearing loss in adulthood (mean age at diagnosis = 32.6, standard deviation = 13.2 years), and received their first CI at much older ages compared to the children (mean age at implantation = 56.1, standard deviation = 13.5 years). Demographics for both groups are shown in Table 4.1. Ten children and three adults were bilaterally-implanted; however, due to the large difference in number of bilaterally-implanted individuals between the groups, only first-implanted CIs were included in this study.

Pediatric and adult CI users were recruited from hospitals and audiology clinics in the greater Seattle area. All participants used oral communication and were native speakers of American English. To minimize potential sources of variability, all participants utilized Advanced Bionics HiRes 90K devices and HiFocus 1J electrode arrays. In addition, individuals with electrode positioners were excluded from this study because 1) positioners are designed to place the electrodes closer to target neurons, and distance could affect auditory detection thresholds (e.g., DeVries et al., 2016; Long et al., 2014), and 2) intracochlear impedance as assessed by EFI are likely sensitive to the volume of the fluid-filled space around the electrode array; positioners fill much of that space and could alter the impedance estimates.

Table 4-1. Demographic information for all child and adult participants.

a. EVA: enlarged vestibular aqueduct. b. DFNB1: genetic nonsyndromic hearing loss.

<i>Subject:</i>	<i>Gender:</i>	<i>Etiology of Deafness:</i>	<i>Ear:</i>	<i>Age at Testing:</i>	<i>Age at Implantation (yrs):</i>	<i>Duration of Deafness (yrs):</i>
<b><u>Children</u></b>						
P02	M	EVA	R	11.8	1.1	1.0
P03	M	Unknown	R	12.9	1.4	1.1
P04	F	Unknown	R	13.2	1.5	0.8
P05	M	DFNB1	R	17.7	4.1	3.0
P06	F	Unknown	R	17.2	4.3	2.5
P07	F	Unknown	R	13.3	1.9	0.4
P08	M	EVA	L	15.3	2.9	0.7
P09	F	Unknown	L	13.5	2.6	1.3
P10	M	DFNB1	L	13.3	1.1	0.9
P11	F	DFNB1	R	13.3	1.4	1.2
P12	M	DFNB1	R	13.3	1.7	1.4
<b>Mean</b>				<b>14.1</b>	<b>2.2</b>	<b>1.3</b>
<b>SD</b>				<b>1.9</b>	<b>1.2</b>	<b>0.8</b>
<b><u>Adults</u></b>						
S28	F	Autoimmune	R	74.9	69.7	18.8
S29	M	Unknown	L	84.0	76.8	30.3
S38	M	Otosclerosis	L	49.8	46.2	28.3
S39	F	Hereditary	R	53.4	30.1	9.1
S41	M	Maternal Rubella	L	48.8	42.9	1.2
S45	F	Hereditary	R	62.7	54.0	32.0
S46	M	Unknown	R	67.2	64.2	48.2
S48	F	Autoimmune	R	59.4	58.0	22.0
S50	F	Measles	R	76.5	61.1	41.1
S52	F	Unknown	R	70.1	66.0	21.1
S55	F	Hereditary	R	63.7	48.5	7.3
<b>Mean</b>				<b>64.6</b>	<b>56.1</b>	<b>23.6</b>
<b>SD</b>				<b>11.4</b>	<b>13.5</b>	<b>14.3</b>

## *Assessments*

### Auditory Detection Thresholds

Thresholds with monopolar and focused electrical stimulation were obtained using a fast sweep procedure with current steering (as in Bierer et al., 2015a, and based on Sek et al., 2005). Focused stimulation was implemented with the steered quadrupolar (sQP) electrode configuration, which involves four intracochlear electrodes: the two middle electrodes serve as active electrodes and the two outer electrodes serve as the returns. Sigma ( $\sigma$ ) specifies the amount of current delivered through the intracochlear return electrodes, with the remainder flowing through an extracochlear ground. Greater sigma values indicate greater current focusing, such that  $\sigma = 0$  represents monopolar stimulation, in which all return current flows through the extracochlear electrode, and  $\sigma = 1$  represents the greatest degree of focusing, in which the intracochlear electrodes carry all return current.

The sQP configuration allows current steering between the two active electrodes. Alpha ( $\alpha$ ) is the steering coefficient: when  $\alpha = 0$ , all steered current passes through the more apical active electrode; when  $\alpha = 1$ , all steered current passes through the more basal active electrode. Electrode channel number (i.e., “channel 4”) is typically defined by the basal active electrode. However, because the sQP configuration requires four electrodes, this arrangement is not possible to steer current for electrode 2. Instead, the same set of electrodes as steering for channel 3 are used, and an  $\alpha$  value of 0 is set to center the current on electrode 2. This arrangement is “channel 2” despite electrode 2’s position as the apical active electrode.

Thresholds were obtained with  $\sigma = 0$  (monopolar) and  $\sigma = 0.9$  (focused) electrode configurations. The sigma value for focused thresholds was chosen because it allowed for the presentation of current levels that were perceptible but below voltage compliance limits. These limits, defined as the maximum voltage supported by the device divided by the impedance of that electrode channel, were calculated for each subject at the beginning of the test session. For one participant (P08), however, a sigma value of 0.8 was used because this subject’s thresholds with  $\sigma = 0.9$  could not be obtained without reaching voltage compliance limits. As the sigma value affects threshold levels, sensitivity analyses were performed by removing this subject’s data. The overall findings of this study were unaffected.

Stimuli were biphasic, charge-balanced pulse trains with the cathodic phase leading. Pulses were 97 microseconds ( $\mu\text{s}$ ) in duration and were presented at a rate of 997.9 pulses per second. Each pulse train was 200.4 milliseconds (ms) long. The Bionic Ear Data Collection System (BEDCS version 1.18.315; Advanced Bionics, LLC) and custom software in MATLAB (The MathWorks, Inc.) were used to present stimuli and record participant responses.

MCL was first determined behaviorally for each channel using the Advanced Bionics clinical loudness scale (Advanced Bionics, LLC). Current level was increased manually until subjects reported a loudness rating of “6” (“Most Comfortable”). MCL was then set as the upper limit for stimulation, ensuring that presented current levels for each participant would not exceed comfortable levels.

For each electrode configuration, participants performed threshold sweeps of electrodes 2 through 15 in which  $\alpha$  was changed in steps of 0.1. Thresholds for electrodes 1 and 16 were not acquired because the focused electrode configuration necessitates flanking electrodes. Participants performed one forward sweep (from apical to basal electrodes) and one backward sweep (from basal to apical electrodes). Each sweep utilized a Békésy-like tracking procedure (Békésy, 1947): the listener was instructed to hold down the spacebar when they could perceive the signal and to release the spacebar when they could no longer perceive it. Each electrode and alpha step is called a “channel.” For example, electrode 6 with  $\alpha = 0.4$  is channel 6.4. The detection thresholds from the forward and backward sweep(s) utilized in this study were those averaged for each non-steered integer electrode channel; for example, channel 6.0.

Current level requirements varied greatly between the electrode configurations as well as subjects’ individual electrode channels, and thus the logarithmic decibel scale was used. This scale is considered to be more suitable than a linear scale for reporting such values.

A one-way multivariate analysis of covariance (MANCOVA) was conducted to compare threshold profiles of early-implanted children and late-implanted adults. Focused and monopolar thresholds across the electrode array were entered as dependent variables and “Group” (children or adults) as the independent variable. “Electrode” was included as a covariate to control for site-specific effects in threshold profiles. Correction for multiple comparisons (for the two ANOVAS within the analysis) was performed using a Bonferroni adjustment (adjusted  $\alpha = 0.025$ ).

### Channel-to-channel Variability

MATLAB (The MathWorks, Inc.) was utilized to assess the variation in monopolar and focused thresholds across the electrode array. The standard deviation of the signed differences in thresholds between each adjacent electrode was calculated from the threshold profiles of each participant. Use of the standard deviation instead of the mean difference between channels results in a measurement of the local variability, as opposed to the absolute magnitude of variability of all channels (e.g., Pfingst & Xu, 2005). Calculations were performed separately for thresholds from each electrode configuration. A one-way multivariate analysis of variance (MANOVA) was performed to compare threshold variability values between children and adults. Channel-to-channel variability of focused and monopolar thresholds were dependent variables, with “Group” as the independent variable. Correction for multiple comparisons was performed at  $\alpha = 0.025$ .

### Dynamic Range

The differences between MCL and threshold in dB for each electrode were calculated for each participant for both the monopolar and the focused electrode configurations. For a subset of channels for one child and for several adult subjects (P08, S29, S38, S41, S45, S50, S55), voltage compliance limits of the device were reached prior to reaching MCL. The level of current at compliance limit was not the true MCL for such channels, and therefore focused MCL and dynamic range values for those channels were not included in the analyses.

A one-way MANCOVA was conducted to compare dynamic range between early-implanted children and late-implanted adults. Monopolar and focused dynamic range were dependent variables, “Group” was the independent variable of interest, and “Electrode” was included as a covariate to account for variance in dynamic range due to electrode site. Because MCL has previously been found to contribute to dynamic range to a greater extent than thresholds (Firszt et al., 2002; Robinson et al., 2012), an additional one-way MANCOVA testing a separate hypothesis was performed with focused and monopolar MCLs as the dependent variables. “Group” was set as the dependent variable to determine whether hearing history significantly affected MCLs obtained from each electrode configuration. “Electrode” was again included as a covariate to control for site-specific effects on MCL. Correction for multiple comparisons for each analysis was performed at  $\alpha = 0.025$ .

## Electrical Field Imaging (EFI)

Data were collected using BEDCS (version 1.18315; Advanced Bionics, LLC). The stimuli were biphasic monopolar pulses (100  $\mu$ s in duration, 50 or 100  $\mu$ A in amplitude) delivered at a rate of 16.6 per second. Ten pulses were presented consecutively on each electrode site, while the voltage on every electrode was recorded sequentially at a 56 kHz sampling rate.

The data were analyzed offline using MATLAB (The MathWorks, Inc). After averaging, signal amplitude at each recording electrode was calculated as half the difference between the positive and negative voltage excursions, then scaled to units of resistance. Following the approach of Vanpoucke et al. (2004), the 16x16 matrix of EFI impedances was transformed to solve a lumped parameter resistor network representing current flow along and out of the cochlea. Three types of resistance were of interest: longitudinal ( $R_{long}$ ; resistance to current flow from the electrode along the length of the cochlea), transverse ( $R_{trans}$ ; resistance to current flow out of the cochlea to the ground electrode), and total ( $R_{total}$ ; resistance from all current pathways out of the cochlea) resistances. The resistor components of the network were labeled  $R_{long}$  and  $R_{trans}$  for each electrode position. The values were estimated using least squares optimization, with a localized weighting scheme to improve fitting of the EFI profiles and a regularization constraint to impose some degree of smoothness on the  $R_{long}$  and  $R_{trans}$  values across electrodes. One additional resistor value was defined for each stimulating electrode, as the peak of the reconstructed EFI profile (based on the solution to the ladder network). This value was termed  $R_{total}$ , as it represented the total resistance encountered by the stimulation electrode, incorporating all possible current pathways in the resistor network model.

This analysis results in estimates for  $R_{long}$ ,  $R_{trans}$ , and  $R_{total}$  for each electrode of a subjects' array. A one-way MANCOVA was performed to examine potential differences in these intracochlear resistance values between children and adults.  $R_{trans}$ ,  $R_{long}$ , and  $R_{total}$  were dependent variables and "Group" was the independent variable. "Electrode" was included as a covariate to control for site-specific effects on intracochlear resistance. Correction for multiple comparisons was performed at  $\alpha = 0.017$ .

## RESULTS

Table 4-2. Means and standard deviations (SD) of values from each assessment.

<i>Assessment</i>	<u>Group</u>			
	<i>Children</i>		<i>Adults</i>	
	Mean	SD	Mean	SD
Focused Thresholds (dB/1 $\mu$ a)	43.0	6.8	45.6	4.2
Monopolar Thresholds (dB/1 $\mu$ a)	32.0	4.4	32.7	4.4
Focused Threshold Variability (dB/1 $\mu$ a)	2.5	0.9	2.6	1.0
Monopolar Threshold Variability (dB/1 $\mu$ a)	1.7	1.1	1.9	1.0
Focused Dynamic Range (dB/1 $\mu$ a)	10.2	5.6	10.6	3.7
Monopolar Dynamic Range (dB/1 $\mu$ a)	11.5	3.7	11.8	3.5
Focused MCL (dB/1 $\mu$ a)	52.4	5.2	55.7	3.9
Monopolar MCL (dB/1 $\mu$ a)	42.9	2.7	44.0	2.7
Longitudinal Resistance ( $\Omega$ )	644.00	895.0	515.00	544.0
Transverse Resistance ( $\Omega$ )	22,178	17,509	15,242	8,472
Total Resistance ( $\Omega$ )	1,433	696	1,219	544

Focused thresholds were higher than monopolar thresholds in early-implanted children as well as in late-implanted adults, consistent with previous studies conducted in adults with CIs (e.g., Bierer, 2007). However, on average, children's thresholds were observed to be lower than those of adults for both electrode configurations (see Figure 4.1). A difference of 2.6 dB was found for focused thresholds and 0.7 dB for monopolar thresholds. Uncorrected group means and standard deviations for focused and monopolar thresholds are shown in Table 4.2. Statistical comparison of thresholds between early-implanted children and late-implanted adults revealed a significant effect of group on auditory perception thresholds [ $F(2,300) = 9.5, p < 0.001$ , Wilk's  $\lambda = 0.941$ , partial  $\eta^2 = 0.059$ ] when controlling for electrode site. Focused thresholds were significantly different between the children and adults [ $F(1,301) = 16.1, p < 0.001$ , partial  $\eta^2 = 0.051$ ]. Significant differences were not observed between groups for monopolar thresholds [ $F(1,301) = 1.9, p = 0.17$ ,

partial  $\eta^2 = 0.006$ ]. No significant effects of electrode site on threshold profiles were found [ $F(2,300) = 2.5$ ,  $p = 0.08$ , Wilk's  $\lambda = 0.983$ ] suggesting that group means were not significantly driven by thresholds from specific electrodes.

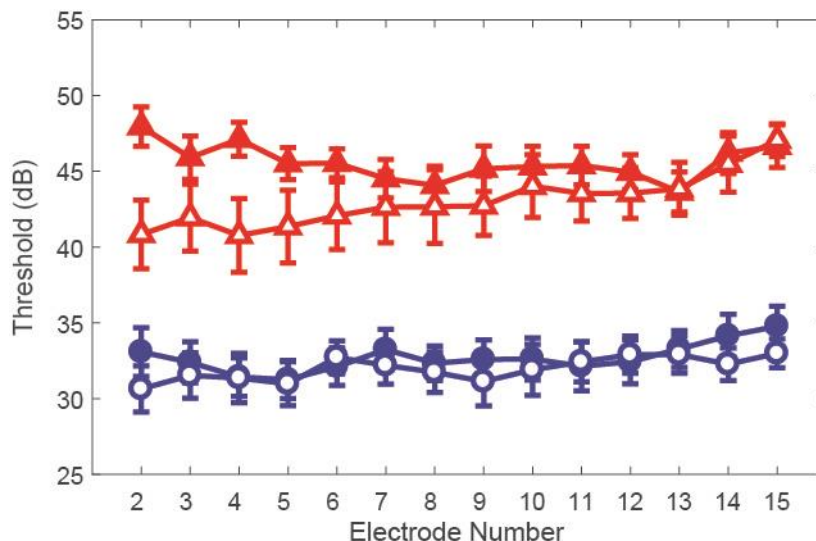


Figure 4-1. Auditory detection thresholds for early-implemented children and late-implemented adults across the electrode array.

Thresholds from the monopolar electrode configuration are indicated by blue circles and those from the focused configuration by red triangles. Adult data are represented by filled symbols and child data by open symbols. Error bars represent one standard error of the mean.

Channel-to-channel variability for the monopolar electrode configuration ranged from 0.29 to 3.91 dB in children and from 0.8 to 4.16 dB in adults (shown in Figure 4.2). With the focused electrode configuration, threshold variability ranged from 0.3 to 3.43 dB in children and from 1.10 to 4.12 dB in adults. Average channel-to-channel variability was similar between groups and was larger for focused compared to monopolar thresholds (means and standard deviations shown in Table 4.2). Statistical examination of threshold variance between early-implemented children and late-implemented adults revealed no significant effect of group [ $F(2,19) = 0.6$ ,  $p = 0.943$ , Wilk's  $\lambda = 0.994$ ]. Channel-to-channel variability did not differ significantly between children and adults for either electrode configuration [focused:  $F(1,20) = 0.07$ ,  $p = 0.80$ ; monopolar:  $F(1,20) = 0.10$ ,  $p = 0.76$ ].

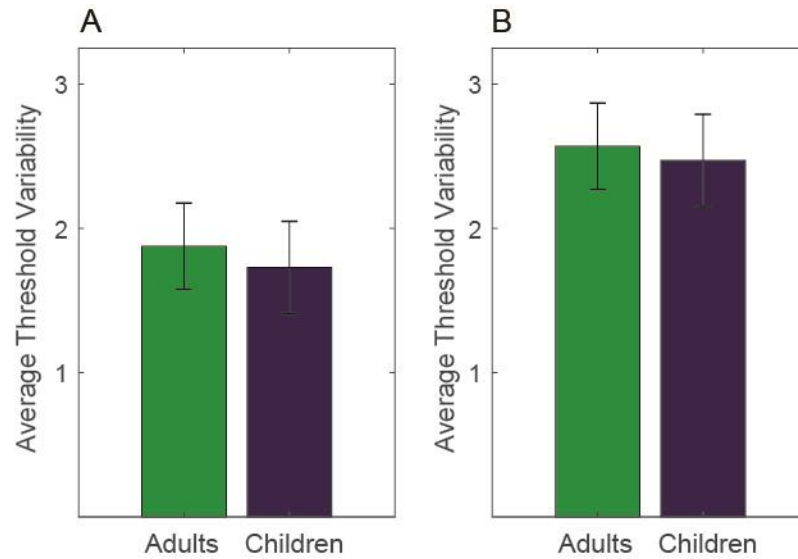


Figure 4-2. Average channel-to-channel variability for early-implemented children and late-implemented adults.

(A) Variability of monopolar thresholds. (B) Variability of focused thresholds. Adult data are in green and child data are in purple. Error bars represent one standard error of the mean.

Similarly, as shown in Figure 4.3, there was no significant effect of group on dynamic range, controlling for electrode site [ $F(2,259) = 0.30, p = 0.74, \text{Wilk's } \lambda = 0.998$ ]. Values were very similar between groups (see Table 4.2). The effect of electrode site on dynamic range was also not significant [ $F(2,249) = 3.9, p = 0.051, \text{Wilk's } \lambda = 0.977$ ].

