



Published in final edited form as:

Ear Hear. 2019 ; 40(2): 295–311. doi:10.1097/AUD.0000000000000615.

Reducing Simulated Channel Interaction Reveals Differences in Phoneme Identification between Children and Adults with Normal Hearing

Kelly N. Jahn¹, Mishaela DiNino¹, and Julie G. Arenberg¹

¹Department of Speech and Hearing Sciences, University of Washington, Seattle, Washington, USA

Abstract

Objectives—Channel interaction, the stimulation of overlapping populations of auditory neurons by distinct cochlear implant (CI) channels, likely limits the speech perception performance of CI users. This study examined the role of vocoder-simulated channel interaction in the ability of children and adults with normal hearing to recognize spectrally degraded speech. The primary aim was to determine the interaction between number of processing channels and degree of simulated channel interaction on phoneme identification performance as a function of age for children with normal hearing (cNH), and to relate those findings to adults with normal hearing (aNH) and to CI users.

Design—Medial vowel and consonant identification of cNH (age 8-17 years) and young aNH were assessed under six (for children) or nine (for adults) different conditions of spectral degradation. Stimuli were processed using a noise-band vocoder with 8, 12, and 15 channels and synthesis filter slopes of 15 (aNH only), 30 and 60 dB/octave (all NH subjects). Steeper filter slopes (larger numbers) simulated less electrical current spread and therefore less channel interaction. Spectrally degraded performance of the NH listeners was also compared to the unprocessed phoneme identification of school-aged children and adults with CIs.

Results—Spectrally degraded phoneme identification improved as a function of age for cNH. For vowel recognition, cNH exhibited an interaction between the number of processing channels and vocoder filter slope, whereas aNH did not. Specifically, for cNH, increasing the number of processing channels only improved vowel identification in the steepest filter slope condition. Additionally, cNH were more sensitive to changes in filter slope. As the filter slopes increased, cNH continued to receive vowel identification benefit beyond where aNH performance plateaued or reached ceiling. For all NH participants, consonant identification improved with increasing filter slopes, but was unaffected by the number of processing channels. Although cNH made more phoneme identification errors overall, their phoneme error patterns were similar to aNH.

Address correspondence to Kelly N. Jahn, University of Washington, Department of Speech and Hearing Sciences, 1417 NE 42nd Street, Box 354875, Seattle, WA 98105; jahnk@uw.edu.

Conflicts of Interest: No conflicts of interest are declared.

Portions of this work were presented at the 2017 annual convention of the American Auditory Society (AAS) in Scottsdale, AZ, the 2017 Conference on Implantable Auditory Prostheses (CIAP) in Lake Tahoe, CA, and the 2017 American Cochlear Implant Alliance (ACIA) 15th Pediatric Symposium on Cochlear Implants in Children in San Francisco, CA.

Furthermore, consonant identification of adults with CIs was comparable to aNH listening to simulations with shallow filter slopes (15 dB/octave). Vowel identification of earlier-implanted pediatric ears was better than that of later-implanted ears, and more comparable to cNH listening in conditions with steep filter slopes (60 dB/octave).

Conclusions—Recognition of spectrally degraded phonemes improved when simulated channel interaction was reduced, particularly for children. cNH showed an interaction between number of processing channels and filter slope for vowel identification. The differences observed between cNH and aNH suggest that identification of spectrally degraded phonemes continues to improve through adolescence, and that children may benefit from reduced channel interaction beyond where adult performance has plateaued. Comparison to CI users suggests that early implantation may facilitate development of better phoneme discrimination.

Keywords

cochlear implant; speech perception; channel interaction; spectral resolution; vocoder; pediatric

Introduction

Cochlear implants (CIs) are highly successful neural prostheses that can restore auditory perception for individuals with severe-to-profound sensorineural hearing loss. Despite this success, children and adults with CIs exhibit substantial variability in speech perception and spoken language outcomes (e.g., Niparko et al. 2010; Holden et al. 2013). One factor that likely contributes to this variability is reduced spectral resolution resulting from a limited number of implanted electrodes. Moreover, although modern CIs transmit acoustic information using 12 to 22 electrical contacts, speech perception performance of high-performing adults with CIs (aCI) does not improve beyond stimulation of 7-8 channels (Dorman and Loizou, 1997; Fishman et al. 1997; Friesen et al. 2001; Nie et al. 2006).

On the other hand, acoustic simulations of CI processing show that enhancing the degree of spectral resolution can improve speech perception performance of adults with normal hearing (aNH) beyond where aCI performance plateaus (e.g., Shannon et al. 1995; Dorman et al. 1997; Dorman and Loizou, 1997; Fishman et al. 1997; Loizou et al. 1999; Friesen et al. 2001; Shannon et al. 2004; Nie et al. 2006). For instance, while sentence recognition asymptotes around four channels for aNH, identification of spectrally degraded speech in the presence of background noise can continue to improve as the number of channels is increased up to twenty (Shannon et al. 1995; Dorman et al. 1998; Friesen et al. 2001; Shannon et al. 2004).

Overall, this evidence from adult listeners suggests that many CI recipients cannot take advantage of all the spectral information provided by the implant and that a small number of implanted electrodes is not the only factor that contributes to spectral smearing. Another likely contributor to poor spectral resolution with an implant is channel interaction, which occurs when electrical current from distinct implant channels stimulates overlapping populations of auditory neurons. Channel interaction theoretically limits the number of independent channels of information available to a listener, smearing the spectral detail provided by a CI (for review, see Bierer 2010). Accordingly, several investigations have used

noise-band vocoder processing to assess how simulating varying degrees of electrical current spread impacts the speech perception performance of aNH (Fu and Nogaki 2004; Litvak et al. 2007b; Bingabr et al. 2008; Stafford et al. 2014).

In those studies, the attenuation slopes of the vocoder synthesis filters were manipulated to simulate different degrees of electrical current spread. Specifically, steeper filter slopes represented less spectral overlap between adjacent vocoder filters and therefore less channel interaction. Results showed that reducing the simulated spread of neural excitation led to improved sentence-in-noise and phoneme identification for aNH (Fu and Nogaki 2004; Litvak et al. 2007b; Bingabr et al. 2008; Stafford et al. 2014), as long as spectral resolution was adequate (>8 channels; Bingabr et al. 2008). Further, extremely steep filter slopes (>100 dB/octave) were often detrimental to speech perception (Warren et al. 2004; Healy and Steinbach 2007; Bingabr et al. 2008), especially when there were a small number of processing channels (Bingabr et al. 2008). This evidence supports the idea that the interaction between electrical current spread and the number of processing channels can affect transmission of spectral information and, consequently, speech perception outcomes.

While spectrally degraded speech perception of aNH has been studied extensively, less is known about the ability of children with normal hearing (cNH) to process CI-simulated speech. To date, vocoder studies with cNH have primarily focused on the number of spectral channels needed to achieve high levels of speech perception. Like adults, vocoded speech perception performance of cNH improves as the number of channels increases (Eisenberg et al. 2000; Dorman et al. 2000; Nittrouer and Lowenstein 2009; Nittrouer et al. 2010; Newman and Chatterjee 2013; Warner-Czyz et al. 2014; Newman et al. 2015). Perhaps more importantly, however, these investigations have shown that the ability to recognize spectrally degraded speech improves with age for cNH, and that young children require more channels than older children and adults to achieve asymptotic speech perception performance (Dorman et al. 2000; Eisenberg et al. 2000, 2002; Vongpaisal et al. 2012; Roman et al. 2017).