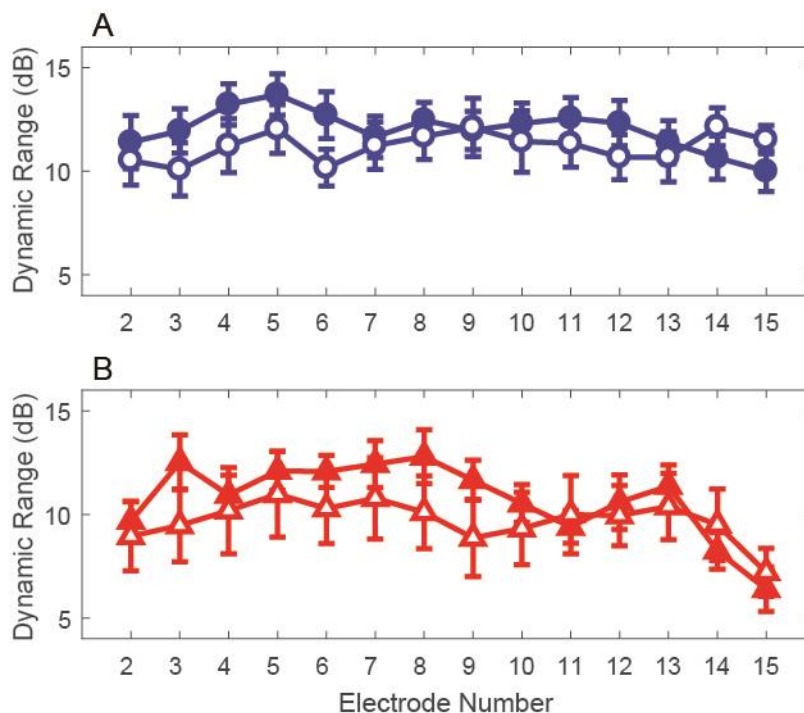


Figure 4-3. Dynamic range for early-implemented children and late-implemented adults across the electrode array.

(A) Dynamic range from the monopolar electrode configuration. (B) Dynamic range from the focused electrode configuration. Adult data are represented by filled symbols and child data by open symbols. Error bars represent one standard error of the mean.

While dynamic range itself did not significantly differ between early-implemented children and late-implemented adults, a significant effect of group was found for the upper limit of the dynamic range [ $F(2,259) = 15.9, p < 0.001$ , Wilk's  $\lambda = 0.891$ , partial  $\eta^2 = 0.109$ ]. Children had significantly lower MCLs for the focused [ $F(1,260) = 31.9, p < 0.001$ , partial  $\eta^2 = 0.1099$ ] and monopolar [ $F(1,260) = 11.7, p = 0.001$ , partial  $\eta^2 = 0.043$ ] electrode configurations. Uncorrected means and standard deviations are in Table 4.2. The analysis also indicated a significant effect of electrode site on MCLs [ $F(2,259) = 6.7, p = 0.001$ , Wilk's  $\lambda = 0.951$ , partial  $\eta^2 = 0.049$ ]. This finding was significant for MCLs obtained with the monopolar configuration only [monopolar:  $F(1,260) = 9.2, p = 0.003$ , partial  $\eta^2 = 0.034$ ; focused:  $F(1,260) = 0.01, p = 0.93$ ]. MCLs obtained with monopolar stimulation demonstrated a consistent increase from the most apical (electrode 2) to the most basal (electrode 15) channel. This effect was greater for the children than for the adults (see Figure 4.4).

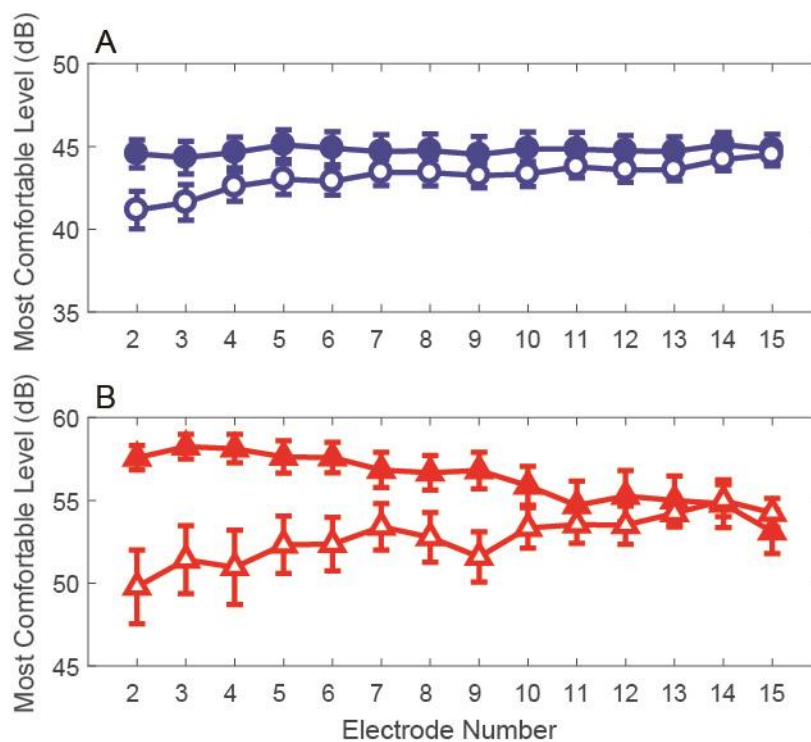


Figure 4-4. Most Comfortable Levels for early-implemented children and late-implemented adults across the electrode array.

(A) Most Comfortable Levels from the monopolar electrode configuration. (B) Most Comfortable Level from the focused electrode configuration. Adult data are represented by filled symbols and child data by open symbols. Error bars represent one standard error of the mean.

Analysis of intracochlear resistance values between early-implemented children and late-implemented adults revealed a significant effect of group [ $F(3,323) = 9.4, p < 0.001$ , Wilk's  $\lambda = 0.920$ , partial  $\eta^2 = 0.080$ ] as well as electrode site [ $F(3,323) = 10.4, p < 0.001$ , Wilk's  $\lambda = 0.912$ , partial  $\eta^2 = 0.088$ ] after correcting for multiple comparisons ( $\alpha = 0.017$ ). Children demonstrated significantly higher  $R_{trans}$  [ $F(1,325) = 20.9, p < 0.001$ , partial  $\eta^2 = 0.060$ ] and  $R_{total}$  [ $F(1,325) = 9.9, p = 0.002$ , partial  $\eta^2 = 0.030$ ] compared to adults. No significant differences between groups were observed for  $R_{long}$  [ $F(1,325) = 2.7, p = 0.097$ ], although on average, children did exhibit higher  $R_{long}$  values. This non-significant result is likely due to the large standard deviations in  $R_{long}$  values of children and adults (see Table 4.2 and Figure 4.5). Electrode site was found to significantly affect  $R_{total}$  [ $F(1,325) = 6.6, p = 0.010$ , partial  $\eta^2 = 0.020$ ] and  $R_{long}$  [ $F(1,325) = 19.0, p < 0.001$ , partial  $\eta^2 = 0.055$ ]. The location of the electrode in the array influenced the mean values for both groups: on average, both types of resistance were greater for basal electrodes

compared to apical electrodes. Electrode site was not found to significantly influence  $R_{trans}$  [ $F(1,325) = 2.2, p = 0.135$ ].

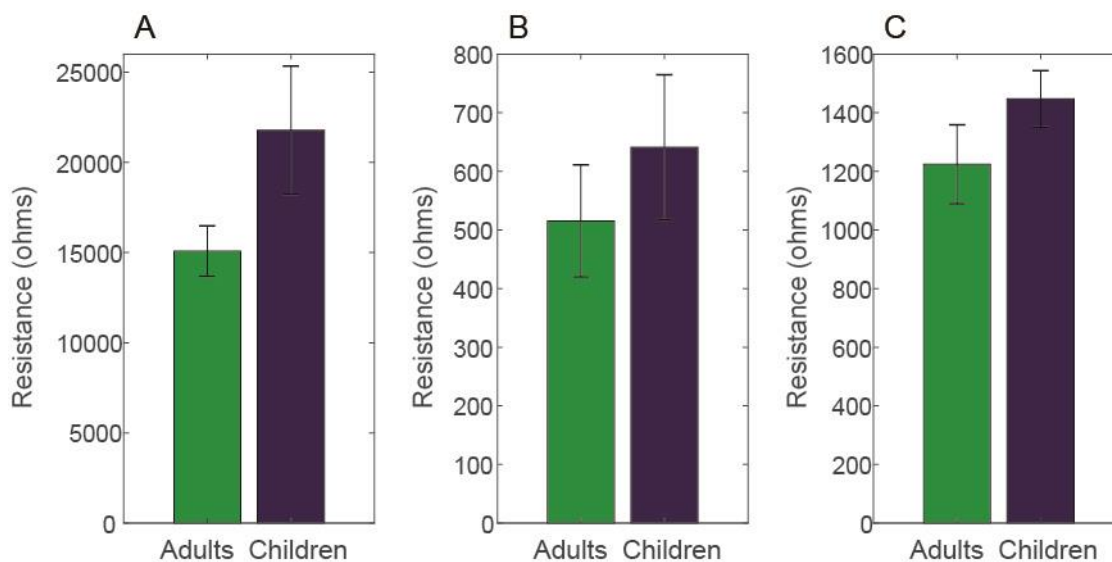


Figure 4-5. Average intracochlear resistance values for early-implemented children and late-implemented adults.

(A) Transverse resistance. (B) Longitudinal resistance. (C) Total resistance. Adult data are in green and child data are in purple. Error bars represent one standard error of the mean.

## DISCUSSION

This study compared factors related to the ENI in early-implemented children and late-implemented adults. It was hypothesized that the divergent hearing histories of these groups would differentially affect the ENI. In correspondence with this hypothesis, auditory detection thresholds for the focused and monopolar electrode configuration were found to be significantly lower in early-implemented children compared to late-implemented adults. The difference between groups was much larger for the focused thresholds, possibly because focused stimulation is more sensitive to the quality of the ENI (Bierer, 2007). It is unlikely that systematic between-group differences in electrode position contributed to this finding, as the size of the cochlea is adult-like at birth and all participants had the same electrode array. A more probable explanation is that this result stems from less neuronal degeneration in early-implemented children compared to late-implemented adults.

The lower thresholds observed in the children may indicate that they have healthier ENIs than do the adults.

Individuals with severe hearing loss have significantly lower spiral ganglion counts than their hearing counterparts (e.g., Otte et al., 1978). Further evidence from studies in cats indicates that spiral ganglion death occurs gradually in the absence of auditory stimulation (e.g., Leake & Hradek, 1998). While the progression of spiral ganglion loss is believed to be slower in humans, these findings suggest that relatively long periods of deafness (which occurred for most adults in this study) may have reduced spiral ganglion survival. Normal aging has also been found to decrease spiral ganglion counts (Makary et al., 2011). However, electrical stimulation of spiral ganglion neurons in deafened animal models has been observed to promote survival of these neurons (Leake et al., 1991). It seems likely that early-implanted children would have greater neural survival (resulting from less neuronal degeneration) compared to late-implanted adults. The results of this investigation add further evidence to this premise.

Despite having lower auditory perception thresholds, pediatric CI users did not have larger dynamic ranges relative to adults because their MCLs were also significantly lower. The range of perceptible and comfortable current seem to be shifted to lower levels in early-implanted children. This further supports the hypothesis that children have healthier and/or a greater density of spiral ganglion neurons. In fact, MCLs of particular channels for only one child (P08) but those for six adult subjects (S29, S38, S41, S45, S50, S55) could not be measured without reaching the voltage compliance limits of the CI, suggesting that a long duration or progression of hearing loss can influence the amount of current required to reach both threshold and MCL. Future studies utilizing in vivo behavioral and objective measures of neuronal health and density (e.g., Hughes et al., 2018; Schwartz-Leyzac & Pfungst, 2016; Zhou et al., 2015) are necessary to further elucidate neuronal status in children and adults with CIs.

While the results described thus far may indicate overall healthier ENIs in children, early-implanted children had significantly higher intracochlear resistance compared to late-implanted adults. EFI is a measure of voltage flow through the cochlea from each electrode and provides information about the physical characteristics of the fluid-filled space surrounding the electrode array. This measure is not the same as electrode impedance; rather, it characterizes the resistivity of the area around the electrodes. Previous research has demonstrated that intracochlear resistance values from EFI were positively related to the distance of electrodes from the inner wall of the

cochlea and were sensitive to the scalar location of the electrode array (Bierer et al., 2015b). Yet, there is no evidence to suggest that these variables would differ between children and adults. While both groups may have variation in electrode position within each of their CI electrode arrays, it is highly unlikely that all electrodes of pediatric CI users' implants would be further away from the inner wall or in a different scala compared to adults, resulting in greater current resistance levels. This is particularly evident because children's thresholds were found to be lower than those of the adult group: increased distance of the electrode from the target neurons would result in increased threshold for that electrode as well. It is also unlikely that adults would have a significantly higher number of electrodes in scala tympani, which has previously been found to result in lower  $R_{trans}$  levels (Bierer et al., 2015b). Voltage is sensitive to the size of the electrode array and the cochlea (Vanpoucke et al., 2004), but the cochlea is adult-sized at birth (e.g., Pelliccia et al., 2014) and all participants in this study used the same type of electrode array without a positioner. Therefore, these factors should not have contributed to the EFI results.

Instead, the higher intracochlear resistances in early-implanted children may be due to greater levels of bone and tissue growth, which can increase resistance to electrical current flow (Spelman et al., 1982). Ossification and fibrous tissue growth have been found to occur after cochlear implantation in adults due to trauma from electrode array insertion (Li et al., 2007; Seyyedi & Nadol, 2014). However, children could have greater cochlear ossification and new tissue growth around the electrode array after CI surgery compared to adults because temporal bone growth is dynamic throughout childhood (Dahm et al., 1993). While the finding of increased intracochlear resistance in children may seem contradictory to the lower auditory perceptual thresholds observed in this group, these results may be consistent with Ohm's Law ( $Voltage = Impedance * Resistance$ ). Increased intracochlear resistance ( $R$ ) in children would lead to less current flow ( $I$ ) along the cochlea and away from the neural targets to equal the same driving voltage gradient ( $V$ ) at the spiral ganglion as in the late-implanted adults. Less current ( $I$ ) would therefore be required to achieve threshold in individuals with higher intracochlear resistance ( $R$ ). Yet, continued study is needed to better understand how the complex geometry of the cochlea affects the spread of excitation with CI stimulation.

Additionally, EFI is a measure of current flow near the implanted electrodes, not the spiral ganglion, and there are other factors that can influence the voltage gradient at a single neuron. Therefore, an exploratory regression analysis was performed to investigate whether intracochlear

resistance levels of the participants in this study affected auditory detection thresholds. A linear mixed-effects model was performed in R using the lme4 package (Bates et al., 2015). Thresholds across the electrode array were specified as the dependent variable, with group membership, measures of intracochlear resistance, and type of electrode configuration for obtaining thresholds as the independent variables. “Subject” was included as a random intercept in the model. Results indicated a significant relation of Rlong, which reflects global current flow within the cochlea, to thresholds of participants ( $\beta = 0.49$ , standard error = 0.14,  $p < 0.001$ ). A significant interaction between “Group” and “Stimulus Type” was also found ( $\beta = 1.9$ , standard error = 0.48,  $p < 0.001$ ), demonstrating that thresholds were significantly different between groups even when intracochlear resistance was accounted for. Altogether, these results indicate that hearing history can alter physical characteristics of the cochlea, which in turn influences auditory signal transmission and perception.

Yet, no significant differences in channel-to-channel variability were observed between groups. Consistent with previous studies in adults (e.g., Bierer, 2007; Bierer et al., 2015a; Long et al., 2014), both pediatric and adult CI users in this study exhibited relatively uniform monopolar thresholds, but variable focused thresholds, across the electrode array. These results demonstrate that children exhibited similar variation in ENI quality as that observed in the adult group. Indeed, examinations of individual participant data revealed that pediatric participants had regions of relatively elevated focused thresholds with their thresholds profiles, akin to late-implanted adults. Although early implantation has been associated with better outcomes for children (e.g., Wang et al., 2008), this result demonstrates that even early-implanted children have electrode-neuron interfaces that are of poorer quality than others within their implant.

Given the results from the present study, early-implanted children may benefit from updated CI programming recommendations. Several prior studies have shown that experimental programming processing strategies can improve speech perception in adults with CIs (e.g., Arenberg et al., 2018; Garadat et al., 2013). However, all but one recent study (e.g., Noble et al., 2016) assessing novel CI programming strategies have been conducted in adults (Azadpour & Smith, 2016; Koning & Wouters, 2016; Müller et al., 2018; Nogueira et al., 2016; Padilla et al., 2017), most of whom were postlingually-deafened and late-implanted. As results from this investigation demonstrated fundamental differences in the ENI between early-implanted children and late-implanted adults, and because the ENI is a critical component for auditory signal

transmission through the CI, programming strategies that may enhance outcomes for late-implanted adults with CIs may not be as advantageous for early-implanted children. Similarly, children could potentially benefit from processing strategies which have had minimal success in some adults. For example, if early-implanted children have greater neuronal densities than do late-implanted adults, processing strategies using focused stimulation (to reduce current interaction between channels) may produce greater improvements in children's speech identification scores compared to those previously found in adults (Berenstein et al., 2008; Srinivasan et al., 2013). In the current study, children were also found to exhibit regions of relatively poorer electrode-neuron interfaces within their electrode array; techniques which have been used in adults to improve transmission of auditory signals through suboptimal electrode-neuron interfaces could therefore be applied in children (e.g., Bierer and Litvak, 2016). Future studies on optimization of programming strategies for early-implanted children are warranted.

While this study found dissimilarities in the ENI between two extreme populations with CIs, a similar investigation could be extended to other groups of CI users, such as children who had progressive hearing loss and adults who lost their hearing as children. Investigations of the ENI in various populations of CI users could provide additional information about the effect of hearing history on auditory signal transmission and auditory perception of individuals with CIs.

## CONCLUSION

Auditory detection thresholds and the upper limit of the dynamic range were found to be lower for early-implanted children compared to late-implanted adults, suggesting that children may have overall healthier electrode-neuron interfaces. The results of this study may be explained by differences in neuronal survival, etiology of deafness, or other factors related to intracochlear physical characteristics in these two populations. Still, channel-to-channel threshold variability was not found to differ significantly between groups, indicating that early-implanted children exhibit similar variation in ENI quality across the electrode array as adults. Future studies should investigate the physiological mechanisms that contribute to ENI quality in early-implanted children and late-implanted adults with CIs.

## CHAPTER 5 : SPEECH-BASED SPECTRAL RESOLUTION BY AGE AND HEARING MODALITY

### ABSTRACT

**Purpose:** Phonetic cue weighting assesses the weight an individual places on a cue when categorizing speech sounds. Such tests can be used to measure one's reliance on the various frequency and temporal components of an auditory signal. This study utilized a cue-weighting test of speech-based spectral resolution, with a continuum of speech sounds ranging from /ba/ to /da/. Two cues related to the place of articulation contrast were manipulated: a fine-resolution cue, formant transitions, and a coarse-resolution cue, spectral tilt. Higher reliance on the formant cue compared to the spectral tilt cue indicates likelihood of better spectral resolving capabilities. In prior studies, normal hearing (NH) adults relied almost exclusively on the formant cue, but CI listeners, who have poor spectral resolution, relied more on the spectral tilt cue to categorize the sounds. However, it was not yet known whether children would demonstrate these same cue-weighting patterns: NH children have been found to place less perceptual weight on certain phonemic cues compared to NH adults, and early-implanted children with CIs have distinct hearing histories compared to late-implanted adults that could affect spectral resolution and perception of speech sounds. Thus, the aim of the current study was to investigate the cue-weighting patterns related to spectral resolution of children with NH and with CIs.

**Methods:** Children and adults with these two hearing modalities performed the test of speech-based spectral resolution. For participants with CIs, phonetic cue usage was compared to performance on a closed-set test of vowel identification as well as a standardized measure of spectral ripple discrimination, the Spectral-temporally Modulated Ripple test (SMRT).

**Results:** NH participants relied much more on the formant cue and less on the spectral tilt cue compared to adults with CIs, similar to results of previous studies in adults. Children demonstrated cue-weighting patterns comparable to those of adults with the same hearing modality, but placed less perceptual weight on phonemic cues. This could indicate immature perceptual abilities of these groups. All NH participants and some adults with CIs exhibited cue trading but children with CIs used both cues to a similar extent. These results suggest that prelingually-deafened, early-implanted children differ from postlingually-deafened, later implanted adults in their integration

of acoustic cues for speech sounds. In addition, formant cue usage was positively related to SMRT thresholds of children with CIs but not adults, suggesting that both measures might tap into common auditory processing mechanisms for early-implanted children. Greater ability to use the formant cue in categorization was also significantly correlated with higher vowel recognition scores for early-implanted children. Further, bilaterally-implanted CI users exhibited distinct cue-weighting patterns with individual CIs, demonstrating that factors related to each CI could contribute to perceptual cue weighting on this test.

**Conclusion:** Results of this study indicated that hearing modality and age can both influence cue-weighting patterns from the test of speech-based spectral resolution. Children may exhibit perceptual weighting that is more akin to those of adults with increasing age and/or time with the CI. Future directions of this work include examinations of variables that may contribute to cue-weighting strategies of child and adult CI users.

## INTRODUCTION

Speech contains many cues in the frequency domain, and thus the ability to resolve spectral cues within the speech signal is important for identifying that sound. Degrading spectral information has been found to decrease normal hearing (NH) listeners' ability to accurately identify speech (e.g., ter Keurs et al., 2002). Further, phonemes with similar spectral cues are those most often confused when spectral information is lost or distorted (DiNino et al., 2016). Cochlear implants (CIs) are highly successful auditory prostheses, yet speech perception performance varies greatly among both adults (Holden et al., 2013) and children (Wang et al., 2008) with these devices. All CI users have limited spectral resolving capabilities compared to NH listeners due to the small number of frequency channels in the implant (Friesen et al., 2001; Henry et al., 2005). Yet, spectral resolution can be further decreased by additional factors, such as variables related to the electrode-neuron interface (Bierer, 2007); individual differences in the spectral resolution of CI users therefore likely contributes to the variance in speech identification performance of these individuals.

Several assessments have been utilized extensively to probe the spectral resolution of individuals with CIs. These include spectral ripple discrimination (Henry and Turner, 2003; Supin et al., 1994) electrode pitch ranking (Nelson et al., 2005), and psychophysical tuning curves (Bierer & Faulkner, 2010). Performance on such tests have been consistently found to relate to speech recognition scores of CI listeners (e.g. Henry et al., 2005; Kenway et al., 2015; Won et al., 2007), but the exact mechanism underlying this relationship is unclear. The stimuli in the aforementioned tests do not contain linguistic elements, and even the broadband noise stimuli of widely-used spectral ripple tests do not represent the important spectral modulations of speech (Saoji et al., 2009; Singh & Theunissen, 2003). In addition, assessments of spectral resolution measure discrimination of spectral content, whereas speech identification also involves categorization of speech sounds based on their representative phonemic characteristics (see Holt & Lotto, 2010 for review; Liberman et al., 1957). Speech recognition tests themselves are also not ideal for assessing one's true perception of speech sounds through the implant: many factors, such as lexical knowledge, context effects, and working memory capacity, can influence performance on these measures. An assessment of spectral resolution that is analogous to the process of speech perception may thus better predict the speech recognition abilities of CI users.

Accordingly, Winn and Litovsky (2015) recently developed a linguistic test of spectral resolution in which a listener utilizes phonetic cues to categorize /ba/ and /da/ stimuli. These sounds are contrasted by place of articulation, and perception of this consonant feature is greatly affected by reduced frequency resolution (Bilger & Wang, 1986; Munson et al., 2002). Therefore, categorizing /ba/ and /da/ is particularly difficult for individuals with hearing loss and with CIs. Yet, even with broadened auditory filters, individuals with hearing impairment can somewhat resolve the place of articulation feature, although they are likely using different phonetic cues than NH listeners are. Winn and Litovsky's test of speech-based spectral resolution takes advantage of this redundancy of speech: co-occurring cues exist for any contrast, and individuals place greater perceptual weight on some cues than others. In addition, a listener can compensate for the decrement of one cue by utilizing a different cue (Repp, 1982), and this is termed "cue trading." Thus, two individuals could perform similarly on a test of speech recognition but utilize different phonetic cues. Identifying the cues that a CI user is relying on to categorize speech sounds provides a greater amount of information about that individual's speech recognition abilities than does a classic test of speech perception.

To create stimuli for this test of speech-based spectral resolution, naturally-spoken /ba/ and /da/ stimuli were orthogonally manipulated on a continuum of two cues that are present in the place of articulation contrast of /ba/ and /da/: formant transitions and spectral tilt (the relative balance of high and low frequency energy in the spectrum). Resolving formant cues requires a high level of spectral resolution, but spectral tilt is a coarse-resolution cue that is accessible even with poor spectral resolution (Alexander & Kluender, 2009). NH listeners have been found to utilize the onset frequency of the second formant (F2) to perceive place of articulation but weight spectral tilt higher when formant information in the contrast was degraded or missing (e.g., Alexander & Kluender, 2008). Accordingly, individuals with good spectral resolution should be able to utilize the formant cue to categorize the sounds, whereas those with poor spectral resolution will not have access to the formant cue and should place greater perceptual weight on the spectral tilt cue. Additionally, unlike traditional tests of spectral resolution, this assessment does not solely test perception and discrimination of spectrally-related cues; it evaluates one's ability to actually use a spectral cue to categorize speech sounds.

In a prior study, Winn and Litovsky (2015) utilized this test with adults with NH and with CIs. Both groups were presented with unprocessed stimuli, but NH listeners also heard stimuli that

were spectrally degraded with vocoder processing. In the unprocessed condition, Winn and Litovsky found that NH adults utilized the formant cue to a greater extent than did adults with CIs. Vocoding the stimuli reduced NH listeners' weighting of the formant cue to the level of CI users. Conversely, CI listeners placed greater perceptual weight on the spectral tilt cue compared to NH adults in the unprocessed condition, but use of this cue was similar between CI listeners and NH adults listening to vocoded stimuli. These results indicate that this test is valid for assessing spectral resolution: listeners in both groups were able to perceive place of articulation to categorize the /ba-/da/ contrast but utilized different cues to do so, depending on the level of spectral degradation as a result of the CI or of vocoder processing.

Winn and Litovsky (2015) also found that greater use of the formant cue by participants with CIs was significantly related to higher monosyllabic word recognition scores, demonstrating that results obtained from this test were able to predict speech recognition performance of adult CI users. In addition, both groups of listeners demonstrated cue trading: NH adults placed high perceptual weight on the formant transition and low perceptual weight on spectral tilt. Adults with CIs weighted spectral tilt highly and the formant cue to a much lesser extent. However, it was not yet known if children with NH or with CIs would exhibit the same pattern of phonetic cue perception as has been found in adults, or if results on the test of speech-based spectral resolution would predict speech identification abilities of pediatric CI users.

Previous studies have found that NH children's patterns of cue weighting can differ from those of NH adults. Children exhibit difficulty when categorizing stimuli with ambiguous phonemic boundaries and seem to require greater salience of cues compared to adults (Hazan & Barrett, 2000; Morrongiello et al., 1984). Many studies which have observed divergent cue weighting patterns for place of articulation found those differences between very young children and adults (e.g., Mayo & Turk, 2004), but identification of consonants in challenging listening conditions has been found to be immature until age 15 (Johnson, 2000). Thus, the age at which consonant categorization is adult-like is unknown. Further, NH children's spectral resolving capabilities are poorer than those of NH adults (Peter et al., 2014). Young NH children have more difficulty accurately identifying vocoded speech sounds (Eisenberg et al., 2000) and perceiving speech in noise (Fallon et al., 2000), compared to older children and adults due to immaturity of sensory processing. Therefore, even NH children may weight a fine-resolution formant cue less than adults with NH.

Differences in phonetic cue weighting are also likely to be found between early-implanted children and late-implanted adults with CIs. In addition to potential developmental effects that could result in distinct patterns of perceptual weights of these groups, prior research has found differences in auditory perception related to the electrode-neuron interface between early-implanted children and late-implanted adults (DiNino et al., 2017). Results from that study suggested that early-implanted children may have healthier electrode-neuron interfaces compared to late-implanted adults; therefore, auditory transmission, and perhaps spectral resolution, through the implant may be better for that group of children. Early-implanted children with CIs may therefore weight a formant cue to a greater extent when categorizing /ba/ and /da/ compared to late-implanted adults.

In addition, early-implanted children and late-implanted adults learned language with different hearing modalities. Such children were born deaf or lost their hearing early in life, and thus learned language with the CI as their sole auditory input. Yet, many late-implanted adults acquired language with acoustic hearing prior to severe to profound hearing loss and receipt of a CI. Categorical perception and cue usage may thus differ between these groups: late-implanted adults often change their auditory perceptual strategies after receipt of the implant (e.g., Winn et al., 2012), but early-implanted children acquire their perceptual strategies with CI input.

The goal of this study was to examine the influence of age group (child or adult) and hearing modality (acoustic hearing or CI) on speech-based spectral resolution, as assessed by Winn and Litovsky's (2015) cue-weighting task. It was expected that adults with NH and with CIs would demonstrate the same pattern of results observed by Winn and Litovsky: NH adults should weight the formant cue more heavily than adults with CIs, whereas adults CI users should place greater perceptual weight on the spectral tilt cue compared to NH individuals. In addition, due to decreased auditory processing abilities in children compared to adults, it was hypothesized that NH children would be unable to utilize the fine-resolution formant cue at the same level as NH adults. Further, children with CIs should place lower perceptual weight on the formant cue compared to both groups of NH participants due to the limited spectral resolution of the implant. However, early-implanted children with CIs may utilize the formant cue to a greater extent than late-implanted adults with CIs because of differences in hearing history between these groups. In addition to examining cue-weighting patterns of children and adults with CIs, results of the test of speech-

based spectral resolution were compared to these individuals' performance on vowel identification and spectral ripple discrimination tasks.

## METHODS

### Participants

Two groups of CI users participated in this study: ten prelingually-deafened children (mean age at first visit = 13.8 years) and thirteen postlingually-deafened adults (mean age at first visit = 67.4 years). Every child received their first implant prior to 5 years of age (mean age at first CI = 1.9 years). Every adult received a CI later in life (mean age at first CI = 57.8). All ten of the children were bilaterally-implanted and each ear was tested separately, for a total of 20 ears. Seven of the adults were bilaterally-implanted, and therefore 20 adult ears were also tested for the adult group. All participants with CIs used oral communication and were native speakers of American English. Most used Advanced Bionics HiRes90K devices; five of the adults with CIs (SC01, SC03, SC05, SC06, and SC07) utilized Cochlear devices. CI participant demographics are shown in Table 5.1. These individuals performed tests for several studies during the same visit, and as a result the assessments for this investigation were performed over two to three visits. The order of the tests that participants performed was randomized.

Table 5-1. Demographics of participants with cochlear implants.

EVA: enlarged vestibular aqueduct. DFNB1: genetic nonsyndromic hearing loss.

<u>Children</u>						<u>Adults</u>					
Subject:	Ear:	Etiology of Deafness:	Age at First Testing:	Age at Implantation (yrs):	Duration of Deafness (yrs):	Subject:	Ear :	Etiology of Deafness	Age at First Testing:	Age at Implantation (yrs):	Duration of Deafness (yrs):
P01	R	Unknown	15.7	2.3	1.5	S22	R	Hereditary	76.9	66.7	11.8
	L			12.1	11.3	S23/S36	L	Unknown	72.1	62.0	3.9
P02	R	EVA	11.8	1.1	1.0		R			64.5	6.5
	L			3.1	3.0	S29	L	Unknown	87.8	76.8	29.8
P03	R	Unknown	12.9	1.4	1.1		R			85.7	38.7
	L			5.6	5.3	S39/S30	R	Hereditary	53.4	30.1	9.1
P04	R	Unknown	13.2	1.5	0.8		L			40.1	19.1
	L			4.5	3.8	S45	R	Hereditary	65.3	54.0	32.0
P06	R	Unknown	17.2	4.3	2.5	S47/S51	R	Unknown	39.1	36.4	10.3
	L			11.0	9.1		L			38.4	12.4
P07	R	Unknown	13.3	1.9	0.4	S50	R	Measles	76.5	61.1	41.1
	L			4.9	3.5	S52	R	Unknown	70.9	66.0	21.1
P09	L	Unknown	13.5	2.6	1.3	SC01	R	Unknown	64.1	62.3	37.3
	R			3.9	2.7	SC03	L	Unknown	70.5	64.7	12.2
P10	L	DFNB1	13.3	1.1	0.9		R			66.0	13.5
	R			5.1	4.9	SC05	L	Hereditary	62.0	58.1	1.2
P11	R	DFNB1	13.3	1.4	1.2		R			59.9	3.0
	L			10.2	10.0	SC06	R	Unknown	71.0	66.2	49.4
P12	R	DFNB1	13.3	1.7	1.4	SC07	R	Hereditary	66.7	46.5	30.4
	L			10.2	10.0		L			59.5	43.5

Fourteen NH adults between the ages of 24 and 66 (mean age = 41.9 years) and thirty-seven NH children between the ages of 8 and 17 (mean age = 12.7 years) were also included in this study. The NH adult group consisted of eight “young adults” between the ages of 24 and 32 (mean age = 28.0 years) and six “older adults,” aged 53 to 66 years (mean age = 60.4) to control for potential effects of aging on cue weighting patterns. In addition, the older NH adults were closer in age to the adult participants with CIs. The entire NH adult was not composed of older adults because hearing abilities often decline with age, and thus NH individuals of older ages are not as common (and are more difficult to recruit) as young NH adults. A greater number of NH children than adults were recruited so that data could be acquired from at least three children of each 1-year age bin. This allowed for a representative sample of potential effects of development on weighting of phonetic cues. All NH participants completed a screening to verify hearing at 25 dB HL at 250, 500, 750, 1000, 2000, and 4000 Hz. NH young adults and children were additionally required to perceive 8000 Hz at 25 dB HL.

#### Assessments

All participants performed the test of speech-based spectral resolution, but only participants with CIs were tested on vowel identification and spectral ripple discrimination. Stimuli for all assessments were presented through a Crown D75 amplifier and an external A/D device (SIIF USB SoundWave 7.1) and were played through speakers in a double-walled sound-attenuating booth (IAC RE-243). Participants were seated in front of a computer screen at 90 degrees azimuth to the speaker. NH participants performed the task with both ears simultaneously. Participants with CIs were tested with each ear separately and turned off the CI or hearing aid in the contralateral ear during testing.

### ***Test of Speech-Based Spectral Resolution***

#### **Stimuli Creation**

Six naturally-spoken sounds from a male, native speaker of American English were utilized as stimuli for this test. These included the contrast of interest, /ba/ and /da/, as well two additional contrasts: /sha/ and /sa/, and /la/ and /ra/. All stimuli manipulations were performed in Praat (Boersma & Weenink, 2013).

The /ba/-/da/ contrast was created by orthogonal manipulation of the tokens on a continuum of the 1) second formant transition and 2) spectral tilt at the onset of the syllable. The formant

continuum was created by first downsampling the /ba/ token to 10,000 Hz and estimating 12 linear predictive coding (LPC) coefficients below 5,000 Hz. The sound was then inverse filtered by the LPC to remove formant peaks, creating a “source” stimulus that allowed for filtering by a different formant structure. As shown in Figure 5.1, the formant contours of the original /ba/ and /da/ tokens were extracted and six intermediate formant contours were created, for a total of eight formant steps from /ba/ to /da/. Each step of the stimuli was low-pass filtered at 3,500 Hz and added to the original /ba/ sound that had been high-pass filtered above 3,500 Hz. This restored the frequency energy above 10,000 Hz to the manipulated stimuli.

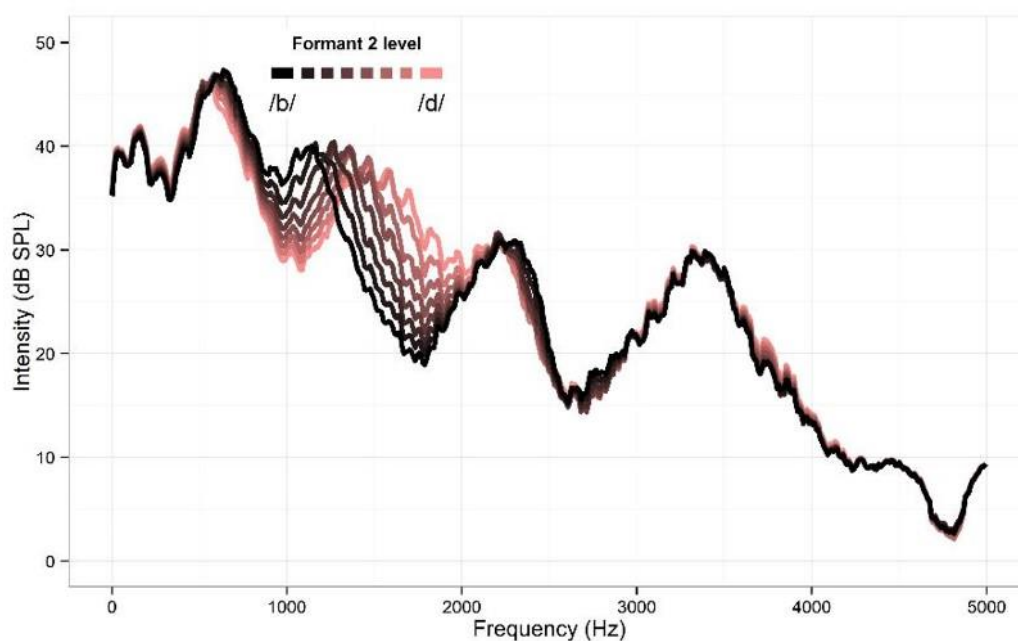


Figure 5-1. Formant transition manipulation for /ba/-/da/ stimuli, from Winn and Litovsky (2015).