Because children require greater spectral resolution than adults to accurately recognize vocoded speech (e.g., Eisenberg et al. 2000; Dorman et al. 2000), channel interaction might be more detrimental to the speech perception performance of children than it is for adults. Furthermore, a pre-lingually deafened child must rely upon the degraded auditory signal from a CI to develop spoken language, and spectral smearing from channel interaction may further limit speech and language development. It is crucial, then, to understand how channel interaction impacts speech perception in children. To date, the effects of varying the degree of CI-simulated channel interaction by changing the vocoder filter slopes has not been investigated in cNH, and little is known about the effects of channel interaction in the pediatric CI population.

Accordingly, the primary goal of the present study was to characterize the perceptual implications of simulated channel interaction in school-aged cNH, and to relate those findings to aNH. Specifically, the interaction between the number of processing channels and filter slope on phoneme identification performance was evaluated using a noise-band vocoder technique similar to that adopted in prior investigations with adults (e.g., Litvak et

al. 2007b; Bingabr et al. 2008). Based on previous evidence, it was expected that phoneme identification performance would improve for NH listeners as the vocoder filter slopes increased, and that this improvement would be more pronounced in conditions with more processing channels. Moreover, because children rely more on spectral information than adults in vocoder studies, it was hypothesized that children would show more phoneme identification improvement than adults in response to increases in filter slope. We also expected that spectrally degraded phoneme identification performance would improve with age for the cNH, in accordance with the development of higher-order auditory and executive processes that extends into adolescence (Gomes et al. 2000; Moore 2002; Roman et al. 2017).

Of note, previous vocoder studies with cNH have used monosyllabic words and sentences with preschool-level vocabulary as the target speech stimuli (Dorman et al. 2000; Eisenberg et al. 2000, 2002; Roman et al. 2017). Spectro-temporally complex stimuli such as words and sentences are less susceptible to spectral degradations than simpler stimuli, as they can be identified using a variety of acoustic-phonetic cues or top-down lexico-semantic processes (Shannon et al. 2004; McGettigan et al. 2007). Thus, phoneme identification (i.e., medial vowels and consonants) was selected as the speech perception outcome of interest in the present study for several reasons. First, spectral manipulation of vowels and consonants can be well-controlled (DiNino et al. 2016), which is of particular interest to the current investigation. Specifically, accurate vowel identification depends strongly on an individual's ability to resolve formant frequency information, making vowel perception particularly vulnerable to spectral degradation (Xu et al. 2005; Nie et al. 2006; Gaudrain et al. 2007; DiNino et al. 2016). Consequently, in vocoder studies, adult listeners require a larger number of channels to reach asymptotic vowel identification performance relative to more complex stimuli such as words and sentences (Dorman et al. 1998; Friesen et al. 2001; Shannon et al. 2004).

On the other hand, a listener can use both spectral and temporal cues to achieve accurate consonant identification (Xu et al. 2005; Nie et al. 2006; DiNino et al. 2016). Despite consonants being more robust to spectral distortion than vowels, perception of consonant place of articulation depends primarily on spectral cues, whereas manner and voicing features can be distinguished by temporal cues (Rosen 1992; Xu et al. 2005; Nie et al. 2006; DiNino et al. 2016). Importantly, temporal cues that are important for conveying voicing and manner information are preserved in noise-band vocoder processing (e.g., Shannon et al. 1995). This distinction allows for analysis of how the transmission of specific consonant acoustic-phonetic features is uniquely distorted by spectral degradation, a process that has not yet been evaluated in vocoder studies with cNH.

Accordingly, the present study will also assess how spectral degradation affects phoneme confusions and the transmission of consonant features for cNH. Consistent with prior studies, it was expected that consonant place of articulation would be the most vulnerable feature to spectral degradation, and that consonants would be confused most often with others of the same manner and voicing, but different place of articulation (Miller and Nicely 1955; Johnson 2000; Xu et al. 2005; DiNino et al. 2016). Further, it was expected that lax vowels would be more susceptible to the spectral degradation than cardinal vowels, as they

have more neighbors in acoustic space for which to be confused (DiNino et al. 2016). We did not expect differences in phoneme error patterns across experimental conditions, as the vocoder manipulations were designed to globally alter the spectral information in the signal, rather than selectively distort information in specific frequency regions.

Additionally, it was hypothesized that vowel and consonant confusions would be similar for children and adults, but that children would make more errors overall. This is consistent with phoneme feature analyses that have been conducted with cNH under other listening conditions that were designed to distort signal information (i.e., noise and reverberation; Johnson 2000). If phoneme confusions are similar across age groups, this could suggest that differences in non-sensory cognitive processing might account for any observed differences in overall performance.

Finally, because the ultimate goal of this research is to understand factors limiting speech and language outcomes for individuals with CIs, a preliminary comparison between NH listeners and CI users was conducted. The CI users in this study completed phoneme identification tasks while participating in other experiments in our laboratory. Since many lines of research suggest that channel interaction limits speech outcomes for CI listeners, it was predicted that the CI users' phoneme identification performance would be comparable to that of NH participants listening in conditions with relatively shallow vocoder filter slopes, representing relatively high levels of simulated channel interaction. This would corroborate the spirit of the noise-band vocoder technique as a tool for assessing how spectral manipulations relate to CI performance.

Materials and Methods

A. Participants

Twenty-eight cNH initially enrolled in the study, but one child (age 10 years) was ultimately excluded due to an inability to accurately identify the phonemic stimuli (see *Procedures* for further detail). Participants with NH included 12 adults (5 males) between the ages of 21 and 35 years ($M = 27.35$ years, $SD = 4.58$ years) and 27 children (13 males) between the ages of 8 and 17 years ($M = 12.96$ years, $SD = 3.08$ years). To ensure similar representation across age, at least two cNH were recruited from each year between age 8 and 17. A hearing screening was performed on all NH participants to verify thresholds at 20 dB hearing level (HL) or better at octave frequencies 250 to 8000 Hz bilaterally (ANSI 1996).

Participants with CIs included 10 children (5 males) between the ages of 13 and 17 years ($M = 14.34$ years, $SD = 1.82$ years) and 14 adults (8 males) between the ages of 37 and 84 years ($M = 61.53$ years, $SD = 12.78$ years). The cCI were pre-lingually deafened, with onset of severe-to-profound bilateral sensorineural hearing loss prior to three years of age, except for one subject (subject P05). That subject's second-implanted ear was diagnosed with profound hearing loss at age 5. Each cCI was bilaterally implanted, and each child received his or her first CI by age 4 years. The aCI were all post-lingually deafened, except for one pre-lingually deafened subject (self-report onset of severe-to-profound hearing loss at age 4). All but one adult were unilaterally implanted. Each ear of the bilaterally implanted participants was tested separately, yielding an effective sample size of $N = 15$ aCI ears and $N = 20$ cCI

ears. Each CI participant used an Advanced Bionics HiRes 90K device. Demographic information for the cCI and aCI is shown in Tables 1 and 2, respectively.

Participants were recruited from the University of Washington campus in Seattle, Washington and the surrounding community, were native speakers of American English, and used spoken language to communicate. All cCI were enrolled in a mainstream educational setting. All but one participant had no self-report history of neurological impairments or developmental delay. One cCI (subject P01) reportedly had mild Asperger's syndrome. Each adult participant provided written informed consent. Children gave informed written assent and their parents or guardians provided informed written consent. Participants were paid for their time. Experimental procedures were approved by the University of Washington Human Subjects Division.

B. Speech stimuli

Phoneme identification was assessed with medial vowels and consonants. Vowel stimuli consisted of a closed set of ten recorded vowels presented in /hVd/ context (/i/, “heed”; /ɪ/, “hid”; /e/, “hayed”; /ɛ/, “head”; /æ/, “had”; /a/, “hod”; /u/, “who'd”; /ʊ/, “hood”; /oʊ/, “hoed”; /ʌ/, “hud”) that were recorded from one female talker native to the Pacific Northwest region of the United States. Consonant stimuli consisted of a closed set of sixteen recorded consonants presented in /aCa/ context (/p/, “aPa”; /t/, “aTa”; /k/, “aKa”; /b/, “aBa”; /d/, “aDa”; /g/, “aGa”; /f/, “aFa”; /θ/, “aTHa”; /s/, “aSa”; /ʃ/, “aSHa”; /v/, “aVa”; /z/, “aZa”; /dʒ/, “aJa”; /m/, “aMa”; /n/, “aNa”; /l/, “aLa”) that were recorded from one male talker (consonant stimuli were the same as those used by Shannon et al. 1999, which were created by Tyler, Preece, and Lowder at the University of Iowa Department of Otolaryngology, 1989). Speech stimuli were presented to NH participants under various vocoder-manipulated conditions, using a technique described in the following section. The CI users listened to unprocessed speech stimuli through their personal sound processors.

C. Vocoder signal processing

For NH listeners, CI processing was simulated acoustically using a noise-band vocoder technique adapted from Litvak et al. (2007), which was designed to model the Fidelity 120 signal processing strategy employed in current Advanced Bionics devices (Advanced Bionics Corp., Valencia, CA). Acoustic speech stimuli were digitally sampled at 17,400 Hz and, depending on the experimental condition, bandpass filtered into 8, 12 or 15 pseudo-logarithmically-spaced frequency bands spanning a frequency range of 250-8700 Hz. The analysis filters were FFT-based, similar to those used in Advanced Bionics speech processing strategies (Litvak, personal communication). The output of each filter was input into an envelope detector with a low-pass filter cutoff of 68 Hz. The extracted envelopes were used to modulate a broadband noise with a center frequency identical to that of the corresponding analysis band. The slope of the attenuation of the noise-band filters was either 15, 30 or 60 dB/octave, depending on the experimental condition. The modulated noise bands were summed and presented to NH listeners via a loudspeaker.

Spectral information was manipulated in two ways to create a total of nine possible experimental vocoder conditions. The number of processing channels, either 8, 12 or 15, was

changed to simulate varying degrees of spectral resolution, with a smaller number of channels (i.e., 8) representing poorer spectral resolution. Table 3 outlines the frequency allocations for each channel in each spectral resolution condition. Additionally, the slope of the synthesis filter outputs was manipulated to simulate varying degrees of channel interaction; that is, a steeper filter slope (larger number; i.e., 60 dB/octave) represented less spectral overlap between adjacent channels and therefore less channel interaction, whereas a shallower filter slope (smaller number; i.e., 15 dB/octave) represented greater spectral overlap between adjacent channels and therefore more channel interaction. These channels and filter slopes were selected to achieve a difficulty level appropriate for both children and adults, based on parameters used in our lab and in prior vocoder studies (e.g., Eisenberg et al. 2000; Fu and Nogaki 2005; Litvak et al. 2007b; Bingabr et al. 2008; DiNino et al., 2016). Consequently, the filter slopes were not designed to necessarily mimic the intracochlear excitation patterns observed in response to broad or more spatially-restricted CI electrode configurations. Adults were tested in all nine experimental conditions (3 processing channel conditions \times 3 filter slope conditions). To avoid floor effects and significant fatigue, the children did not complete testing in the 15 dB/octave filter slope condition, yielding a total of six experimental conditions for the cNH (3 processing channel conditions \times 2 filter slope conditions).

D. Procedures

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 a calibrated level of 60 dB-A through a Bose 161 speaker placed at 0° azimuth and 1 meter from the participant's head. Custom software (ListPlayer2 version 2.2.11.52, Advanced Bionics, Valencia, CA) was used to present the stimuli and to record participant responses. Following presentation of each speech token, a graphical user interface with the possible /hVd/ or /aCa/ choices was displayed on a computer screen in front of the participant, and he or she selected his or her response using a computer mouse.

Vowels and consonants were presented in separate blocks, and the order of blocks was randomized for each participant. Prior to each experimental block, the participants completed at least one, and a maximum of three, practice lists consisting of one presentation of each unprocessed vowel or consonant token. To demonstrate understanding of the stimuli and the ability to accurately read the answer choices prior to testing, NH participants were required to achieve 100% correct in the unprocessed condition before moving forward with the study. As previously stated, one cNH (age 10 years) could only achieve 67% correct across three unprocessed practice runs and was therefore unable to complete the study. Following the unprocessed practice runs, participants completed one vocoded (with 15 channels and a 60 dB/octave filter slope) practice list consisting of one repetition of each vowel or consonant token. To familiarize participants with the vocoded stimuli and the task prior to experimental testing, they could repeat the practice tokens as many times as desired, and performance feedback was provided.

Experimental runs were initiated following the unprocessed and vocoded practice runs. Each vowel and consonant speech token was repeated six times, allowing collection of six data

points per token per experimental condition. Children completed two runs of three repetitions each. As adults participated in three more experimental conditions than the children, they completed three runs consisting of two repetitions to avoid fatigue. However, during pilot testing, three children completed three runs consisting of two repetitions each, and three adults completed two runs consisting of three repetitions each. A technical error led to one child completing all six repetitions of vowels within the same run. Despite these differences, the participants' data fell within the range of scores for their respective age groups. Furthermore, removing those data from statistical analyses did not change the results.

Within each block, stimuli were pseudo-randomly interleaved to minimize systematic confounding effects of order, fatigue, and practice across experimental conditions. A percent correct score was calculated for each experimental condition for both vowels and consonants. Feedback was not provided during the experimental runs and participants were not allowed to repeat presentations of any of the speech tokens. Participants were provided with ample breaks and refreshments as needed throughout testing.

CI listeners completed testing with the same sound-booth set-up, stimulus level (60 dB-A), and computer interface as the NH listeners. However, the CI users only listened to the speech stimuli in the unprocessed condition using their personal sound processors. To familiarize the participants with the testing procedure and the stimuli, each CI user completed at least one practice run prior to experimental testing. The cCI only completed vowel identification testing because they were participating in a larger study of spectral resolution in our lab, for which vowels were the speech stimuli of interest. Due to time constraints, we could not add a consonant identification task to the cCI test battery. The aCI completed both vowel and consonant identification testing. Bilateral CI listeners were tested with one implant at a time. For unilateral CI users, an earplug was placed in the non-implanted ear to avoid possible contribution from any residual hearing.

E. Analyses

Statistical analyses were performed using IBM SPSS Statistics (IBM Corp., 2016). Data for the NH participants were analyzed using a repeated measures analysis-of-variance (ANOVA), with a 3 (filter slope) \times 3 (channel) design for the adults and a 2 (filter slope) \times 3 (channel) design for the children. Because many aNH performed near ceiling and because the CI data were highly variable, percent correct scores were converted to rationalized arcsine units (RAUs) for statistical analyses to normalize the variance across groups and across conditions (Studebaker 1985). After converting to RAUs, the assumption of homogeneity of variance was still violated when comparing NH to CI data, so equal variances were not assumed for those analyses. Planned follow-up simple effects *t*-tests with Dunn-Sidak adjustments for multiple comparisons were implemented to assess differences in overall phoneme identification across conditions (adjusted $\alpha = 0.0057$ for 9 comparisons and $\alpha = 0.0085$ for 6 comparisons). Separate analyses were conducted for vowel and consonant identification measures for each age group. In cases where Mauchly's test of sphericity was significant, Greenhouse-Geisser adjusted *F*-tests are reported. Mean

performance (in percent correct) and standard deviations for all NH conditions are reported in Table 2. ANOVA results are reported in Table 3.