Lighter-colored lines are the most like /d/ and black lines are most like /b/.

After creation of the /ba/-/da/ formant continuum, the spectral tilt (from the first to fourth formant at the syllable onset) of the stimuli was altered on a continuum of five steps within each formant step (see Figure 5.2). A filter which utilized logarithmic multiplication of the stimuli amplitude spectrum either amplified or attenuated frequency energy above 800 Hz to create varying linear slopes of spectral tilt. The formant continuum stimuli were then each multiplied by the five filters of shifted spectral tilt to create stimuli that contained one step of each cue. These

were added to the original /ba/ sound at the beginning of the consonant burst. The vowel nucleus in all stimuli were identical at 80 ms after the burst to the end of the stimulus.

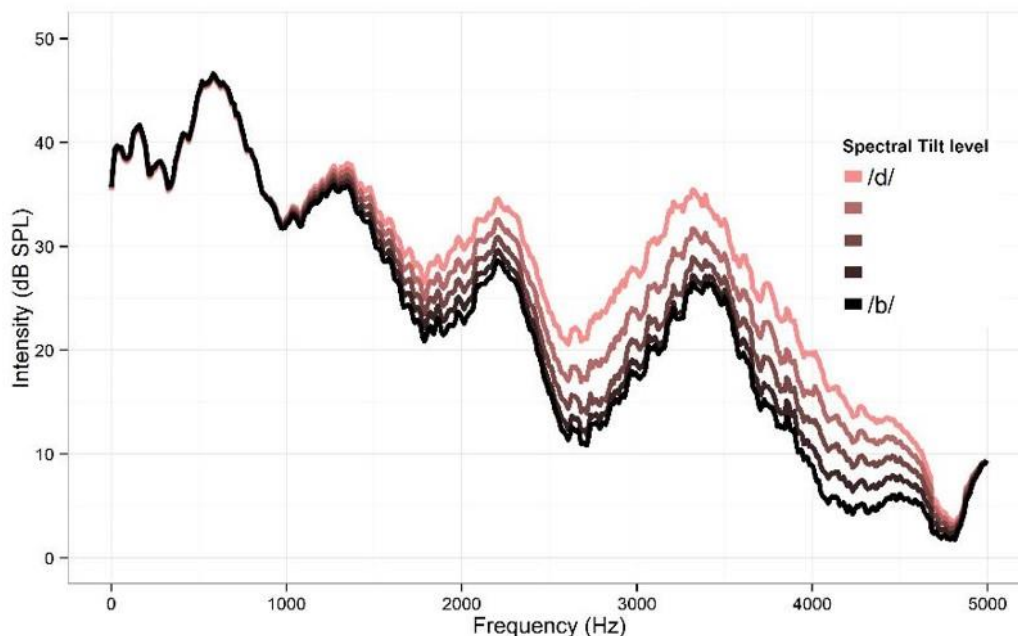


Figure 5-2. Spectral tilt manipulation for /ba/-/da/ stimuli with formant structure held constant, from Winn and Litovsky (2015).

Lighter-colored lines are the most like /d/ and black lines are most like /b/.

Stimuli with a diffuse-falling spectrum (decreasing spectral tilt and F2 frequency at the onset of the transition) are perceived acoustically as “ba,” whereas stimulus with a diffuse-rising spectrum (increasing spectral tilt and F2 frequency at the onset of the transition) are perceived acoustically as “da” (Stevens & Blumstein, 1978). Formant and spectral tilt steps between the /ba/ and /da/ stimulus endpoints were acoustically more ambiguous. The formant contrast between step 1 to step 8 was a range of about 1000 to 1800 Hz, whereas the range of spectral tilt was between about 800 to 6000 Hz. Therefore, individuals with good spectral resolution would be able to perceive the formant cue to categorize the speech sounds. Individuals with poorer spectral resolution would not be able to resolve the fine-resolution formant cue, and would instead place high perceptual weight on the spectral tilt cue.

The additional two contrasts were used as fillers for the test so that participants would not become more sensitive to the manipulation of the /ba/ and /da/ contrast. The /sha/ and /sa/ contrast

was created by gradually blending the sounds to form a seven-step continuum from /sha/ to /sa/. The /la/ and /ra/ contrast consisted solely of the original recorded signals.

### **Procedure**

Participants performed a one-interval, six-alternative forced-choice task of speech sound categorization. Subjects were seated in front of a computer screen that contained six boxes, each labelled with “ba,” “da,” “sha,” “sa,” “la,” or “ra.” Stimuli were presented at 65 dB SPL. After presentation of a sound, participants were asked to select the sound they heard by using a computer mouse to click the appropriate box. Each participant completed one practice run in which only the endpoint stimuli were presented, to familiarize the participant with the testing procedure. Participants then performed five test runs with one repetition of each /ba/-/da/ step of formant and spectral tilt per run. Data from the practice run were not included in the analysis. Participants were allowed to repeat the stimuli during all runs but were told to “make their best guess” from their first impression of the sound instead of repeating it, if possible.

### **Analysis**

Only responses to the /ba/-/da/ continuum were analyzed. Binomial logistic regressions were used to determine the perceptual weighting of the formant and spectral tilt cues for each group (NH adults, NH children, CI adults, and CI children), as well as for each individual NH participant and for each CI of participants with implants. Participant’s responses (“ba” or “da”) were converted to 0 or 1 and used as the dependent variable. The steps of formant and tilt that elicited each response comprised the independent variables. This analysis calculated slope coefficients for perceptual weighting of the formant and spectral tilt cues. Higher coefficients indicated greater use of the cue. Coefficients equal to zero indicated inability to utilize the cue, and coefficients less than zero indicated use of the cue in the wrong expected direction.

### ***Vowel Identification***

Children and adults with CIs performed a closed-set test of vowel identification. Ten vowels in /hVd/ context (/i/, /ɪ/, /eɪ/, /ɛ/, /æ/, /ɑ/, /u/, /ʊ/, /o/, /ʌ/) naturally spoken by a female talker were presented at 60 dB SPL. Stimuli were presented, and participants responses recorded, through List Player (version 2.2.11.52, Advanced Bionics, LLC). Participants were seated in front of a computer screen that contained ten boxes labelled with each of the possible response (i.e., “heed”,

“hid”, “hood,” etc.). Following presentation of a vowel sound, participants were asked to select the box that was labelled with the sound they perceived.

Each participant completed one practice run, in which they could repeat the presented sound and received feedback, with three repetitions of each vowel. They then performed two test runs with three repetitions of each vowel, in which they could not repeat the presentation of the stimulus and no feedback was given. If a participant’s scores were greater than 10% apart, they performed a third test run. Scores from all test runs were averaged for each participant. Practice runs were not included in the average score. Percent correct scores were converted to rationalized arcsine units (RAU) to normalize error variance (Studebaker, 1985).

### ***The SMRT***

Participants with CIs also performed a test of broadband spectral discrimination, the SMRT. Stimuli consisted of 202 summed pure tones with drifting phase, as in Aronoff and Landsberger (2013). Stimuli were presented at 65 dB SPL in a 3IFC one-down one-up adaptive procedure with ten reversals. The interface for this test consisted of three boxes on a computer screen labelled with either “1,” “2,” or “3.” Each box was highlighted in red during presentation of the corresponding first, second, or third sound in a trial. Participants were asked to select the box of the sound that “sounded different” from the others.

Thresholds for each run were calculated based on the average of the last six reversals, with higher thresholds indicating better spectral discrimination abilities. Each participant completed one practice run and two test runs. The practice run was identical to the test run but this data was not included in the calculation of participants’ average SMRT thresholds. Repetition of the sounds was not allowed during practice or in test runs. If a participant’s thresholds from two test runs were greater than one ripple per octave (RPO) different, he or she completed a third run. Results from all test runs were averaged to determine the mean SMRT threshold for each subject. SMRT data were not collected from one adult subject (SC03) due to time constraints.

The CI processor distorts spectral ripple stimuli at high RPO values (O’Brien & Winn, 2017; DiNino & Arenberg, 2018); still, such tests are widely used in studies of CI listeners. One aim of the current study was to compare CI users’ performance on the test of speech-based spectral resolution (a linguistic test that assesses use of spectral cues) to performance on a spectral ripple task (a non-linguistic assessment of spectral discrimination). As spectral ripple tests may have limited utility and/or predictive power for CI users who achieve high RPO values, the test of

speech-based spectral resolution may be more suitable for assessing spectral resolution in children and adults with CIs.

### Statistical Analyses

All statistical comparisons were conducted in IBM SPSS Statistics (IBM Corp., 2010). Mixed-model regression analyses were performed to examine the effect of age group and hearing modality on perceptual cue weighting. “Hearing group” (NH or CI) and “Age group” (child or adult) were included as independent variables in these models. Individual participants’ slope coefficients for each cue from the logistic regression analysis were utilized as dependent variables. Formant coefficients were the dependent variable for the first model, and spectral tilt coefficients were the dependent variables for the second. “Subject” was set as a random intercept in these models. In addition, an independent t-test was performed to compare formant and slope coefficients between the young and older NH adults.

As cue-trading had been observed by Winn and Litovsky (2015) for adult listeners, the next set of analyses investigated whether cue trading existed in the group of children with CIs. A correlation analysis was performed between the formant and spectral tilt slope coefficients for children in this group. A separate correlation analysis was conducted with the adult CI formant and spectral tilt cue usage for comparison.

SMRT thresholds and vowel identification scores were next compared between groups of CI users. Repeated-measures mixed-model regression analyses were performed with “Subject” as the random intercept and CI age group (child or adult) as the predictor variable. SMRT thresholds of individual participants were the dependent variable for the first model and vowel identification scores in RAU were dependent variables for the second model. All child and some adult participants were bilaterally-implanted, and therefore “Ear Implanted” (first- or second-implanted CI) was set as a random slope, and an unstructured covariance matrix was specified for both models.

The last set of analyses examined the relation of CI users’ perceptual weighting of formant and spectral tilt cues to performance on vowel identification scores in quiet and on the SMRT. Separate mixed-model regressions were performed for each group of CI users, with each test of interest (vowel identification scores in quiet or SMRT thresholds) as the dependent variable. Formant and spectral tilt slope coefficients from each CI were included as independent variables

in all analyses and “Subject” was again set as the random intercept. Only half of the adult participants with CIs had two implants, but all children with CIs were bilaterally-implanted; therefore “Ear Implanted” was included as a random slope in the analyses for children’s data. An unstructured covariance matrix was specified for these models.

## RESULTS

The patterns of cue weighting were similar between the NH groups. Both children and adults with NH weighted the formant cue highly, although children exhibited lower formant slope coefficients compared to adults (NH adults:  $\beta = 2.12$ ; NH children:  $\beta = 1.82$ ). All NH participants placed a relatively lower perceptual weight on spectral tilt. Slope coefficients for this cue were almost identical between age groups of NH listeners (NH adults:  $\beta = 0.61$ ; NH children:  $\beta = 0.62$ ). Participants with CIs also demonstrated analogous perceptual weighting patterns across age groups within this hearing modality. Both child and adult CI users had lower formant slope coefficients relative to NH participants (CI adults:  $\beta = 0.38$ , CI children:  $\beta = 0.35$ ). Adult CI users demonstrated greater use of the spectral tilt cue than NH adults ( $\beta = 0.83$ ). However, children with CIs were not able to use spectral tilt ( $\beta = 0.38$ ) as effectively as participants in the other groups.

Psychometric functions for each group are shown in Figures 5.3 and 5.4. NH adults and NH children exhibited steeply sloping functions for the formant cue and relatively flat functions for the spectral tilt cue. Children and adults with CIs both demonstrated shallower functions than NH participants for the formant cue. As perceptual weight for spectral tilt was similar between all groups, the CI users had moderately sloping functions for use of this cue.

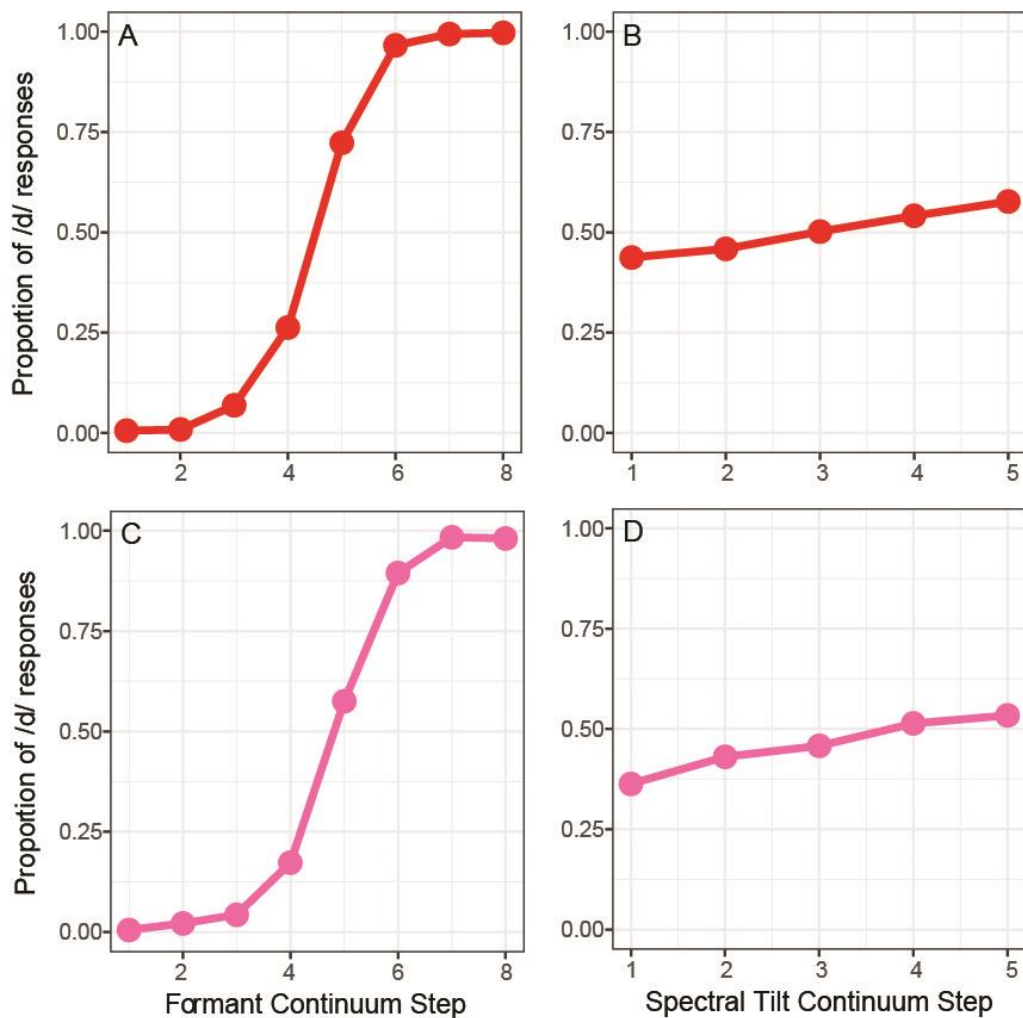


Figure 5-3. Psychometric functions of children and adults with NH.

Proportion of /d/ responses as a function of the cue continuum step for the formant (A, C) and spectral tilt (B, D) cues, for NH adults (in red; A, B) and children (in pink; C, D). Steeper psychometric functions represent greater perceptual weighting.

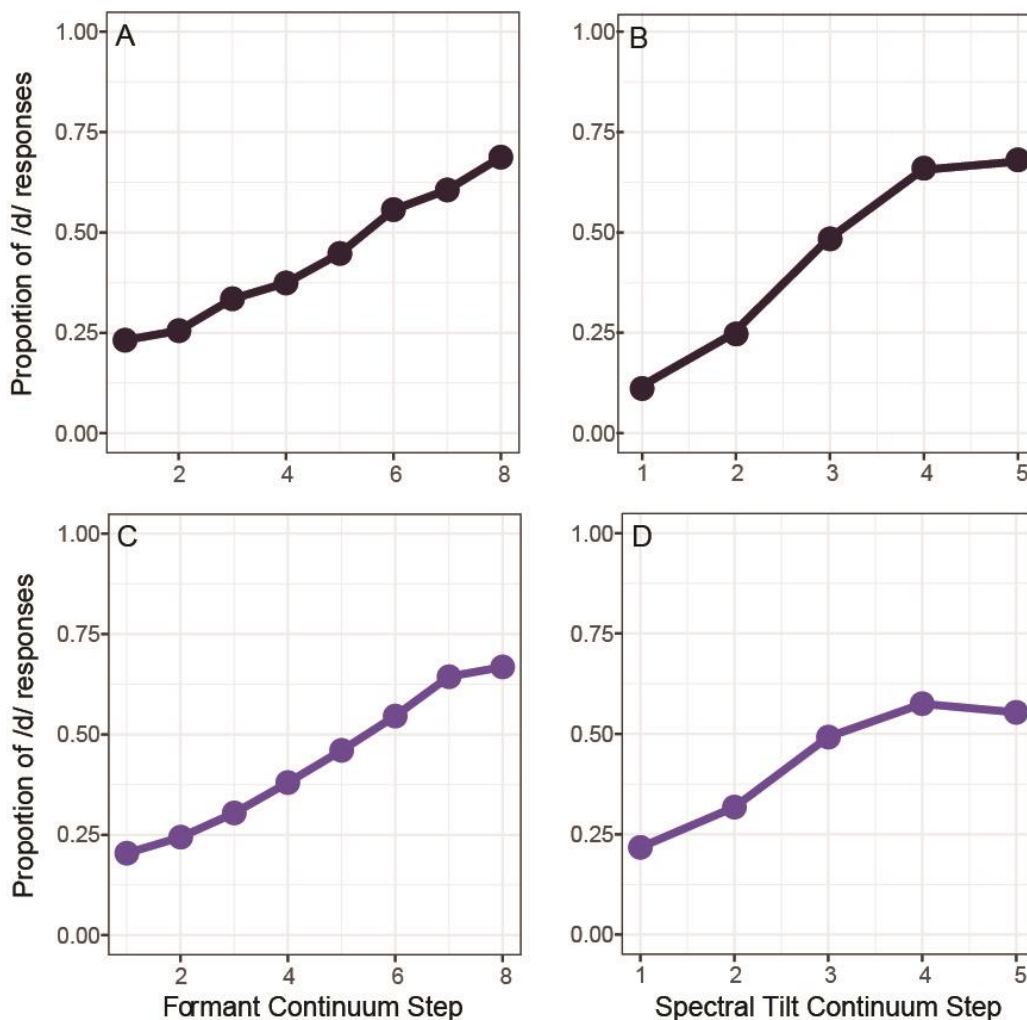


Figure 5-4. Psychometric functions of children and adults with CIs.

Proportion of /d/ responses as a function of the cue continuum step for the formant (A, C) and spectral tilt (B, D) cues, adult (in dark purple; A, B) children (in light purple; C, D) CI users.

Steeper psychometric functions represent greater perceptual weighting.

Accordingly, mixed-model regression analyses revealed a significant effect of hearing modality [ $F(1,89) = 6.8, p = 0.011$ ] but not age group [ $F(1,89) = 1.4, p = 0.24$ ] on perceptual weighting of formant transitions. No significant effect of either hearing [ $F(1,89) = 1.8, p = 0.19$ ] or age [ $F(1,89) = 0.9, p = 0.34$ ] was found on spectral tilt weighting. The comparison of cue-weighting patterns between young and older adult NH groups indicated that perceptual weighting of formant [ $t(12) = 0.75, p = 0.47$ ] and spectral tilt [ $t(12) = 0.72, p = 0.92$ ] did not differ significantly between these groups. Therefore, the older age of NH adults did not significantly affect use of either cue.

A correlation analysis to determine whether children with CIs demonstrated cue trading resulted in a strong, significant positive correlation between formant and spectral tilt cue coefficients ( $R^2 = 0.64, p = < 0.001$ ). Instead of demonstrating a cue trading relationship, this result indicates that early-implanted children utilized both cues to an equal extent. Further, unlike the results observed by Winn and Litovsky (2015), a significant negative relationship was not found between the formant and spectral tilt cues ( $R^2 = 0.007, p = 0.717$ ) for the adults with CIs, signifying that cue trading did not occur in this group of CI users. Figure 5.5 shows the relation between perceptual weighting of the formant and tilt cues for children and adults with CIs.

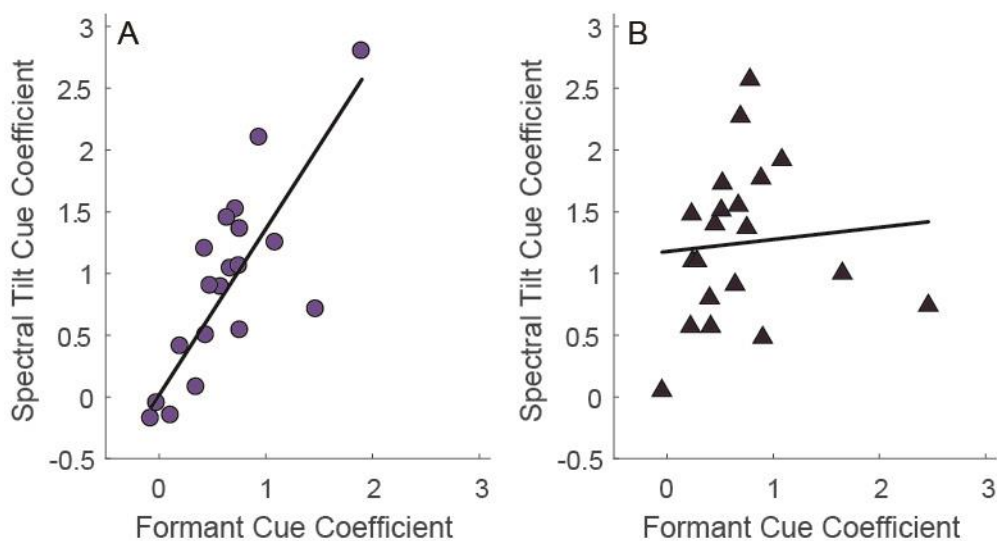


Figure 5-5. Relation between perceptual weighting of formant and spectral tilt cues for children and adults with CIs.

Formant cue coefficients plots against spectral tilt coefficients of child (A) and adult (B) participants with CIs. Each point represents data from one CI. Children's data are represented by light purple circles and adults' data by dark purple triangles. The black lines on each plot indicate the lines of best fit.

SMRT thresholds ranged between 0.58 and 7.68 (mean = 2.9; standard deviation = 2.0) RPO for pediatric CI users and between 0.85 and 6.15 (mean = 2.8; standard deviation = 1.8) RPO for adults with CIs. Age group did not have a significant effect on SMRT thresholds [ $F(1,23.2) = 0.17, p = 0.69$ ]. Vowel identification scores in quiet ranged from 12.0 to 123.0 (mean = 85.5;

standard deviation = 32.8) RAU in children and from 60.3 to 123.0 (mean = 90.8; standard deviation = 16.3) RAU in adults. Significant differences in vowel identification scores were also not observed between age groups of CI participants [ $F(1, 25.2) = 0.025, p = 0.88$ ]. These results indicate that, on average, children and adults with CIs exhibited similar performance on a test of spectral discrimination and on vowel identification in quiet.

The last set of analyses was conducted to determine whether CI users' perceptual weighting of either formant transitions or spectral tilt significantly predicted spectral discrimination thresholds or vowel identification scores. Formant cue weighting was found to significantly predict SMRT thresholds of children with CIs [ $F(1,11.9) = 11.5, p = 0.005$ ], but not vowel identification scores [ $F(1,11.3) = 1.8, p = 0.21$ ]. Spectral tilt coefficients of children with CIs were not significantly related to either assessment [SMRT thresholds:  $F(1,15.6) = 0, p = 0.99$ ; vowel identification:  $F(1,8.8) = 2.5, p = 0.15$ ]. However, because a strong, positive correlation was observed between the formant and spectral tilt slope coefficients of children with CIs, these analyses were conducted again with separate models for formant and spectral tilt cue usage. A Bonferroni correction for multiple comparisons was applied at  $\alpha = 0.025$ . The separate models to examine the effects of perceptual weighting of cues revealed significant relations between formant [ $F(1,17.5) = 7.6, p = 0.013$ ] and spectral tilt [ $F(1,9.3) = 11.7, p = 0.007$ ] cue coefficients and vowel identification scores in quiet (see Figure 5.6). Results of the revised models demonstrated an even stronger relationship between perceptual weighting of the formant cue and SMRT thresholds [ $F(1,11.1) = 36.3, p < 0.001$ ]. The effect of spectral tilt cue usage on SMRT performance was significant as well [ $F(1,11.0) = 11.4, p = 0.006$ ], although the effect was lower than that of the formant cue (see Figure 5.7). As the children with CIs were found to utilize both cues to a similar extent, it is unsurprising that both cues would significantly predict performance on the vowel identification test and on the SMRT.

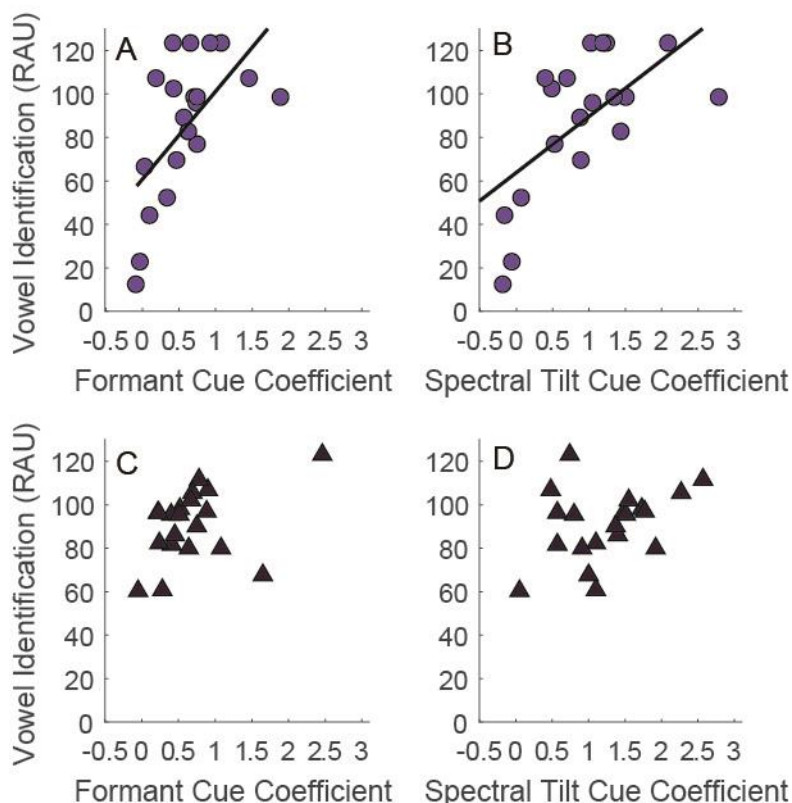


Figure 5-6. Relation between SMRT thresholds and phonemic cue coefficients.

SMRT thresholds in RPO plotted against formant (A, C) and spectral tilt (B, D) cue coefficients for pediatric (A, B) and adult (C, D) CI users. Each point represents data from one CI. Children's data are shown as light purple circles and adults' data as dark purple triangles.

The black lines indicate the lines of best fit for significant relationships.

Mixed-model regression analyses for adult CI users indicated no significant relations between weighting of the formant or spectral tilt cues on vowel recognition performance [formant:  $F(1,17) = 3.4$ ,  $p = 0.08$ ; spectral tilt:  $F(1,17) = 2.5$ ,  $p = 0.14$ ; results shown in Figure 5.6] or on the SMRT [formant:  $F(1,15) = 0.85$ ,  $p = 0.37$ ; spectral tilt:  $F(1,15) = 0.15$ ,  $p = 0.7$ ; results shown in Figure 5.7]. No correlation between perceptual weights for formant and spectral tilt cues was found in this group, so additional analyses with separate models for each cue were not conducted.

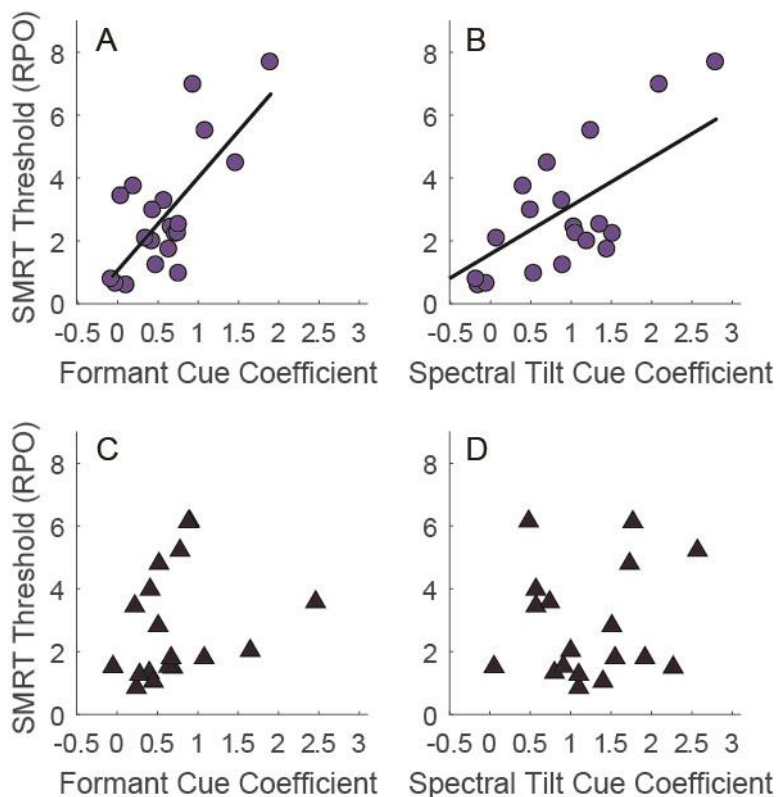


Figure 5-7. Relation between vowel identification scores and phonemic cue coefficients.

Vowel identification scores in RAU plotted against formant (A, C) and spectral tilt (B, D) cue coefficients for children (A, B) and adults (C, D) with CIs. Each point represents data from one CI. Children's data are represented by light purple circles and adults' data as dark purple triangles. The black lines indicate the lines of best fit for significant relationships.

## DISCUSSION

This study investigated speech-based spectral resolution of children with NH and with CIs. The cue-weighting patterns of pediatric participants were compared to those of adults with the same hearing modality. It was hypothesized that both groups of NH listeners would place greater perceptual weight on the fine-resolution formant cue, while individuals with CIs would exhibit less use of the formant cue and instead rely heavily on spectral tilt. It was anticipated that age-related differences would be found between the child and adult groups. Further, early-implanted children and late-implanted adults could differ in their perceptual weighting strategies because of their distinct hearing histories.

Group analyses showed that both child and adult CI users placed less perceptual weight on the formant cue compared to NH listeners. In addition, adults with CIs used the coarse-resolution spectral tilt cue to a greater extent than NH adults. These results are consistent with findings from the previous study in adults that utilized this test (Winn & Litovsky, 2015). The CI limits spectral resolution by broadening frequency bins (due to small number of channels in the implant) and therefore narrowband formant frequency cues are difficult for CI listeners to resolve (Henry & Turner, 2003). To discriminate changes in spectral tilt, however, a listener merely needs to perform a comparison between large spectral regions. Accordingly, Lindholm et al. (1988) found that hearing impaired individuals, who also exhibit broadened auditory filters, utilized spectral tilt more than formant transitions for classifying stop consonants. Similarly, Alexander and Kluender (2009) observed that hearing impaired listeners weighted spectral tilt more highly than NH listeners when categorizing stimuli with unambiguous place of articulation F2 cues. The results from the present study are consistent with previous findings of greater perceptual weighting of spectral tilt in listeners for whom access to formant cues is decreased. In addition, despite having a hearing history that could potentially increase spectral resolving capabilities compared to late-implanted adults, cue-weighting patterns of early-implanted children with CIs were still driven by the limitations of the implant: as a group, these children were not able to utilize formant cues to the same extent as listeners with NH.

Children's overall cue weighting patterns were observed to be similar to those of adults with the same hearing modality; yet, children used the cues less effectively compared to the adult groups. NH children placed less perceptual weight on the formant transition cue than did adults with NH, and children with CIs weighted both cues to a lesser extent than did adult CI users. These results may be indicative of potential developmental effects in cue-weighting patterns. Prior studies have observed lower perceptual weights for phonetic cues in children compared to adults (e.g., Giezen et al., 2010). These findings could be the result of differences in auditory sensitivity between children and adults or immature perceptual strategies. Further, the cue-weighting test used in this study assesses spectral resolution, which has previously been found to improve with age in NH children (Peter et al., 2014) and with longer time with the CI for pediatric CI users (DiNino et al., 2018). A longitudinal study of children with these hearing modalities would provide valuable information related to the development of speech-based spectral resolution.

In the current study, all NH listeners and some adults with CIs demonstrated a cue-trading relationship, while children with CIs did not. These results suggest an immaturity of pediatric CI users in prioritizing either cue. Further, the children with CIs in this study were prelingually-deafened and thus learned language with the CI. Because these children never had much (if any) acoustic hearing, they never had experience with fine-resolution formant cues for categorizing speech sounds. As a result, they seem to show no preference for either cue and weight spectral tilt and formant transitions to a similar extent. This finding is in contrast to the observed results from the late-implanted adults: some exhibited cue-trading, but most demonstrated no relationship between use of the two cues. The results of this study demonstrate that hearing history can influence auditory perception, as differences in perceptual strategies for categorizing speech sounds were found between early-implanted children and late-implanted adults.

Winn and Litovsky (2015) observed that CI users with larger perceptual weights for the formant cue were those with higher word recognition scores. Thus, individuals with cue-weighting patterns more similar to NH listeners were able to perform better on a test of speech recognition. The manipulation of the formant cue in the test of speech-based spectral resolution covers the frequency range that is essential for speech (Kates, 2013), and therefore it follows that a relationship between perceptual weighting of the formant cue and speech identification performance should exist. Vowel identification in particular is dependent on resolution of formant cues, as the contrasts between vowels are primarily spectral. Consistent with this reasoning and with results from the prior study, child CI users in the current study demonstrated a significant relationship between formant cue weighting and vowel identification scores. This relationship was not significant for adults with CIs, but a positive relation between these variables was observed (see Figure 5.6). These findings thus provide additional evidence that outcomes on the test of speech-based spectral resolution can predict CI user speech identification performance. Further research utilizing various speech perception tests is necessary to fully examine the predictive abilities of this speech-based assessment of spectral resolution.

Perceptual weighting of both formant and spectral tilt cues were significantly related to performance on the SMRT in children, but not in adults, with CIs. These results in the adult group are in contrast with those from a prior study which found a significant relationship between performance on a traditional spectral ripple test and use of the formant cue in adult CI users (Winn et al., 2016). These discrepant findings are likely due to differences in the stimuli: unlike traditional

spectral ripple discrimination tasks, the SMRT utilizes spectral ripples with drifting modulation phases, which introduces a temporal component into the signal. Spectral cues may therefore not relate as strongly to performance on the SMRT compared to performance on other tests of spectral resolution. In addition, the finding of no relationship between use of the formant cue and SMRT thresholds in adults with CIs could be due to testing procedures, as each ear was tested separately in the present study. Individuals utilize both ears in everyday listening situations and thus a bilateral or bimodal condition could better represent their abilities. This effect may be larger for adults than for children who have developed with the CI as their sole auditory input.

Although bilateral testing may be more akin to real-world listening situations, testing individual CIs allowed for assessment of phonetic cue weighting while minimizing between-subject factors such as intelligence quotient (IQ), cognitive abilities, and reading performance. In this study, perceptual weighting of phonetic cues was found to differ between the two CIs of most bilaterally-implanted participants. These results indicate that at least some factors that contribute to cue-weighting patterns are related to individual CIs, rather than global perceptual strategies. In addition, some child and adult participants with CIs were able to utilize the formant cue to a greater extent than other CI users, indicating potentially better spectral resolving abilities of these individuals; yet, even these individuals differed in their cue-weighting patterns between ears. Future directions of this work include investigation of listening experience and other central and peripheral determinants that may influence perceptual cue weighting with individual CIs.

## CONCLUSION

Results from this study demonstrated that speech-based spectral resolution is driven mainly by hearing modality. However, hearing history of CI users and age group of both NH and CI participants played a role. These findings indicate that developmental effects may influence perceptual weighting of the phonetic cues utilized in this test of speech-based spectral resolution. Results from this study also showed differences in speech-related auditory perception strategies between early-implanted children and late-implanted adults with CIs. In addition, bilaterally-implanted participants with CIs were found to weight formant and spectral tilt cues differentially with separate ears. Therefore, in addition to predicting speech recognition scores of CI users with their everyday clinical processing program, cue-weighting assessments have utility for studies of novel CI program processing strategies: such tests could be used to determine how the

experimental program alters perception of particular phonetic cues through the implant. Cue-weighting patterns may thus provide more detailed information about the implant than traditional tests of spectral resolution or assessments of speech identification performance.