In addition to comparing average performance across conditions, phoneme confusions and the identification of consonant acoustic-phonetic features were assessed for cNH and adults. To perform these analyses, phoneme error patterns were compared across age groups using confusion matrices, sequential information analysis (SINFA; Wang and Bilger 1973) and perceptual distance (Zaar and Dau 2015). Together, these analyses provide a quantitative and qualitative comparison of the phoneme error patterns of children and adults.

Sequential information analysis (SINFA)—Identification of the consonant acoustic-phonetic features of voicing, manner, and place of articulation was quantified for each experimental condition using SINFA (Wang and Bilger 1973). SINFA, which is based on the analyses described by Miller and Nicely (1955), uses participants' phoneme confusions to determine the percent of information transmitted for each phonetic feature. Because the contributions of each feature to overall consonant identification are not independent from one another (e.g., since all nasal consonants are voiced, there is redundancy of the voicing feature across all nasal phonemes), SINFA implements an iterative analysis procedure to eliminate this internal redundancy. For example, in the first iteration of SINFA, percent of information transmitted is quantified for the feature with the most perceptual importance for identification. Then, in the second iteration, the effect of that feature is held constant while the feature with the second-highest perceptual importance is identified and quantified, and so on, until the contributions of all specified features have been determined.

In the present study, three iterations of SINFA were performed, one for each of the three consonant features of voicing, manner, and place of articulation. Consonant feature transmission was quantified for each filter slope for the NH participants and the aCI. SINFA was not performed on vowel identification, as vowel features (F1, F2, duration) are not independent of each other, and thus SINFA results are much less accurate for vowel confusion patterns than for consonant confusions (DiNino et al. 2016). The consonant feature classification criteria were adapted from Xu et al. (2005) and Miller and Nicely (1955) and are outlined in Table 4.

Phoneme confusions—To assess whether children and adults make similar phonemic errors under spectrally-degraded listening conditions, phoneme confusion matrices were quantitatively compared across age groups using a perceptual distance analysis (e.g., Zaar and Dau 2015; DiNino et al. 2016). In this analysis, each cell of two confusion matrices is compared to determine the difference in confusion patterns across phonemes. Differences can range from 0% to 100%. A perceptual difference of 0% indicates that the responses to a presented token were the same between matrices. Conversely, a perceptual difference of 100% means that the responses to a presented token were entirely different between matrices. To determine whether a difference between matrices is meaningful, the perceptual distance results are compared to a baseline value. Meaningful differences between matrices are defined as perceptual distance values that are higher than the baseline value. The baseline for this study was calculated as the perceptual distance between each subject's confusion matrices from the first three stimulus repetitions and the second three stimulus

repetitions in the 15-channel, 30 dB/octave filter slope condition. These values presumably represent listener uncertainty and this method has been used in other investigations that have employed the perceptual distance analysis (e.g., Zaar and Dau 2015; DiNino et al. 2016).

Results

A. Individual phoneme identification performance

Individual phoneme identification performance as a function of age is shown in Figures 1 and 2. Figure 1 shows the relationship between phoneme identification and age for two conditions that have the same number of channels (15 channels), but different filter slopes (30 dB/octave versus 60 dB/octave). Alternatively, Figure 2 shows this relationship for two experimental conditions that have the same filter slope (30 dB/octave) but differ in the number of channels (8 channels versus 15 channels).

The statistical significance of the relationship between phoneme identification and age was assessed for the condition with 15 channels and a 30 dB/octave filter slope only, since individual phoneme identification performance across conditions was highly correlated. This experimental condition was chosen because it has been used as a control condition in previous investigations, allowing for across-study comparisons (Litvak et al. 2007b; DiNino et al. 2016). Age significantly predicted vowel identification ($R^2 = 0.50$, $F(1, 25) = 25.20$, $p < 0.001$) and consonant identification ($R^2 = 0.26$, $F(1, 24) = 8.32$, $p = 0.008$) performance for cNH, but not for aNH ($p > 0.05$). Specifically, phoneme identification performance of the cNH improved with age.

B. Assessing the interaction between number of channels and filter slope

The main effects of number of processing channels and filter slope on phoneme identification performance, and the interaction between those variables, were assessed for three age groups: younger cNH (age 8-12 years; $N = 14$, $M = 10.34$ years, $SD = 1.50$ years), older cNH (age 13-17 years; $N = 13$, $M = 15.77$ years, $SD = 1.24$ years), and aNH. The age cut-off between younger and older children was selected as the age where the largest difference in phoneme identification performance occurred (again, in the 15-channel, 30 dB/octave filter slope condition). Specifically, performance for each pair of sequential years was averaged together and compared to the performance for the next pair of years. The largest difference in phoneme identification performance occurred between the 11-12-year-olds and the 13-14-year-olds, with differences of 11.4% (vowels) and 12.0% (consonants). On the other hand, the largest difference for any of the other comparisons was 6.4%. Please refer to Table 5 for the means and standard deviations in each condition for each age group and Table 6 for the ANOVA data. Tables 7 and 8 show the mean differences in overall phoneme identification scores between filter slopes for each number of channels, and between processing channel conditions within each filter slope, respectively.

Vowel identification performance

Adults with normal hearing: Average vowel identification performance in each condition for aNH is plotted in Figure 3A. For aNH, there were significant main effects of filter slope and number of channels on vowel identification performance, but the interaction between the

two variables was not statistically significant (Table 6). The planned multiple comparisons (adjusted $\alpha = 0.0057$) showed that vowel performance was significantly better with 12 channels than with 8 channels for the 15 dB/octave and 30 dB/octave filter slopes ($p < 0.0057$). The difference between 8 and 12 channels in the 60 dB/octave filter slope condition was not significant after adjusting for multiple comparisons. Furthermore, regardless of filter slope, vowel recognition plateaued at 12 channels; that is, performance did not improve when the number of channels was increased from 12 to 15.

Additionally, significantly poorer vowel identification performance was observed in the 15 dB/octave filter slope condition compared to 30 dB/octave, for 8, 12 and 15 channels ($p < 0.0057$). Performance plateaued in the 30 dB/octave filter slope condition; that is, across all numbers of channels tested, additional improvement in vowel identification performance was not observed when filter slope was further increased from 30 dB/octave to 60 dB/octave. While the aNH may have reached a performance ceiling at 12 channels in the 30 and 60 dB/octave filter slope conditions, note that performance also plateaued at 12 channels in the 15 dB/octave filter slope conditions, where scores were not at ceiling.

Comparison to adults with cochlear implants: Average vowel identification performance of the aCI ($M = 80.67\%$, $SD = 18.67$) is plotted in Figure 3A alongside the aNH data. Vowel performance of the aCI was better than that of aNH listening in the 15 dB/octave filter slope conditions, and comparable to aNH performance in the 30 and 60 dB/octave filter slope conditions, irrespective of the number of vocoded processing channels ($p > 0.05$). Figure 4 shows individual vowel identification performance for the CI users. Note that vowel identification performance of the aCI is highly variable.

Children with normal hearing: Figure 3(B-C) plots average vowel identification performance for the older (Figure 3B) and younger (Figure 3C) cNH. For both age groups, there were significant main effects of filter slope and number of channels on vowel identification performance (Table 6). There was also a significant interaction between filter slope and number of channels for the older cNH, and this interaction approached significance for the younger cNH ($p = 0.056$). The interaction was ordinal in nature, suggesting that cNH's vowel identification performance depended on both the number of channels and the filter slope, and that the distinct effects of each of those factors was dependent upon the other.

Specifically, vowel identification performance of younger cNH did not improve as the number of channels was increased in the 30 dB/octave filter slope condition. However, when filter slope was steeper (60 dB/octave), vowel identification performance did improve when the number of channels increased from 8 to 15 ($p < 0.0085$), but not from 8 to 12. For older cNH, vowel performance was significantly better for 15 channels than for 8 channels ($p < 0.0085$) in the 30 dB/octave filter slope condition, but there was no difference in vowel identification performance between 8 and 12 channels. When filter slope was steeper (60 dB/octave), vowel identification of older cNH improved when the number of channels increased from 8 to 12 ($p < 0.0085$), and did not change when the number of channels increased from 12 to 15. Like the adults, it is possible that the older cNH reached a performance ceiling in the 12-channel, 60 dB/octave filter slope condition; however, vowel

identification also plateaued at 12 channels in the 30 dB/octave filter slope condition even though performance was not at ceiling. Furthermore, both younger and older cNH showed better vowel identification performance in the steeper filter slope condition (60 dB/octave) compared to the shallower filter slope condition (30 dB/octave) for 12 and 15 channels ($p < 0.0085$), but not for 8 channels.

Comparison to children with cochlear implants: Vowel identification performance of the cCI is shown in Figure 3B alongside the data of the age-matched older cNH. A paired *t*-test was used to compare vowel identification performance of the first-implanted ears to the performance of the second-implanted ears. Vowel identification performance of the first-implanted ears ($M = 85.60\%$, $SD = 15.90$) was significantly better than that of the second-implanted ears ($M = 70.10\%$, $SD = 31.83$), $t(9) = 2.30$, $p = 0.047$, $d = 0.61$. Additionally, the average age of implantation of the first-implanted ears ($M = 2.23$ years, $SD = 1.13$) was significantly earlier than that of the second-implanted ears ($M = 8.14$ years, $SD = 3.70$), $t(9) = 6.23$, $p < 0.001$, $d = 2.46$.

Furthermore, the first-implanted ears of the cCI performed similarly to age-matched cNH listening in the 12-channel, 60 dB/octave filter slope condition ($p > 0.05$). On the other hand, vowel identification performance of the second-implanted ears was most comparable to that of cNH listening in the 8-channel, 30 dB/octave filter slope condition ($p > 0.05$). Figure 4 shows individual phoneme identification performance for the pediatric CI ears. Each cCI is represented by a unique symbol. Like the aCI, vowel identification performance of the cCI was highly variable. All but three cCI performed worse with the second-implanted ear compared to the first-implanted ear. Finally, vowel identification performance of the first-implanted and second-implanted cCI ears was not significantly different than performance of the aCI ($p > 0.05$).

Consonant identification performance

Participants with normal hearing: One child with NH (age 9.72 years) withdrew from the study prior to completing consonant recognition testing, yielding a sample size of $N = 13$ for the younger children for consonant analyses (mean age = 10.39 years). Figure 5(A-C) shows the consonant identification data for the NH participants. For consonant identification, there were significant main effects of filter slope for all age groups. There were also significant main effects of number of channels for adults and older children. The interaction between filter slope and number of channels was not significant for adults and younger children; however, this interaction was significant for older children (see Tables 5 and 6 for descriptive statistics and ANOVA results, respectively).

Follow-up comparisons showed that the pattern of consonant recognition performance across conditions was essentially the same for all age groups. Specifically, for all participants, consonant recognition improved for all numbers of channels when the filter slope was increased from 30 dB/octave to 60 dB/octave ($p < 0.001$). For adults, consonant identification also improved for all numbers of channels when filter slope was increased from 15 dB/octave to 30 dB/octave ($p < 0.001$). Additionally, consonant recognition did not improve as the number of channels increased from 8 to 12 or from 12 to 15 for any groups.

Comparison to adults with cochlear implants: Individual consonant identification performance of the adult CI users is shown in Figure 4. Like vowel recognition, their consonant identification performance was highly variable. Figure 5A plots overall consonant performance for the adult CI users ($M = 72.20\%$, $SD = 21.01$), which was comparable to aNH in the 15 dB/octave filter slope condition, irrespective of number of processing channels ($p > 0.05$).

C. Sequential information analysis (SINFA)

Figure 6 and Table 9 show the results of the SINFA for the NH participants and the aCI. Because number of processing channels did not impact overall consonant identification, percent of information transmitted was collapsed across processing channels within each filter slope condition. The amount of information transmitted for the manner feature was high in all conditions and across all age groups, consistent with relatively good preservation of temporal envelope cues that are important for discriminating consonants on the basis of manner (Xu et al. 2005). For the place and voicing features, the amount of information transmitted increased as the filter slope increased for all age groups. Of the three features, transmission of place of articulation was most affected by the vocoder manipulations, especially for the cNH. Similarly, transmission of the place of articulation feature was poorest for aCI and was comparable to the aNH listening in the shallowest filter slope condition (15 dB/octave). However, transmission of the manner feature was not as high for the aCI as it was for the NH listeners.

D. Phoneme confusions

Phoneme confusions of the NH listeners were evaluated quantitatively via perceptual distance analysis (e.g., Zaar and Dau 2015). The within-subject baseline perceptual distances, represented by dashed lines in Figure 7, were 18.61% for vowels and 15.10% for consonants. Within each age group, perceptual distances across conditions did not exceed the baseline values, suggesting that phoneme errors were not affected by the number of channels or the filter slope. Because of this, between-group comparisons were made in the condition with 15-channels and a 30 dB/octave filter slope (Figure 7). Again, this condition was selected to allow for across-study comparisons. For each age group comparison, the perceptual distance was less than the baseline perceptual distance. This suggests that the overall phoneme confusion patterns of the children and adults were quantitatively similar.

To qualitatively assess phoneme error patterns for one condition, confusion matrices for the 15-channel, 30 dB/octave filter slope condition are presented in Figures 8 (vowels) and 9 (consonants). In each figure, the ordinate shows the phoneme that was presented, and the abscissa shows the phoneme was perceived by the listener. Cells with darker shading received a higher percentage of responses than cells with lighter shading. Shading along the diagonal indicates correct responses, whereas off-diagonal shading represents incorrect responses. In Figure 8, adjacent vowels have similar second formant frequencies, while vowels on opposite ends of the matrix have similar first formant frequencies. Listeners made the most identification errors for the vowels /eɪ/, /ʌ/, and /ʊ/ (i.e., the speech tokens HAYED, HUD and HOOD, respectively). Specifically, the vowel /eɪ/ was often confused with /i/ and /I/, whereas /ʌ/ and /ʊ/ were often confused with each other.

In Figure 9, consonants are ordered by manner of articulation. Consonant confusions tended to cluster around the diagonal, indicating that confusions occurred between tokens that were most similar in manner. Listeners made the most confusions for the consonants /p/, /b/, /k/, /g/, and /θ/ (i.e., the speech tokens APA, ABA, AKA, AGA, and ATHA, respectively). Specifically, /p/ and /k/ were often confused with /t/, /b/ was often confused with /p/, /g/ was often confused with /d/, and /θ/ was often confused with /v/. Although phoneme confusion patterns were similar across age groups, children made more errors overall than adults.

E. Test-retest comparisons

Finally, test-retest data were analyzed to assess the possible contribution of learning or fatigue effects to overall vocoder performance. The difference between performance on the first three and second three repetitions of the stimuli, in percent correct, was calculated. Overall, there was a slight improvement in phoneme identification performance over time for the NH participants. For vowels, the average difference between the first and second repetitions was 4.31% ($SD = 5.69$) for the younger cNH, 5.95% ($SD = 3.51$) for the older cNH, and 4.98% ($SD = 3.91$) for the aNH. For consonants, the average difference between the first and second repetitions was 1.90% ($SD = 5.87$) for the younger cNH, 2.53% ($SD = 4.22$) for the older cNH, and 2.92% ($SD = 2.15$) for the aNH. One-way ANOVAs revealed that there was no difference in test-retest performance between age groups for either vowel identification ($F(2, 36) = 0.45, p = 0.642$) or consonant identification ($F(2, 35) = 0.17, p = 0.844$).