## CHAPTER 6 : CONCLUSION

The results of these experiments can be summarized by the following:

1. Vocoder simulations of suboptimal CI electrode-neuron interfaces in NH adults decreases vowel and consonant identification performance relative to a control vocoder condition
2. NH listeners confused consonants of similar manner of articulation, but vowel confusions were unidirectional and shifted away from the frequency region of spectral degradation
3. NH listeners' vowel confusions resulting from spectral degradation in particular frequency regions were similar to those of two example CI users with poor electrode-neuron interfaces covering those same frequency regions
4. Chronological age predicted vocoded vowel identification performance of NH children but not unprocessed vowel identification of early-implanted children with CIs
5. Amount of time with the CI was related to both performance on vowel recognition in quiet and the Spectral-temporally Modulated Ripple Test (SMRT) in early-implanted children
6. SMRT thresholds were a significant predictor of vocoded vowel identification scores of children with NH and of vowel identification in quiet scores of children with CIs
7. Perception of the second formant (F2) feature of vowel stimuli was also related to performance on the SMRT in both groups of children
8. Auditory detection thresholds were lower in early-implanted children compared to late-implanted adults with CIs, particularly when using a focused electrical configuration
9. Channel-to-channel variability in auditory detection thresholds did not differ between child and adult CI users, indicating that early-implanted children exhibit variance in electrode-neuron interface quality that is similar to that observed in late-implanted adults
10. Dynamic ranges of auditory perception were similar between early-implanted children and late-implanted adults with CIs, but the upper and lower limits were shifted downward in children relative to adults
11. Early-implanted children demonstrated higher levels of intracochlear resistance compared to adults, potentially indicative of greater levels of ossification and tissue growth after CI surgery in the pediatric population

12. NH children and adults were able to utilize a fine-resolution cue to a greater extent than children and adults with CIs to categorize speech sounds, although children with NH placed less perceptual weight on the formant cue compared to NH adults
13. Early-implanted children and late-implanted adults with CIs demonstrated similar patterns of phonetic cue weighting, but children made less efficient use of both cues
14. NH individuals demonstrated cue trading: high perceptual weighting of the fine-resolution cue and low weighting of coarse-resolution cue
15. Early-implanted pediatric CI users tended to use both cues to a similar degree, indicative of altered perceptual strategies for categorizing speech compared to individuals with (or who have previously had) normal, acoustic hearing

Overall, the results of these experiments provide a better understanding of the respective contributions of hearing history and the electrode-neuron interface on auditory perception of individuals with CIs. These experiments were also the first to demonstrate differences in auditory perception and physical aspects of the cochlea between early-implanted children and late-implanted adults. Yet, there are still many future directions of this research.

Previous studies in adults with CIs have demonstrated that channels with relatively elevated focused thresholds also exhibit broad psychophysical tuning curves (PTCs), indicating poor spatial selectivity (Bierer & Faulkner, 2010). The existence of particular channels with broad PTCs have been related to poor performance on narrowband spectral ripple density resolution in the corresponding frequency range (Anderson et al., 2011). However, the relation between suboptimal electrode-neuron interfaces and spatial selectivity is unknown in early-implanted children. An investigation of PTCs and narrowband spectral ripple discrimination could be conducted to determine whether suboptimal electrode-neuron interfaces affect spectral discrimination to the same extent in early-implanted children as has been previously found in late-implanted adults.

Similarly, evidence from Experiment 1 showed that that distortions of particular frequency regions leads to predictable confusions in NH listeners. These confusions were similar to those of two example adult CI users with elevated focused thresholds in the same frequency region as the manipulated channels in the vocoder simulation. This indicates that CI users' vowel errors and confusions can be related to poor transmission of formant frequencies carried by electrodes with

elevated focused thresholds. Experiment 1 examined the effects of large regions (4 channels) of frequency loss or distortion, but a future study could investigate whether individual channels with suboptimal electrode-neuron interfaces result in vowel confusions related to the frequency information normally transmitted through that channel. Such a study could help to determine the spectral cues degraded for children and adults with CIs and inform modifications to their program to attempt to provide the missing information.

The results of Experiment 3 suggested that early-implanted children with CIs have healthier electrode-neuron interfaces compared to late-implanted adults, presumably due to better spiral ganglion neuron survival. These results coincide with temporal bone studies that have found greater degrees of spiral ganglion degeneration with aging (Makary et al., 2011) and with prolonged periods of hearing loss (Otte et al., 1978), but potential differences in the status of the electrode-neuron interface between early-implanted children and late-implanted adults have not been determined physiologically. Future research utilizing objective assessments of spiral ganglion survival, such as the electrically-evoked compound action potential (ECAP; Schwartz-Leyzac & Pfingst, 2016), should be performed in these groups of CI users to follow up on the research in this dissertation.

Experiments 3 and 4 found differences in the results of psychoacoustic measures between early-implanted children and late-implanted adults. These measures were related auditory perceptual thresholds, spectral resolution, and speech identification, but additional comparisons between these groups could be made with other auditory tasks. For example, these could include assessments of temporal resolution or spatial localization. Such investigations in early-implanted children and late-implanted adults would provide information about potential differences between these groups in the perception of additional auditory domains.

The experiments in this dissertation, along with the aforementioned next steps of this research, could lead to development of CI programming processing strategies tailored towards early-implanted children. Recent investigations conducted in the Arenberg Lab and others have shown that adult CI user speech perception can be improved by optimizing spectral resolution through device programming (Bierer & Litvak, 2016). This dissertation identified differences in auditory perception and the electrode-neuron interface between children and adults with CIs; these results thus provided a foundation of knowledge on which to develop novel stimulation paradigms

to optimize transmission of spectral features for early-implanted children. This could be a significant step toward enhancing speech recognition abilities for children with CIs.

## BIBLIOGRAPHY

- [1] Ainsworth, W. A. (1972). Duration as a Cue in the Recognition of Synthetic Vowels. *The Journal of the Acoustical Society of America*, 51(2B), 648–651. <https://doi.org/10.1121/1.1912889>
- [2] Alexander, J. M., & Kluender, K. R. (2008). Spectral tilt change in stop consonant perception. *The Journal of the Acoustical Society of America*, 123(1), 386–396. <https://doi.org/10.1121/1.2817617>
- [3] Alexander, J. M., & Kluender, K. R. (2009). Spectral tilt change in stop consonant perception by listeners with hearing impairment. *Journal of Speech, Language, and Hearing Research: JSLHR*, 52(3), 653–670. [https://doi.org/10.1044/1092-4388\(2008/08-0038\)](https://doi.org/10.1044/1092-4388(2008/08-0038))
- [4] Allen, P., & Wightman, F. (1992). Spectral pattern discrimination by children. *Journal of Speech and Hearing Research*, 35(1), 222–233.
- [5] Anderson, E. S., Nelson, D. A., Kreft, H., Nelson, P. B., & Oxenham, A. J. (2011). Comparing spatial tuning curves, spectral ripple resolution, and speech perception in cochlear implant users. *The Journal of the Acoustical Society of America*, 130(1), 364–375. <https://doi.org/10.1121/1.3589255>
- [6] Arenberg, J.G., Parkinson, W.S., Litvak, L., Chen, C., Kreft, H.A., & Oxenham, A.J. (2018). A dynamically focusing cochlear implant strategy can improve vowel identification in noise. *Ear and Hearing*.
- [7] Aronoff, J. M., & Landsberger, D. M. (2013). The development of a modified spectral ripple test. *The Journal of the Acoustical Society of America*, 134(2), EL217-222. <https://doi.org/10.1121/1.4813802>
- [8] Assmann, P. F., & Katz, W. F. (2005). Synthesis fidelity and time-varying spectral change in vowels. *The Journal of the Acoustical Society of America*, 117(2), 886–895.
- [9] Azadpour, M., & Smith, R. L. (2016). Enhancing speech envelope by integrating hair-cell adaptation into cochlear implant processing. *Hearing Research*, 342, 48–57. <https://doi.org/10.1016/j.heares.2016.09.008>
- [10] Bahng, J., Hedrick, M., & von Hapsburg, D. (2014). Weighting of static and transition cues in voiceless fricatives and stops in children wearing cochlear implants. *Clinical and*

- Experimental Otorhinolaryngology, 7(4), 254–259.  
<https://doi.org/10.3342/ceo.2014.7.4.254>
- [11] Bates, D, Maechler, M, Bolker, B, Walker, S (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67:1-48
- [12] Békésy, G. v. (1947). A New Audiometer. *Acta Oto-Laryngologica*, 35(5–6), 411–422.  
<https://doi.org/10.3109/00016484709123756>
- [13] Berenstein, C. K., Mens, L. H. M., Mulder, J. J. S., & Vanpoucke, F. J. (2008). Current steering and current focusing in cochlear implants: comparison of monopolar, tripolar, and virtual channel electrode configurations. *Ear and Hearing*, 29(2), 250–260.
- [14] Bierer, J.A., Bierer, S. M., Kreft, H. A., & Oxenham, A. J. (2015a). A fast method for measuring psychophysical thresholds across the cochlear implant array. *Trends in Hearing*, 19. <https://doi.org/10.1177/2331216515569792>
- [15] Bierer, J.A. (2007). Threshold and channel interaction in cochlear implant users: evaluation of the tripolar electrode configuration. *The Journal of the Acoustical Society of America*, 121(3), 1642–1653.
- [16] Bierer, J.A. (2010). Probing the electrode-neuron interface with focused cochlear implant stimulation. *Trends in Amplification*, 14(2), 84–95.  
<https://doi.org/10.1177/1084713810375249>
- [17] Bierer, J.A. & Faulkner, K. F. (2010). Identifying cochlear implant channels with poor electrode-neuron interface: partial tripolar, single-channel thresholds and psychophysical tuning curves. *Ear and Hearing*, 31(2), 247–258.  
<https://doi.org/10.1097/AUD.0b013e3181c7daf4>
- [18] Bierer, J. A., Faulkner, K. F., & Tremblay, K. L. (2011). Identifying cochlear implant channels with poor electrode-neuron interfaces: electrically evoked auditory brain stem responses measured with the partial tripolar configuration. *Ear and Hearing*, 32(4), 436–444. <https://doi.org/10.1097/AUD.0b013e3181ff33ab>
- [19] Bierer, J.A., & Middlebrooks, J. C. (2002). Auditory cortical images of cochlear-implant stimuli: dependence on electrode configuration. *Journal of Neurophysiology*, 87(1), 478–492. <https://doi.org/10.1152/jn.00212.2001>

- [20] Bierer, J.A., & Nye, A. D. (2014). Comparisons between detection threshold and loudness perception for individual cochlear implant channels. *Ear and Hearing*, 35(6), 641–651. <https://doi.org/10.1097/AUD.0000000000000058>
- [21] Bierer, S.M., Shea-Brown, E., Bierer, J.A. (2015b) Current spread in the cochlea: Insights from CT and Electrical Field Imaging. Poster presented at the Conference on Implantable Auditory Prostheses, Tahoe, CA. T38.
- [22] Bilger, R. C., & Wang, M. D. (1976). Consonant confusions in patients with sensorineural hearing loss. *Journal of Speech and Hearing Research*, 19(4), 718–748.
- [23] Bingabr, M., Espinoza-Varas, B., & Loizou, P. C. (2008). Simulating the effect of spread of excitation in cochlear implants. *Hearing Research*, 241(1–2), 73–79. <https://doi.org/10.1016/j.heares.2008.04.012>
- [24] Blamey, P. J., Pyman, B. C., Gordon, M., Clark, G. M., Brown, A. M., Dowell, R. C., & Hollow, R. D. (1992). Factors predicting postoperative sentence scores in postlinguistically deaf adult cochlear implant patients. *The Annals of Otology, Rhinology, and Laryngology*, 101(4), 342–348. <https://doi.org/10.1177/000348949210100410>
- [25] Boersma, P., and Weenink, D. (2013). “Praat: Doing phonetics by computer (version 5.3.56),” from <http://www.praat.org>.
- [26] Boëx, C., Kós, M.-I., & Pelizzone, M. (2003). Forward masking in different cochlear implant systems. *The Journal of the Acoustical Society of America*, 114(4 Pt 1), 2058–2065.
- [27] Boons, T., Brokx, J., Frijns, J., Philips, B., Vermeulen, A., Wouters, J., & van Wieringen, A. (2013). Newborn hearing screening and cochlear implantation: impact on spoken language development. *B-ENT, Suppl 21*, 91–98.
- [28] Boothroyd, A., Mulhearn, B., Gong, J., & Ostroff, J. (1996). Effects of spectral smearing on phoneme and word recognition. *The Journal of the Acoustical Society of America*, 100(3), 1807–1818.
- [29] Dahm, M. C., Shepherd, R. K., & Clark, G. M. (1993). The postnatal growth of the temporal bone and its implications for cochlear implantation in children. *Acta Oto-Laryngologica. Supplementum*, 505, 1–39.

- [30] Dawes, P., & Bishop, D. V. M. (2008). Maturation of visual and auditory temporal processing in school-aged children. *Journal of Speech, Language, and Hearing Research: JSLHR*, 51(4), 1002–1015. [https://doi.org/10.1044/1092-4388\(2008/073\)](https://doi.org/10.1044/1092-4388(2008/073))
- [31] de Jong, M. A. M., Briare, J. J., & Frijns, J. H. M. (2017). Learning Effects in Psychophysical Tests of Spectral and Temporal Resolution. *Ear and Hearing*. <https://doi.org/10.1097/AUD.0000000000000499>
- [32] DeVries, L., Scheperle, R., & Bierer, J. A. (2016). Assessing the Electrode-Neuron Interface with the Electrically Evoked Compound Action Potential, Electrode Position, and Behavioral Thresholds. *Journal of the Association for Research in Otolaryngology: JARO*, 17(3), 237–252. <https://doi.org/10.1007/s10162-016-0557-9>
- [33] DiNino, M., Arenberg, J.G. (2018). Age-related performance on vowel identification and the Spectral-temporally Modulated Ripple Test in children with normal hearing and with cochlear implants. *Trends in Hearing*, 22, 1-20. <https://doi.org/10.1177/2331216518770959>
- [34] DiNino, M., O'Brien, G., & Arenberg, J.A. (2017). Assessment of the electrode-neuron interface in children with cochlear implants compared to adults. Paper presented at the Conference on Implantable Auditory Prostheses, Tahoe, CA.
- [35] DiNino, M., Wright, R. A., Winn, M. B., & Bierer, J. A. (2016). Vowel and consonant confusions from spectrally manipulated stimuli designed to simulate poor cochlear implant electrode-neuron interfaces. *The Journal of the Acoustical Society of America*, 140(6), 4404–4418. <https://doi.org/10.1121/1.4971420>
- [36] Dorman, M. F., Hannley, M. T., Dankowski, K., Smith, L., & McCandless, G. (1989). Word recognition by 50 patients fitted with the Symbion multichannel cochlear implant. *Ear and Hearing*, 10(1), 44–49.
- [37] Dorman, M. F., Loizou, P. C., Kemp, L. L., & Kirk, K. I. (2000). Word recognition by children listening to speech processed into a small number of channels: data from normal-hearing children and children with cochlear implants. *Ear and Hearing*, 21(6), 590–596.
- [38] Dorman, M. F., Loizou, P. C., & Rainey, D. (1997). Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. *The Journal of the Acoustical Society of America*, 102(4), 2403–2411.

- [39] Drennan, W. R., Won, J. H., Timme, A. O., & Rubinstein, J. T. (2016). Nonlinguistic outcome measures in adult cochlear implant users over the first year of implantation. *Ear and Hearing*, 37(3), 354–364. <https://doi.org/10.1097/AUD.0000000000000261>
- [40] Dudley, H. (1939). The automatic synthesis of speech. *Proceedings of the National Academy of Sciences of the United States of America*, 25(7), 377–383.
- [41] Eggermont, J. J., Ponton, C. W., Don, M., Waring, M. D., & Kwong, B. (1997). Maturational delays in cortical evoked potentials in cochlear implant users. *Acta Oto-Laryngologica*, 117(2), 161–163.
- [42] Eggermont, Jos J., & Ponton, C. W. (2003). Auditory-evoked potential studies of cortical maturation in normal hearing and implanted children: correlations with changes in structure and speech perception. *Acta Oto-Laryngologica*, 123(2), 249–252.
- [43] Eisenberg, L. S., Shannon, R. V., Martinez, A. S., Wygonski, J., & Boothroyd, A. (2000). Speech recognition with reduced spectral cues as a function of age. *The Journal of the Acoustical Society of America*, 107(5 Pt 1), 2704–2710.
- [44] Eisenberg, Laurie S., Fisher, L. M., Johnson, K. C., Ganguly, D. H., Grace, T., Niparko, J. K., & CDaCI Investigative Team. (2016). Sentence recognition in quiet and noise by pediatric cochlear implant users: Relationships to spoken language. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 37(2), e75-81. <https://doi.org/10.1097/MAO.0000000000000910>
- [45] Eisenberg, Laurie S., Martinez, A. S., Holowecky, S. R., & Pogorelsky, S. (2002). Recognition of lexically controlled words and sentences by children with normal hearing and children with cochlear implants. *Ear and Hearing*, 23(5), 450–462. <https://doi.org/10.1097/01.AUD.0000034736.42644.BE>
- [46] Elliott, L. L. (1979). Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. *The Journal of the Acoustical Society of America*, 66(3), 651–653.
- [47] Finley, C. C., Holden, T. A., Holden, L. K., Whiting, B. R., Chole, R. A., Neely, G. J., ... Skinner, M. W. (2008a). Role of electrode placement as a contributor to variability in cochlear implant outcomes. *Otology & Neurotology: Official Publication of the American*

- Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology, 29(7), 920–928. <https://doi.org/10.1097/MAO.0b013e318184f492>
- [48] Firszt, J. B., Chambers And, R. D., & Kraus, N. (2002). Neurophysiology of cochlear implant users II: comparison among speech perception, dynamic range, and physiological measures. *Ear and Hearing*, 23(6), 516–531. <https://doi.org/10.1097/01.AUD.0000042154.70495.DE>
- [49] Firszt, J. B., Holden, L. K., Skinner, M. W., Tobey, E. A., Peterson, A., Gaggl, W., ... Wackym, P. A. (2004). Recognition of speech presented at soft to loud levels by adult cochlear implant recipients of three cochlear implant systems. *Ear and Hearing*, 25(4), 375–387.
- [50] Friesen, L. M., Shannon, R. V., Baskent, D., & Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. *The Journal of the Acoustical Society of America*, 110(2), 1150–1163.
- [51] Fu, Q. J., & Shannon, R. V. (1999). Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing. *The Journal of the Acoustical Society of America*, 105(3), 1889–1900.
- [52] Garadat, S. N., Zwolan, T. A., & Pfungst, B. E. (2013). Using temporal modulation sensitivity to select stimulation sites for processor MAPs in cochlear implant listeners. *Audiology & Neuro-Otology*, 18(4), 247–260. <https://doi.org/10.1159/000351302>
- [53] Geers, A., Brenner, C., & Davidson, L. (2003). Factors associated with development of speech perception skills in children implanted by age five. *Ear and Hearing*, 24(1 Suppl), 24S–35S. <https://doi.org/10.1097/01.AUD.0000051687.99218.0F>
- [54] Geers, A. E., Moog, J. S., Biedenstein, J., Brenner, C., & Hayes, H. (2009). Spoken language scores of children using cochlear implants compared to hearing age-mates at school entry. *Journal of Deaf Studies and Deaf Education*, 14(3), 371–385. <https://doi.org/10.1093/deafed/enn046>
- [55] Giezen, M. R., Escudero, P., & Baker, A. (2010). Use of acoustic cues by children with cochlear implants. *Journal of Speech, Language, and Hearing Research: JSLHR*, 53(6), 1440–1457. [https://doi.org/10.1044/1092-4388\(2010/09-0252\)](https://doi.org/10.1044/1092-4388(2010/09-0252))

- [56] Gifford, R. H., Shallop, J. K., & Peterson, A. M. (2008). Speech recognition materials and ceiling effects: considerations for cochlear implant programs. *Audiology & Neuro-Otology*, 13(3), 193–205. <https://doi.org/10.1159/000113510>
- [57] Goldwyn, J. H., Bierer, S. M., & Bierer, J. A. (2010). Modeling the electrode-neuron interface of cochlear implants: effects of neural survival, electrode placement, and the partial tripolar configuration. *Hearing Research*, 268(1–2), 93–104. <https://doi.org/10.1016/j.heares.2010.05.005>
- [58] Greenwood, D. D. (1990). A cochlear frequency-position function for several species--29 years later. *The Journal of the Acoustical Society of America*, 87(6), 2592–2605.
- [59] Hall, J. W., Buss, E., Grose, J. H., & Dev, M. B. (2004). Developmental effects in the masking-level difference. *Journal of Speech, Language, and Hearing Research: JSLHR*, 47(1), 13–20. [https://doi.org/10.1044/1092-4388\(2004/002\)](https://doi.org/10.1044/1092-4388(2004/002))
- [60] Halpin, K. S., Smith, K. Y., Widen, J. E., & Chertoff, M. E. (2010). Effects of universal newborn hearing screening on an early intervention program for children with hearing loss, birth to 3 yr of age. *Journal of the American Academy of Audiology*, 21(3), 169–175. <https://doi.org/10.3766/jaaa.21.3.5>
- [61] Harnsberger, J. D., Svirsky, M. A., Kaiser, A. R., Pisoni, D. B., Wright, R., & Meyer, T. A. (2001). Perceptual “vowel spaces” of cochlear implant users: implications for the study of auditory adaptation to spectral shift. *The Journal of the Acoustical Society of America*, 109(5 Pt 1), 2135–2145.
- [62] Hartley, D. E., Wright, B. A., Hogan, S. C., & Moore, D. R. (2000). Age-related improvements in auditory backward and simultaneous masking in 6- to 10-year-old children. *Journal of Speech, Language, and Hearing Research: JSLHR*, 43(6), 1402–1415.
- [63] Hazan, V., & Barrett, S. (2000). The development of phonemic categorization in children aged 6–12. *Journal of Phonetics*, 28(4), 377–396. <https://doi.org/10.1006/jpho.2000.0121>
- [64] Hedrick, M., Bahng, J., von Hapsburg, D., & Younger, M. S. (2011). Weighting of cues for fricative place of articulation perception by children wearing cochlear implants. *International Journal of Audiology*, 50(8), 540–547. <https://doi.org/10.3109/14992027.2010.549515>
- [65] Heffner, R.-M.S., (1937). Notes on the length of vowels. *American Speech*, 12(2), 128–134.