Discussion

In this study, the effects of changing the noise-band vocoder filter slopes and the number of processing channels on vowel and consonant identification performance of school-aged children and adults with NH were assessed. The vocoder parameters were designed to simulate different degrees of CI channel interaction (i.e., filter slopes) and spectral resolution (i.e., number of processing channels). Phoneme identification performance was compared across degraded conditions within each age group to examine the interaction between filter slopes and number of processing channels. Phoneme confusion analyses were also conducted to assess whether cNH make the same phonemic errors as aNH under spectrally-degraded listening conditions. Finally, phoneme identification performance of the NH listeners was compared to that of children and adults with CIs.

Interaction between Filter Slope and Number of Channels

As expected, phoneme identification scores improved with age for the cNH, consistent with prior studies of spectrally-degraded speech perception in cNH (e.g., Dorman et al. 2000; Eisenberg et al. 2000, 2002; Roman et al. 2017). Further, as predicted, cNH were more sensitive to increases in filter slope than aNH for vowel identification performance. Although all NH listeners benefitted from increasing filter slope to some extent, the vowel identification scores of cNH continued to improve beyond the filter slope where adult performance reached ceiling (30 dB/octave). Of note, this occurred for both the younger cNH and the older cNH, suggesting that this difference between children and adults persists through late adolescence. While it cannot be concluded that the children received more

vowel identification *benefit* from increasing filter slopes than adults, they did experience a larger *improvement* in performance when filter slope increased. So, cNH's vowel identification performance can continue to benefit from reductions in simulated channel interaction even when adults have reached a performance plateau or ceiling.

A second major hypothesis was that there would be an interaction between number of processing channels and filter slopes on phoneme identification performance. This hypothesis was confirmed for the vowel identification performance of cNH, but not for aNH. Specifically, the effect of vocoder filter slope on children's vowel identification was mediated by the number of available processing channels. While children received vowel identification benefit from steeper filter slopes, this only occurred when the number of processing channels was relatively high (12-15 channels); conversely, children tended to benefit from more processing channels when filter slope was relatively steep (60 dB/octave), but not when the filter slope was shallow (30 dB/octave). These findings suggest that children may not be able to take full advantage of reduced spectral smearing unless there is *both* a high number of processing channels and a small amount of spectral overlap between adjacent channels. On the other hand, aNH were able to take advantage of additional processing channels (up to 12) for vowel identification even when simulated channel interaction was relatively high (i.e., shallow filter slopes).

Although previous work by Bingabr and colleagues (2008) showed an interaction between the number of processing channels and filter slope on the speech perception performance of aNH, they included a condition with fewer processing channels (4 channels) and a much steeper filter slope (110 dB/octave) than was used in the current study. Thus, an interaction between the variables may have emerged for the adults in the present investigation if we were able to include more experimental conditions. Considering the findings of Bingabr et al. (2008), our results suggest that the interaction between number of processing channels and filter slope might emerge for cNH under different conditions than what is expected for aNH.

Furthermore, vowel identification for adults improved from 8 to 12 channels regardless of the filter slope. This plateau in vowel identification performance at 12 channels for aNH is consistent with prior evidence from adult vocoder studies (Xu et al. 2005; Nie et al. 2006). The present study further demonstrates that adults' ability to use additional processing channels for vowel identification is unaffected by the degree of simulated channel interaction for the filter slopes assessed here. The interaction observed in this study for cNH shows that children's vowel identification performance may similarly plateau around 12 channels, but only when vocoder filter slopes are relatively steep.

In contrast to vowel identification, the effects of number of processing channels and filter slope on consonant identification performance were similar for all three age groups. All NH listeners demonstrated improvement in consonant identification performance as filter slope increased, and the number of processing channels (beyond 8) did not significantly affect consonant identification scores. This observation is consistent with many prior studies on adult consonant identification performance under spectrally degraded conditions, showing a

plateau in consonant scores around 7-8 channels (Shannon et al. 1995; Dorman et al. 1997; Dorman and Loizou, 1997; Fishman et al. 1997; Friesen et al. 2001; Nie et al. 2006).

However, in the present study, aNH and older cNH showed a non-significant trend for scores to improve with increasing numbers of processing channels in the steepest filter slope condition (60 dB/octave). Further, there was a significant interaction between number of processing channels and filter slope for the consonant identification scores of older cNH, though this interaction was not significant for any other group. Because consonant identification tends to plateau around 8 channels, it is possible that an interaction between number of processing channels and filter slope may have emerged for all NH listeners if we included an experimental condition with fewer than 8 processing channels. Regardless, the present study demonstrates that spectrally degraded consonant identification performance of school-aged cNH plateaus before or at 8 spectral channels, though scores may slightly improve with additional channels if the filter slope is relatively steep.

Of note, we did not attempt to directly compare consonant and vowel identification performance to each other; rather, we have assessed how the experimental parameters independently influenced performance within each set of phonemes. Evidently, for the parameters assessed in this study, differences between cNH and aNH are more pronounced for vowel identification than for consonant identification. These differences are not surprising, as consonant identification was expected to be more robust overall to spectral degradation than vowels, and prior studies with adults show that consonant identification performance plateaus with fewer spectral channels than vowel identification (e.g., Xu et al. 2005; DiNino et al. 2016). However, another important consideration is that the speaker and the gender of the speakers differed for the vowel (female) and consonant (male) stimuli in this study. It is well-established that female and male speakers differ in vocal fold size and length, which influence the fundamental frequency and the formant frequencies, respectively (e.g., Peterson and Barney 1952; Klatt and Klatt 1990; Hillenbrand et al. 1995; Hillenbrand et al. 2009). Consequently, speech perception performance for NH and CI listeners depends on the speaker and his or her gender (e.g., Fu et al. 2004; Schwartz and Chatterjee 2012; Meister et al. 2016; DiNino et al. 2016). Thus, direct comparisons between the vowel and consonant stimuli in this study were inappropriate and were not performed.

Overall, the present data show that spectrally degraded phoneme identification and the ability to optimally use additional spectral information is immature through late adolescence. Previously, Eisenberg and colleagues (2000) found that spectrally degraded recognition of pediatric words and sentences, with vocabulary appropriate for 3-5-year-olds, was mature by 12 years of age. The results of the present study extend upon those findings to show that recognition of spectrally-degraded phonemes, which rely heavily on intact spectral information and less on top-down lexico-semantic processes, might mature later than the speech tasks used in Eisenberg et al. (2000). Indeed, our data are consistent with previous evidence suggesting that phoneme recognition in challenging listening environments such as noise and reverberation continues to mature through at least age 15 years (Neuman and Hochberg 1983; Johnson 2000; Talarico et al. 2007). Furthermore, difficult listening tasks such as voice emotion (Chatterjee et al. 2015) and emotional prosody recognition (Tinnemore et al. 2018) continue to mature through age 17-19 years in cNH

when the signal is vocoded. Importantly, the present data also show that reducing spectral smearing can improve the phoneme identification of children, who may be at a developmental disadvantage when the spectral resolution of the signal is poor.

The present findings are also consistent with the literature showing that temporal resolution (e.g., Hall and Grose 1994; Stuart 2005; Hall et al. 2012; Buss et al. 2014; Buss et al. 2017) and spectral resolution (e.g., Allen and Wightman 1992; Peter et al. 2014; Rayes et al. 2014; Kirby et al. 2015; Landsberger et al. 2017; Horn et al. 2017a, b) are poorer for younger children and mature gradually into adolescence. Although temporal cues are preserved in the vocoder simulations, it is possible that younger children may not be able to take advantage of these cues to the same extent as older children and adults. A combination of immature spectral and temporal resolution would unsurprisingly limit spectrally degraded phoneme recognition relative to adults and could lead to greater improvements in performance when more spectral information is introduced. Indeed, this was observed in the present study.

Furthermore, the perception of spectrally degraded speech reflects contributions of peripheral and central auditory processes, as well as executive functioning. In cNH, the auditory cortex (e.g., Ponton et al. 2000; Moore 2002; Moore and Linthicum 2007) and auditory selective attention (e.g., Doyle 1973; Gomes et al. 2000; Werner and Boike 2001) continue to develop through late adolescence. Young children's non-sensory executive functioning may also be limited by immature short-term memory capacity, the use of less effective listening strategies, reduced motivation, and/or higher internal noise levels relative to adults (e.g., Olsho 1985; Schneider et al. 1989; Allen and Wightman 1994; Gomes et al. 2000; Buss et al. 2006; Buss et al. 2009; Buss et al. 2012). These higher-order executive processes are likely mediated by development of the pre-frontal cortex, which extends into early adulthood (e.g., Gomes et al. 2000). Each of these higher-order processes may influence development of spectrally degraded speech perception abilities.

In fact, a recent investigation by Roman and colleagues (2017) showed that measures of auditory attention and short-term memory were significantly correlated with cNH's ability to recognize vocoded words and sentences. Specifically, children who had better speech perception scores also had better auditory attention skills and short-term memory capacities than children with poorer speech perception scores. Thus, because acoustic-phonetic information is distorted by vocoder processing, a cNH must rely on non-sensory neurocognitive processes to accurately identify spectrally degraded speech. Thus, in this study, higher-order auditory and executive function may have been immature in the children relative to the adults, leading to greater cognitive demands of spectrally degraded phoneme identification for the cNH. Accordingly, future investigations may wish to explore the relation of non-sensory factors such as listening effort or reaction time to vocoded speech perception in cNH.

Similar Phoneme Confusions for Children and Adults with Normal Hearing

A secondary aim of the present study was to assess the phoneme confusion patterns of cNH compared to aNH under spectrally degraded listening conditions. As expected, phoneme error patterns were similar across age groups, even if overall accuracy differed. For example, listeners made the most identification errors for the lax vowels /eɪ/, /ʌ/, and /ʊ/. In

perceptual vowel space, lax vowels have many acoustic neighbors with similar formant frequencies. This suggests that, when frequency information is degraded, lax vowels have a larger number of confusable pairs than tense vowels. Accordingly, NH listeners made very few errors for tense vowels such as /i/, /a/, and /u/. The finding that lax vowels were more prone to errors under spectrally degraded listening conditions than tense vowels is consistent with previous adult vocoder studies (DiNino et al. 2016).

Furthermore, as predicted, NH listeners most often confused consonants that differed by place of articulation. For example, /p/, /k/, /t/ were often confused with each other, as were /g/ and /d/. It was uncommon for listeners to confuse vocoded consonants with different manners of articulation, consistent with robust encoding of the temporal envelope with vocoder processing (e.g., Shannon et al. 1995). Further, it is well-established that identification of consonant place of articulation relies on robust spectral information, whereas voicing and manner are more dependent upon temporal cues (e.g., Shannon et al. 1995; Xu et al. 2005; Nie et al. 2006; DiNino et al. 2016).

Consonant confusions were consistent with the results of the SINFA analysis, which quantified the amount of feature information transmitted for place of articulation, voicing, and manner. That is, transmission of place of articulation was most impacted by the vocoder manipulations and drove overall consonant performance. Manner information was relatively robust to spectral degradation, whereas voicing fell in between. Additionally, transmission of place of articulation improved as the vocoder filter slope increased, consistent with a reduction in spectral smearing and enhanced transmission of spectral information. This finding is consistent with prior adult vocoder studies (e.g., Shannon et al. 1995; Xu et al. 2005; DiNino et al. 2016), and with Johnson (2000), who showed that place of articulation was the feature most impacted by noise and reverberation for cNH. Thus, we show that, even though spectral degradation reduces phoneme identification scores of cNH more than those of aNH, they make the same types of phoneme confusions when the signal is distorted. These findings support the previous argument that protracted maturation of cognitive processes such as auditory attention and short-term memory may impact children's overall performance. Because NH listeners made similar errors, but the children made more errors overall, it is possible that listening to the spectrally-degraded stimuli was more cognitively demanding for the children compared to the adults.

Clinical Implications and Comparison to Cochlear Implant Users

If spectrally degraded phoneme identification and non-sensory cognitive factors continue to mature through the late teenage years in cNH, it is reasonable to expect delayed maturation of these abilities in children with congenital hearing loss. On top of this, spectrally degraded input from a CI may further hinder development of these sensory and non-sensory abilities. In this study, we compared phoneme identification performance of a small group of CI users to our NH data. This preliminary data suggested that phoneme identification performance of the second-implanted cCI ears was comparable to cNH in conditions with relatively shallow filter slopes, representing high simulated channel interaction. On the other hand, performance of the first-implanted cCI ears was comparable to cNH listening in conditions with steeper filter slopes (less simulated channel interaction). This finding suggests that

early access to CI stimulation may facilitate development of better spectral resolution abilities in children. Though beyond the scope of the present investigation, a subsequent study from our laboratory will further analyze the relationship between demographic variables and several measures of spectral resolution in cCI.

For aCI, vowel identification performance was like that of aNH listening in conditions with relatively steep filter slopes. On the other hand, consonant identification performance of the aCI was comparable to aNH in conditions with very shallow filter slopes (15 dB/octave), representing a high degree of simulated channel interaction. Because the aCI were older than the aNH (mean age = 61.53 years), it is possible that age-related deficits in temporal processing led to an inability to capitalize on available temporal cues for consonant identification (for review, see Walton 2010). In fact, the SINFA analysis revealed that the percent of information transmitted for manner was much lower for aCI than for aNH, corroborating the idea that the aCI were unable to resolve the relatively robust temporal cues for manner.

Overall, this preliminary comparison to CI listeners suggests that CI channel interaction may play a role in speech perception outcomes, especially for late-implanted cCI. Data from the NH listeners suggest that, unless simulated channel interaction is low, children may not be able to take advantage of the number of channels provided by the implant, and this finding persists through the teenage years. If cNH require a larger reduction in simulated channel interaction in conjunction with a high number of processing channels to achieve asymptotic phoneme identification performance throughout adolescence, a child relying full-time on degraded input from a CI may be expected to have a similar, if not greater, need for reduced spectral smearing.

Thus, the evidence presented here has important implications for programming CIs in the pediatric population. Children who are congenitally deafened must rely upon the degraded input from a CI to develop spoken language. It is therefore essential to design programming strategies that are tailored to the unique needs of congenitally deafened children to facilitate the development of age-appropriate oral language skills. For example, modern CI processing strategies use a spatially broad monopolar (MP) electrode configuration, which results in substantial overlap of electrical stimulation between adjacent channels (Jolly et al. 1996; Kral et al. 1998). This presumably leads to a large amount of channel interaction, and it is possible that this is not the optimal stimulation mode for pediatric CI listeners. Instead, a more spatially-focused electrode configuration such as partial tripolar (Litvak et al. 2007a) might enable cCI to use spectral information from more of their available processing channels. Although speech perception outcomes of aCI with spatially-focused programming strategies are mixed (e.g., Berenstein et al. 2008; Srinivasan et al. 2013; Bierer and Litvak 2016), our findings suggest that these results may not be generalizable to children. Thus, it may be worth exploring the speech perception outcomes of cCI using current focusing in an attempt to limit channel interaction and improve spectral resolution.