- [66] Henry, B. A., & Turner, C. W. (2003). The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. *The Journal of the Acoustical Society of America*, 113(5), 2861–2873.
- [67] Henry, B. A., Turner, C. W., & Behrens, A. (2005). Spectral peak resolution and speech recognition in quiet: normal hearing, hearing impaired, and cochlear implant listeners. *The Journal of the Acoustical Society of America*, 118(2), 1111–1121.
- [68] Hepper, P. G., & Shahidullah, B. S. (1994). Development of fetal hearing. *Archives of Disease in Childhood*, 71(2), F81-87.
- [69] Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, 97(5 Pt 1), 3099–3111.
- [70] Hillenbrand, J. M., & Nearey, T. M. (1999). Identification of resynthesized /hVd/ utterances: effects of formant contour. *The Journal of the Acoustical Society of America*, 105(6), 3509–3523.
- [71] Hirata, Y. (2004). Effects of speaking rate on the vowel length distinction in Japanese. *Journal of Phonetics*, 32(4), 565–589. <https://doi.org/10.1016/j.wocn.2004.02.004>
- [72] Hirata, Y., & Tsukada, K. (2009). Effects of speaking rate and vowel length on formant frequency displacement in Japanese. *Phonetica*, 66(3), 129–149. <https://doi.org/10.1159/000235657>
- [73] Holden, L. K., Finley, C. C., Firszt, J. B., Holden, T. A., Brenner, C., Potts, L. G., ... Skinner, M. W. (2013). Factors affecting open-set word recognition in adults with cochlear implants. *Ear and Hearing*, 34(3), 342–360. <https://doi.org/10.1097/AUD.0b013e3182741aa7>
- [74] Holden, L. K., Firszt, J. B., Reeder, R. M., Uchanski, R. M., Dwyer, N. Y., & Holden, T. A. (2016). Factors affecting outcomes in cochlear implant recipients implanted with a perimodiolar electrode array located in scala tympani. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 37(10), 1662–1668. <https://doi.org/10.1097/MAO.0000000000001241>
- [75] Holt, L. L., & Lotto, A. J. (2010). Speech perception as categorization. *Attention, Perception & Psychophysics*, 72(5), 1218–1227. <https://doi.org/10.3758/APP.72.5.1218>

- [76] Horn, D. L., Dudley, D. J., Dedhia, K., Nie, K., Drennan, W. R., Won, J. H., ... Werner, L. A. (2017). Effects of age and hearing mechanism on spectral resolution in normal hearing and cochlear-implanted listeners. *The Journal of the Acoustical Society of America*, 141(1), 613. <https://doi.org/10.1121/1.4974203>
- [77] Hudspeth, A. J. (1997). How hearing happens. *Neuron*, 19(5), 947–950.
- [78] Hughes, M. L., Choi, S., & Glickman, E. (2018). What can stimulus polarity and interphase gap tell us about auditory nerve function in cochlear-implant recipients? *Hearing Research*, 359, 50–63. <https://doi.org/10.1016/j.heares.2017.12.015>
- [79] IBM SPSS Statistics for Windows, Version 19.0. IBM Corp. Released 2010. Armonk, NY: IBM Corp.
- [80] Idemaru, K., & Holt, L. L. (2013). The developmental trajectory of children's perception and production of English /r/-/l/. *The Journal of the Acoustical Society of America*, 133(6), 4232–4246. <https://doi.org/10.1121/1.4802905>
- [81] Johnson, C. E. (2000). Children's phoneme identification in reverberation and noise. *Journal of Speech, Language, and Hearing Research: JSLHR*, 43(1), 144–157.
- [82] Jolly, C. N., Spelman, F. A., & Clopton, B. M. (1996). Quadrupolar stimulation for Cochlear prostheses: modeling and experimental data. *IEEE Transactions on Bio-Medical Engineering*, 43(8), 857–865. <https://doi.org/10.1109/10.508549>
- [83] Jung, K. H., Won, J. H., Drennan, W. R., Jameyson, E., Miyasaki, G., Norton, S. J., & Rubinstein, J. T. (2012). Psychoacoustic performance and music and speech perception in prelingually deafened children with cochlear implants. *Audiology & Neuro-Otology*, 17(3), 189–197. <https://doi.org/10.1159/000336407>
- [84] Kasturi, K., Loizou, P. C., Dorman, M., & Spahr, T. (2002). The intelligibility of speech with “holes” in the spectrum. *The Journal of the Acoustical Society of America*, 112(3 Pt 1), 1102–1111.
- [85] Kates, J. M. (2013). Improved estimation of frequency importance functions. *The Journal of the Acoustical Society of America*, 134(5), EL459-464. <https://doi.org/10.1121/1.4826148>
- [86] Kenway, B., Tam, Y. C., Vanat, Z., Harris, F., Gray, R., Birchall, J., ... Axon, P. (2015). Pitch Discrimination: An Independent Factor in Cochlear Implant Performance Outcomes. *Otology & Neurotology: Official Publication of the American Otological Society*,

- American Neurotology Society [and] European Academy of Otolology and Neurotology, 36(9), 1472–1479. <https://doi.org/10.1097/MAO.0000000000000845>
- [87] Kirby, B. J., Browning, J. M., Brennan, M. A., Spratford, M., & McCreery, R. W. (2015). Spectro-temporal modulation detection in children. *The Journal of the Acoustical Society of America*, 138(5), EL465-468. <https://doi.org/10.1121/1.4935081>
- [88] Koch, D. B., Osberger, M. J., Segel, P., & Kessler, D. (2004). HiResolution and conventional sound processing in the HiResolution bionic ear: using appropriate outcome measures to assess speech recognition ability. *Audiology & Neuro-Otology*, 9(4), 214–223. <https://doi.org/10.1159/000078391>
- [89] Koch, D. B., Quick, A., Osberger, M. J., Saoji, A., & Litvak, L. (2014). Enhanced hearing in noise for cochlear implant recipients: clinical trial results for a commercially available speech-enhancement strategy. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otolology and Neurotology*, 35(5), 803–809. <https://doi.org/10.1097/MAO.0000000000000301>
- [90] Koning, R., & Wouters, J. (2016). Speech onset enhancement improves intelligibility in adverse listening conditions for cochlear implant users. *Hearing Research*, 342, 13–22. <https://doi.org/10.1016/j.heares.2016.09.002>
- [91] Kral, A., Hartmann, R., Mortazavi, D., & Klinke, R. (1998). Spatial resolution of cochlear implants: the electrical field and excitation of auditory afferents. *Hearing Research*, 121(1–2), 11–28.
- [92] Kral, A., Tillein, J., Heid, S., Hartmann, R., & Klinke, R. (2005). Postnatal cortical development in congenital auditory deprivation. *Cerebral Cortex (New York, N.Y.: 1991)*, 15(5), 552–562. <https://doi.org/10.1093/cercor/bhh156>
- [93] Lammers, M. J. W., Jansen, T. T. G., Grolman, W., Lenarz, T., Versnel, H., van Zanten, G. A., ... Lesinski-Schiedat, A. (2015). The influence of newborn hearing screening on the age at cochlear implantation in children. *The Laryngoscope*, 125(4), 985–990. <https://doi.org/10.1002/lary.25045>
- [94] Landsberger, D. M., Padilla, M., Martinez, A. S., & Eisenberg, L. S. (2017). Spectral-temporal modulated tipple discrimination by Children with cochlear implants. *Ear and Hearing*, 1. <https://doi.org/10.1097/AUD.0000000000000463>

- [95] Lawler, M., Yu, J., & Aronoff, J. M. (2017). Comparison of the Spectral-Temporally Modulated Ripple Test with the Arizona Biomedical Institute Sentence Test in cochlear implant users. *Ear and Hearing*, 38(6), 760–766. <https://doi.org/10.1097/AUD.0000000000000496>
- [96] Leake, P. A., & Hradek, G. T. (1988). Cochlear pathology of long term neomycin induced deafness in cats. *Hearing Research*, 33(1), 11–33.
- [97] Leake, P. A., Hradek, G. T., & Snyder, R. L. (1999). Chronic electrical stimulation by a cochlear implant promotes survival of spiral ganglion neurons after neonatal deafness. *The Journal of Comparative Neurology*, 412(4), 543–562.
- [98] Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 49(2), Suppl 2:467+.
- [99] Li, P. M. M. C., Somdas, M. A., Eddington, D. K., & Nadol, J. B. (2007). Analysis of intracochlear new bone and fibrous tissue formation in human subjects with cochlear implants. *The Annals of Otology, Rhinology, and Laryngology*, 116(10), 731–738. <https://doi.org/10.1177/000348940711601004>
- [100] Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54(5), 358–368.
- [101] Lindholm, J. M., Dorman, M., Taylor, B. E., & Hannley, M. T. (1988). Stimulus factors influencing the identification of voiced stop consonants by normal-hearing and hearing-impaired adults. *The Journal of the Acoustical Society of America*, 83(4), 1608–1614.
- [102] Litvak, L. M., Spahr, A. J., Saoji, A. A., & Fridman, G. Y. (2007). Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. *The Journal of the Acoustical Society of America*, 122(2), 982–991. <https://doi.org/10.1121/1.2749413>
- [103] Loizou, P. C. (1999). Introduction to cochlear implants. *IEEE Engineering in Medicine and Biology Magazine: The Quarterly Magazine of the Engineering in Medicine & Biology Society*, 18(1), 32–42.
- [104] Loizou, Philipos C. (2006). Speech processing in vocoder-centric cochlear implants. *Advances in Oto-Rhino-Laryngology*, 64, 109–143. <https://doi.org/10.1159/000094648>

- [105] Long, C. J., Holden, T. A., McClelland, G. H., Parkinson, W. S., Shelton, C., Kelsall, D. C., & Smith, Z. M. (2014). Examining the electro-neural interface of cochlear implant users using psychophysics, CT scans, and speech understanding. *Journal of the Association for Research in Otolaryngology: JARO*, 15(2), 293–304. <https://doi.org/10.1007/s10162-013-0437-5>
- [106] Makary, C. A., Shin, J., Kujawa, S. G., Liberman, M. C., & Merchant, S. N. (2011). Age-related primary cochlear neuronal degeneration in human temporal bones. *Journal of the Association for Research in Otolaryngology: JARO*, 12(6), 711–717. <https://doi.org/10.1007/s10162-011-0283-2>
- [107] Maxon, A. B., & Hochberg, I. (1982). Development of psychoacoustic behavior: sensitivity and discrimination. *Ear and Hearing*, 3(6), 301–308.
- [108] Mayo, C., & Turk, A. (2004). Adult-child differences in acoustic cue weighting are influenced by segmental context: children are not always perceptually biased toward transitions. *The Journal of the Acoustical Society of America*, 115(6), 3184–3194.
- [109] Mayo, C., & Turk, A. (2005). The influence of spectral distinctiveness on acoustic cue weighting in children's and adults' speech perception. *The Journal of the Acoustical Society of America*, 118(3 Pt 1), 1730–1741.
- [110] Mens, L. H. M. (2007). Advances in cochlear implant telemetry: evoked neural responses, electrical field imaging, and technical integrity. *Trends in Amplification*, 11(3), 143–159. <https://doi.org/10.1177/1084713807304362>
- [111] Mens, L. H. M., & Berenstein, C. K. (2005). Speech perception with mono- and quadrupolar electrode configurations: a crossover study. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 26(5), 957–964.
- [112] Miller, G. A., & Nicely, P. E. (1955). An Analysis of Perceptual Confusions Among Some English Consonants. *The Journal of the Acoustical Society of America*, 27(2), 338–352. <https://doi.org/10.1121/1.1907526>
- [113] Miura, M., Sando, I., Hirsch, B. E., & Orita, Y. (2002). Analysis of spiral ganglion cell populations in children with normal and pathological ears. *The Annals of Otology, Rhinology, and Laryngology*, 111(12 Pt 1), 1059–1065.

- [114] Moberly, A. C., Lowenstein, J. H., Tarr, E., Caldwell-Tarr, A., Welling, D. B., Shahin, A. J., & Nittrouer, S. (2014). Do adults with cochlear implants rely on different acoustic cues for phoneme perception than adults with normal hearing? *Journal of Speech, Language, and Hearing Research: JSLHR*, 57(2), 566–582. [https://doi.org/10.1044/2014\\_JSLHR-H-12-0323](https://doi.org/10.1044/2014_JSLHR-H-12-0323)
- [115] Moore, D. R. (2002). Auditory development and the role of experience. *British Medical Bulletin*, 63, 171–181.
- [116] Moore, J. K., & Guan, Y.-L. (2001). Cytoarchitectural and Axonal Maturation in Human Auditory Cortex. *JARO: Journal of the Association for Research in Otolaryngology*, 2(4), 297–311. <https://doi.org/10.1007/s101620010052>
- [117] Moore, J. K., & Linthicum, F. H. (2007). The human auditory system: a timeline of development. *International Journal of Audiology*, 46(9), 460–478. <https://doi.org/10.1080/14992020701383019>
- [118] Morrongiello, B. A., Robson, R. C., Best, C. T., & Clifton, R. K. (1984). Trading relations in the perception of speech by 5-year-old children. *Journal of Experimental Child Psychology*, 37(2), 231–250.
- [119] Müller, V., Klünter, H., Fürstenberg, D., Meister, H., Walger, M., & Lang-Roth, R. (2018). Examination of prosody and timbre perception in adults with cochlear implants comparing different fine structure coding strategies. *American Journal of Audiology*, 1–11. [https://doi.org/10.1044/2017\\_AJA-17-0046](https://doi.org/10.1044/2017_AJA-17-0046)
- [120] Munson, B., Donaldson, G. S., Allen, S. L., Collison, E. A., & Nelson, D. A. (2003). Patterns of phoneme perception errors by listeners with cochlear implants as a function of overall speech perception ability. *The Journal of the Acoustical Society of America*, 113(2), 925–935.
- [121] Nearey, T. M., & Assmann, P. F. (1986). Modeling the role of inherent spectral change in vowel identification. *The Journal of the Acoustical Society of America*, 80(5), 1297–1308. <https://doi.org/10.1121/1.394433>
- [122] Nelson, D. A., Van Tasell, D. J., Schroder, A. C., Soli, S., & Levine, S. (1995). Electrode ranking of “place pitch” and speech recognition in electrical hearing. *The Journal of the Acoustical Society of America*, 98(4), 1987–1999.

- [123] Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N.-Y., Quittner, A. L., ... CDaCI Investigative Team. (2010). Spoken language development in children following cochlear implantation. *Journal of the American Medical Association*, 303(15), 1498–1506. <https://doi.org/10.1001/jama.2010.451>
- [124] Nittrouer, S. (2002). Learning to perceive speech: how fricative perception changes, and how it stays the same. *The Journal of the Acoustical Society of America*, 112(2), 711–719.
- [125] Nittrouer, S. (2007). Dynamic spectral structure specifies vowels for children and adults. *The Journal of the Acoustical Society of America*, 122(4), 2328–2339. <https://doi.org/10.1121/1.2769624>
- [126] Nittrouer, S., Caldwell-Tarr, A., Moberly, A. C., & Lowenstein, J. H. (2014). Perceptual weighting strategies of children with cochlear implants and normal hearing. *Journal of Communication Disorders*, 52, 111–133. <https://doi.org/10.1016/j.jcomdis.2014.09.003>
- [127] Noble, J. H., Hedley-Williams, A. J., Sunderhaus, L., Dawant, B. M., Labadie, R. F., Camarata, S. M., & Gifford, R. H. (2016). Initial results with image-guided cochlear implant programming in children. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 37(2), e63-69. <https://doi.org/10.1097/MAO.0000000000000909>
- [128] Noble, J. H., Labadie, R. F., Gifford, R. H., & Dawant, B. M. (2013). Image-guidance enables new methods for customizing cochlear implant stimulation strategies. *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society*, 21(5), 820–829. <https://doi.org/10.1109/TNSRE.2013.2253333>
- [129] Nogueira, W., Rode, T., & Büchner, A. (2016). Spectral contrast enhancement improves speech intelligibility in noise for cochlear implants. *The Journal of the Acoustical Society of America*, 139(2), 728–739. <https://doi.org/10.1121/1.4939896>
- [130] O'Brien, E, and Winn, MB (2017). Aliasing of spectral ripple through CI processors: a challenge to the interpretation of correlation with speech recognition scores. Poster presented at the Conference on Implantable Auditory Prostheses, Tahoe, CA. T21b.

- [131] Oh, S.-H., Kim, C.-S., Kang, E. J., Lee, D. S., Lee, H. J., Chang, S. O., ... Koo, J. W. (2003). Speech perception after cochlear implantation over a 4-year time period. *Acta Otolaryngologica*, 123(2), 148–153.
- [132] Otte, J., Schunknecht, H. F., & Kerr, A. G. (1978). Ganglion cell populations in normal and pathological human cochleae. Implications for cochlear implantation. *The Laryngoscope*, 88(8 Pt 1), 1231–1246. <https://doi.org/10.1288/00005537-197808000-00004>
- [133] Padilla, M., Stupak, N., & Landsberger, D. M. (2017). Pitch ranking with different virtual channel configurations in electrical hearing. *Hearing Research*, 348, 54–62. <https://doi.org/10.1016/j.heares.2017.02.013>
- [134] Park, M.-H., Won, J. H., Horn, D. L., & Rubinstein, J. T. (2015). Acoustic temporal modulation detection in normal-hearing and cochlear implanted listeners: effects of hearing mechanism and development. *Journal of the Association for Research in Otolaryngology: JARO*, 16(3), 389–399. <https://doi.org/10.1007/s10162-014-0499-z>
- [135] Parkinson, A. J., el-Kholy, W., & Tyler, R. S. (1998). Vowel perception in prelingually deafened children with multichannel cochlear implants. *Journal of the American Academy of Audiology*, 9(3), 179–190.
- [136] Pelliccia, P., Venail, F., Bonafé, A., Makeieff, M., Iannetti, G., Bartolomeo, M., & Mondain, M. (2014). Cochlea size variability and implications in clinical practice. *Acta Otorhinolaryngologica Italica: Organo Ufficiale Della Societa Italiana Di Otorinolaringologia E Chirurgia Cervico-Facciale*, 34(1), 42–49.
- [137] Peter, V., Wong, K., Narne, V. K., Sharma, M., Purdy, S. C., & McMahon, C. (2014). Assessing spectral and temporal processing in children and adults using temporal modulation transfer function (TMTF), Iterated Ripple Noise (IRN) perception, and spectral ripple discrimination (SRD). *Journal of the American Academy of Audiology*, 25(2), 210–218. <https://doi.org/10.3766/jaaa.25.2.9>
- [138] Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, 24(2), 175–184. <https://doi.org/10.1121/1.1906875>

- [139] Pfungst, B. E., & Xu, L. (2005). Psychophysical metrics and speech recognition in cochlear implant users. *Audiology & Neuro-Otology*, 10(6), 331–341. <https://doi.org/10.1159/000087350>
- [140] Pfungst, B. E., Xu, L., & Thompson, C. S. (2004). Across-site threshold variation in cochlear implants: relation to speech recognition. *Audiology & Neuro-Otology*, 9(6), 341–352. <https://doi.org/10.1159/000081283>
- [141] Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000). Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 111(2), 220–236.
- [142] Ramsden, R., Greenham, P., O’Driscoll, M., Mawman, D., Proops, D., Craddock, L., ... Pickerill, M. (2005). Evaluation of bilaterally implanted adult subjects with the nucleus 24 cochlear implant system. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 26(5), 988–998.
- [143] Ranck, J. B. (1975). Which elements are excited in electrical stimulation of mammalian central nervous system: a review. *Brain Research*, 98(3), 417–440.
- [144] Rayes, H., Sheft, S., & Shafiro, V. (2014). Discrimination of static and dynamic spectral patterns by children and young adults in relationship to speech perception in noise. *Audiology Research*, 4(1). <https://doi.org/10.4081/audiores.2014.101>
- [145] Reeder, R. M., Firszt, J. B., Holden, L. K., & Strube, M. J. (2014). A longitudinal study in adults with sequential bilateral cochlear implants: time course for individual ear and bilateral performance. *Journal of Speech, Language, and Hearing Research: JSLHR*, 57(3), 1108–1126. [https://doi.org/10.1044/2014\\_JSLHR-H-13-0087](https://doi.org/10.1044/2014_JSLHR-H-13-0087)
- [146] Remus, J. J., Throckmorton, C. S., & Collins, L. M. (2007). Expediting the identification of impaired channels in cochlear implants via analysis of speech-based confusion matrices. *IEEE Transactions on Bio-Medical Engineering*, 54(12), 2193–2204.
- [147] Repp, B. H. (1982). Phonetic trading relations and context effects: new experimental evidence for a speech mode of perception. *Psychological Bulletin*, 92(1), 81–110.
- [148] Robinson, E. J., Davidson, L. S., Uchanski, R. M., Brenner, C. M., & Geers, A. E. (2012). A longitudinal study of speech perception skills and device characteristics of adolescent

- cochlear implant users. *Journal of the American Academy of Audiology*, 23(5), 341–349. <https://doi.org/10.3766/jaaa.23.5.5>
- [149] Roman, A. S., Pisoni, D. B., Kronenberger, W. G., & Faulkner, K. F. (2017). Some neurocognitive correlates of noise-vocoded speech perception in children with normal hearing: A replication and extension of Eisenberg et al. (2002). *Ear and Hearing*, 38(3), 344–356. <https://doi.org/10.1097/AUD.0000000000000393>
- [150] Ruffin, C. V., Kronenberger, W. G., Colson, B. G., Henning, S. C., & Pisoni, D. B. (2013). Long-term speech and language outcomes in prelingually deaf children, adolescents and young adults who received cochlear implants in childhood. *Audiology & Neuro-Otology*, 18(5), 289–296. <https://doi.org/10.1159/000353405>
- [151] Ruffin, C. V., Tyler, R. S., Witt, S. A., Dunn, C. C., Gantz, B. J., & Rubinstein, J. T. (2007). Long-term performance of Clarion 1.0 cochlear implant users. *The Laryngoscope*, 117(7), 1183–1190. <https://doi.org/10.1097/MLG.0b013e318058191a>
- [152] Sagi, E., Meyer, T. A., Kaiser, A. R., Teoh, S. W., & Svirsky, M. A. (2010). A mathematical model of vowel identification by users of cochlear implants. *The Journal of the Acoustical Society of America*, 127(2), 1069–1083. <https://doi.org/10.1121/1.3277215>
- [153] Saoji, A. A., Litvak, L., Spahr, A. J., & Eddins, D. A. (2009). Spectral modulation detection and vowel and consonant identifications in cochlear implant listeners. *The Journal of the Acoustical Society of America*, 126(3), 955–958. <https://doi.org/10.1121/1.3179670>
- [154] Schwartz-Leyzac, K. C., & Pfungst, B. E. (2016). Across-site patterns of electrically evoked compound action potential amplitude-growth functions in multichannel cochlear implant recipients and the effects of the interphase gap. *Hearing Research*, 341, 50–65. <https://doi.org/10.1016/j.heares.2016.08.002>
- [155] Sek, A., Alcántara, J., Moore, B. C. J., Kluk, K., & Wicher, A. (2005). Development of a fast method for determining psychophysical tuning curves. *International Journal of Audiology*, 44(7), 408–420.
- [156] Seyyedi, M., & Nadol, J. B. (2014). Intracochlear inflammatory response to cochlear implant electrodes in humans. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of*

Otology and Neurotology, 35(9), 1545–1551.  
<https://doi.org/10.1097/MAO.0000000000000540>

- [157] Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science (New York, N.Y.)*, 270(5234), 303–304.
- [158] Shannon, Robert V., Galvin, J. J., & Baskent, D. (2002). Holes in hearing. *Journal of the Association for Research in Otolaryngology: JARO*, 3(2), 185–199.
- [159] Sharma, A., Kraus, N., McGee, T. J., & Nicol, T. G. (1997). Developmental changes in P1 and N1 central auditory responses elicited by consonant-vowel syllables. *Electroencephalography and Clinical Neurophysiology*, 104(6), 540–545.
- [160] Shepard, R. N. (1972). “Psychological representation of speech sounds,” in *Human Communication: A Unified View*, edited by E. David and D. P. Denes. Cambridge, UK: Cambridge University Press, 67–113.
- [161] Singh, N. C., & Theunissen, F. E. (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. *The Journal of the Acoustical Society of America*, 114(6 Pt 1), 3394–3411.
- [162] Skinner, M. W., Holden, T. A., Whiting, B. R., Voie, A. H., Brunsdon, B., Neely, J. G., ... Finley, C. C. (2007). In vivo estimates of the position of advanced bionics electrode arrays in the human cochlea. *The Annals of Otology, Rhinology & Laryngology. Supplement*, 197, 2–24.
- [163] Snyder, R. L., Bierer, J. A., & Middlebrooks, J. C. (2004). Topographic spread of inferior colliculus activation in response to acoustic and intracochlear electric stimulation. *Journal of the Association for Research in Otolaryngology: JARO*, 5(3), 305–322.  
<https://doi.org/10.1007/s10162-004-4026-5>
- [164] Spelman, F. A., Clopton, B. M., & Pflugst, B. E. (1982). Tissue impedance and current flow in the implanted ear. Implications for the cochlear prosthesis. *The Annals of Otology, Rhinology & Laryngology. Supplement*, 98, 3–8.
- [165] Srinivasan, A. G., Padilla, M., Shannon, R. V., & Landsberger, D. M. (2013). Improving speech perception in noise with current focusing in cochlear implant users. *Hearing Research*, 299, 29–36. <https://doi.org/10.1016/j.heares.2013.02.004>
- [166] Stafford, R. C., Stafford, J. W., Wells, J. D., Loizou, P. C., & Keller, M. D. (2014). Vocoder simulations of highly focused cochlear stimulation with limited dynamic range

- and discriminable steps. *Ear and Hearing*, 35(2), 262–270.  
<https://doi.org/10.1097/AUD.0b013e3182a768e8>
- [167] Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *The Journal of the Acoustical Society of America*, 64(5), 1358–1368.
- [168] Stickney, G. S., Loizou, P. C., Mishra, L. N., Assmann, P. F., Shannon, R. V., & Opie, J. M. (2006). Effects of electrode design and configuration on channel interactions. *Hearing Research*, 211(1–2), 33–45. <https://doi.org/10.1016/j.heares.2005.08.008>
- [169] Supin, A. Y., Popov, V. V., Milekhina, O. N., & Tarakanov, M. B. (1994). Frequency resolving power measured by rippled noise. *Hearing Research*, 78(1), 31–40. [https://doi.org/10.1016/0378-5955\(94\)90041-8](https://doi.org/10.1016/0378-5955(94)90041-8)
- [170] Svirsky, M. A., Teoh, S.-W., & Neuburger, H. (2004). Development of language and speech perception in congenitally, profoundly deaf children as a function of age at cochlear implantation. *Audiology & Neuro-Otology*, 9(4), 224–233. <https://doi.org/10.1159/000078392>
- [171] Syrdal, A. K., & Gopal, H. S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *The Journal of the Acoustical Society of America*, 79(4), 1086–1100. <https://doi.org/10.1121/1.393381>
- [172] Talarico, M., Abdilla, G., Aliferis, M., Balazic, I., Giaprakis, I., Stefanakis, T., ... Paolini, A. G. (2007). Effect of age and cognition on childhood speech in noise perception abilities. *Audiology & Neuro-Otology*, 12(1), 13–19. <https://doi.org/10.1159/000096153>
- [173] ter Keurs, M., Festen, J. M., & Plomp, R. (1992). Effect of spectral envelope smearing on speech reception. I. *The Journal of the Acoustical Society of America*, 91(5), 2872–2880.
- [174] Throckmorton, C. S., & Collins, L. M. (2002). The effect of channel interactions on speech recognition in cochlear implant subjects: predictions from an acoustic model. *The Journal of the Acoustical Society of America*, 112(1), 285–296.
- [175] Tobey, E. A., Thal, D., Niparko, J. K., Eisenberg, L. S., Quittner, A. L., Wang, N.-Y., & CDaCI Investigative Team. (2013). Influence of implantation age on school-age language performance in pediatric cochlear implant users. *International Journal of Audiology*, 52(4), 219–229. <https://doi.org/10.3109/14992027.2012.759666>

- [176] Tyler, R. S., Fryauf-Bertschy, H., Kelsay, D. M., Gantz, B. J., Woodworth, G. P., & Parkinson, A. (1997). Speech perception by prelingually deaf children using cochlear implants. *Otolaryngology--Head and Neck Surgery: Official Journal of American Academy of Otolaryngology-Head and Neck Surgery*, 117(3 Pt 1), 180–187.
- [177] Vanpoucke, F. J., Zarowski, A. J., & Peeters, S. A. (2004). Identification of the impedance model of an implanted cochlear prosthesis from intracochlear potential measurements. *IEEE Transactions on Bio-Medical Engineering*, 51(12), 2174–2183. <https://doi.org/10.1109/TBME.2004.836518>
- [178] Vongpaisal, T., Trehub, S. E., Glenn Schellenberg, E., & van Lieshout, P. (2012). Age-related changes in talker recognition with reduced spectral cues. *The Journal of the Acoustical Society of America*, 131(1), 501–508. <https://doi.org/10.1121/1.3669978>
- [179] Wake, M., Hughes, E. K., Poulakis, Z., Collins, C., & Rickards, F. W. (2004). Outcomes of children with mild-profound congenital hearing loss at 7 to 8 years: a population study. *Ear and Hearing*, 25(1), 1–8. <https://doi.org/10.1097/01.AUD.0000111262.12219.2F>
- [180] Walley, A. C., & Carrell, T. D. (1983). Onset spectra and formant transitions in the adult's and child's perception of place of articulation in stop consonants. *The Journal of the Acoustical Society of America*, 73(3), 1011–1022.
- [181] Wang, M. D., & Bilger, R. C. (1973). Consonant confusions in noise: a study of perceptual features. *The Journal of the Acoustical Society of America*, 54(5), 1248–1266.
- [182] Wang, N.-Y., Eisenberg, L. S., Johnson, K. C., Fink, N. E., Tobey, E. A., Quittner, A. L., ... CDaCI Investigative Team. (2008). Tracking development of speech recognition: longitudinal data from hierarchical assessments in the Childhood Development after Cochlear Implantation Study. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 29(2), 240–245. <https://doi.org/10.1097/MAO.0b013e3181627a37>
- [183] White, M. W., Merzenich, M. M., & Gardi, J. N. (1984). Multichannel cochlear implants. Channel interactions and processor design. *Archives of Otolaryngology (Chicago, Ill.: 1960)*, 110(8), 493–501.

- [184] Winn, M. B., Chatterjee, M., & Idsardi, W. J. (2012). The use of acoustic cues for phonetic identification: effects of spectral degradation and electric hearing. *The Journal of the Acoustical Society of America*, 131(2), 1465–1479. <https://doi.org/10.1121/1.3672705>
- [185] Winn, M. B., Edwards, J. R., & Litovsky, R. Y. (2015). The Impact of auditory spectral resolution on listening effort revealed by pupil dilation. *Ear and Hearing*, 36(4), e153-165. <https://doi.org/10.1097/AUD.0000000000000145>
- [186] Winn, M. B., & Litovsky, R. Y. (2015). Using speech sounds to test functional spectral resolution in listeners with cochlear implants. *The Journal of the Acoustical Society of America*, 137(3), 1430–1442. <https://doi.org/10.1121/1.4908308>
- [187] Winn, M. B., Rhone, A. E., Chatterjee, M., & Idsardi, W. J. (2013). The use of auditory and visual context in speech perception by listeners with normal hearing and listeners with cochlear implants. *Frontiers in Psychology*, 4, 824. <https://doi.org/10.3389/fpsyg.2013.00824>
- [188] Winn, M. B., Won, J. H., & Moon, I. J. (2016). Assessment of spectral and temporal resolution in cochlear implant users using psychoacoustic discrimination and speech cue categorization. *Ear and Hearing*, 37(6), e377–e390. <https://doi.org/10.1097/AUD.0000000000000328>
- [189] Won, J. H., Clinard, C. G., Kwon, S., Dasika, V. K., Nie, K., Drennan, W. R., ... Rubinstein, J. T. (2011). Relationship between behavioral and physiological spectral-ripple discrimination. *Journal of the Association for Research in Otolaryngology: JARO*, 12(3), 375–393. <https://doi.org/10.1007/s10162-011-0257-4>
- [190] Won, J. H., Drennan, W. R., & Rubinstein, J. T. (2007). Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. *Journal of the Association for Research in Otolaryngology: JARO*, 8(3), 384–392. <https://doi.org/10.1007/s10162-007-0085-8>
- [191] Won, J. H., Jones, G. L., Moon, I. J., & Rubinstein, J. T. (2015). Spectral and temporal analysis of simulated dead regions in cochlear implants. *Journal of the Association for Research in Otolaryngology: JARO*, 16(2), 285–307. <https://doi.org/10.1007/s10162-014-0502-8>

- [192] Wright, R., & Souza, P. (2012). Comparing identification of standardized and regionally valid vowels. *Journal of Speech, Language, and Hearing Research: JSLHR*, 55(1), 182–193. [https://doi.org/10.1044/1092-4388\(2011/10-0278\)](https://doi.org/10.1044/1092-4388(2011/10-0278))
- [193] Xu, L., Thompson, C. S., & Pfingst, B. E. (2005). Relative contributions of spectral and temporal cues for phoneme recognition. *The Journal of the Acoustical Society of America*, 117(5), 3255–3267.
- [194] Zaar, J., & Dau, T. (2015). Sources of variability in consonant perception of normal-hearing listeners. *The Journal of the Acoustical Society of America*, 138(3), 1253–1267. <https://doi.org/10.1121/1.4928142>
- [195] Zeng, F.-G., Rebscher, S., Harrison, W., Sun, X., & Feng, H. (2008). Cochlear implants: system design, integration, and evaluation. *IEEE Reviews in Biomedical Engineering*, 1, 115–142. <https://doi.org/10.1109/RBME.2008.2008250>
- [196] Zhang, J., Tyler, R., Ji, H., Dunn, C., Wang, N., Hansen, M., & Gantz, B. (2015). Speech, Spatial and Qualities of Hearing Scale (SSQ) and Spatial Hearing Questionnaire (SHQ) changes over time in adults with simultaneous cochlear implants. *American Journal of Audiology*, 24(3), 384–397. [https://doi.org/10.1044/2015\\_AJA-14-0074](https://doi.org/10.1044/2015_AJA-14-0074)
- [197] Zhou, N. (2017). Deactivating stimulation sites based on low-rate thresholds improves spectral ripple and speech reception thresholds in cochlear implant users. *The Journal of the Acoustical Society of America*, 141(3), EL243. <https://doi.org/10.1121/1.4977235>
- [198] Zhou, N., Kraft, C. T., Colesa, D. J., & Pfingst, B. E. (2015). Integration of pulse trains in humans and guinea pigs with cochlear implants. *Journal of the Association for Research in Otolaryngology: JARO*, 16(4), 523–534. <https://doi.org/10.1007/s10162-015-0521-0>
- [199] Zwolan, T. A., Collins, L. M., & Wakefield, G. H. (1997). Electrode discrimination and speech recognition in postlingually deafened adult cochlear implant subjects. *The Journal of the Acoustical Society of America*, 102(6), 3673–3685.