Conclusions

Overall, the present results are consistent with prior evidence showing that spectrally degraded speech perception improves with age for cNH. We also present evidence that phoneme identification of cNH can improve when the degree of simulated channel interaction is reduced, and that improvement is observed beyond the level at which adult performance plateaus or reaches ceiling (vowel identification). Furthermore, the ability of cNH to achieve vowel identification benefit from reduced simulated channel interaction depends on the number of available processing channels, whereas this interaction is not observed in aNH. Even when overall identification levels are poorer, children make similar phonemic confusions to adults for both consonants and vowels under spectrally degraded listening conditions. Comparison to CI listeners suggests that channel interaction may play a role in the speech perception performance of some children and adults with CIs. These findings highlight the importance of investigating methods to reduce spectral smearing for CI listeners, and especially for children.

Acknowledgments

The authors would like to thank Dr. Leonid Litvak and the Advanced Bionics Corporation for providing the clinical interface and for assistance with developing the vocoder manipulations. We also thank Jacob Hulswit, Ka'iulani Moku-Paiva, and Wendy Parkinson for assistance with data collection and participant recruitment. Further, we thank our research participants for their generous donation of time.

Source of Funding: This research was supported by the National Institutes of Health National Institute on Deafness and Other Communication Disorders Grant No. R01 DC012142 (J.G.A.) and the National Institutes of Health Grant No. T32 DC005361 (K.N.J., M.D., PI David Perkel).

References

- Allen P, Wightman F. 1992; Spectral pattern discrimination by children. *J Speech Hear Res.* 35:222–233. [PubMed: 1735972]
- Allen P, Wightman F. 1994; Psychometric functions for children's detection of tones in noise. *J Speech Hear Res.* 37:205–215. [PubMed: 8170124]
- American National Standards Institute (ANSI). American standard specification for audiometers. American National Standards Institute; New York: 1996.
- Berenstein CK, Mens LHM, Mulder JJS, et al. 2008; Current steering and current focusing in cochlear implants: Comparison of monopolar, tripolar, and virtual channel electrode configurations. *Ear Hear.* 29:250–260. [PubMed: 18595189]
- Bierer JA. 2010; Probing the electrode-neuron interface with focused cochlear implant stimulation. *Trends Amplif.* 14(2):84–95. [PubMed: 20724356]
- Bierer JA, Litvak L. 2016; Reducing channel interaction through cochlear implant programming may improve speech perception: Current focusing and channel deactivation. *Trends Hear.* 20:1–12.
- Bingabr M, Espinoza-Varas B, Loizou PC. 2008; Simulating the effect of spread of excitation in cochlear implants. *Hear Res.* 241(1-2):73–79. [PubMed: 18556160]
- Buss E, Hall JW III, Grose JH. 2006; Development and the role of internal noise in detection and discrimination thresholds with narrow band stimuli. *J Acoust Soc Am.* 120(5):2777–2788. [PubMed: 17139738]
- Buss E, Hall JW III, Grose JH. 2009; Psychometric functions for pure tone intensity discrimination: slope differences in school-aged children and adults. *J Acoust Soc Am.* 125(2):1050–1058. [PubMed: 19206879]

- Buss, E, Hall, JW, , III Grose, JH. Development of auditory coding as reflected in psychophysical performance. In: Werner, LA, Fay, RR, Popper, AN, editors. *Human Auditory Development*. New York, NY: Springer; 2012. 107–136.
- Buss E, Hall JW, Porter HL, et al. 2014; Gap detection in school-age children and adults: Effects of inherent envelope modulation and the availability of cues across frequency. *J Speech Lang Hear Res.* 57:1098–1107. [PubMed: 24686553]
- Buss E, Porter HL, Hall JW, et al. 2017; Gap detection in school-age children and adults: Center frequency and ramp duration. *J Speech Lang Hear Res.* 60:172–181. [PubMed: 28056469]
- Chatterjee M, Zion D, Deroche ML, et al. 2015; Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. *Hear Res.* 322:151–162. [PubMed: 25448167]
- DiNino M, Wright RA, Winn MB, et al. 2016; Vowel and consonant confusions from spectrally manipulated stimuli designed to simulate poor cochlear implant electrode-neuron interfaces. *J Acoust Soc Am.* 140(6):4404–4418. [PubMed: 28039993]
- Dorman MF, Loizou PC. 1997; Speech intelligibility as a function of the number of channels of stimulation for normal-hearing listeners and patients with cochlear implants. *Am J Otol.* 18(suppl):S113–S114. [PubMed: 9391623]
- Dorman MF, Loizou PC, 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. *J Acoust Soc Am.* 102(4):2403–2411. [PubMed: 9348698]
- Dorman MF, Loizou PC, Fitzke J, et al. 1998; The recognition of sentences in noise by normal-hearing listeners using simulations of cochlear-implant signal processors with 6-20 channels. *J Acoust Soc Am.* 104(6):3583–3585. [PubMed: 9857516]
- Dorman MF, Loizou PC, Kemp LL, et al. 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 Hear.* 21:590–596. [PubMed: 11132785]
- Doyle AB. 1973; Listening to distraction: a developmental study of selective attention. *J Exp Child Psychol.* 15:100–115. [PubMed: 4706960]
- Eisenberg LS, Shannon RV, Schaefer Martinez A, et al. 2000; Speech recognition with reduced spectral cues as a function of age. *J Acoust Soc Am.* 107(5):2704–2710. [PubMed: 10830392]
- Eisenberg LS, Martinez AS, Holowecky SR, et al. 2002; Recognition of lexically controlled words and sentences by children with normal hearing and children with cochlear implants. *Ear Hear.* 23:450–462. [PubMed: 12411778]
- Fielden CA, Kluk K, McKay CM. 2013; Place specificity of monopolar and tripolar stimuli in cochlear implants: the influence of residual masking. *J Acoust Soc Am.* 133:4109–4123. [PubMed: 23742363]
- Fishman KE, Shannon RV, Slattery WH. 1997; Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. *J Speech Lang Hear Res.* 40:1201–1215. [PubMed: 9328890]
- Friesen LM, Shannon RV, Baskent D, et al. 2001; Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. *J Acoust Soc Am.* 110(2):1150–1163. [PubMed: 11519582]
- Fu Q, Nogaki G. 2005; Noise susceptibility of cochlear implant users: the role of spectral resolution and smearing. *J Assoc Res Otolaryngol.* 6:19–27. [PubMed: 15735937]
- Gaudrain E, Grimhault N, Healy EW, et al. 2007; Effect of spectral smearing on the perceptual segregation of vowel sequences. *Hear Res.* 231:32–41. [PubMed: 17597319]
- Gomes H, Molholm S, Christodoulou C, Ritter W, Cowan N. 2000; The development of auditory attention in children. *Front Biosci.* 5:d108–120. [PubMed: 10702373]
- Hall JW, Grose JH. 1994; Development of temporal resolution in children as measured by the temporal modulation transfer function. *J Acoust Soc Am.* 96(1):150–154. [PubMed: 7598757]
- Hall JA, Buss E, Grose JH, et al. 2012; Effects of age and hearing impairment on the ability to benefit from temporal and spectral modulation. *Ear Hear.* 33:340–348. [PubMed: 22237164]
- Healy EW, Steinbach HM. 2007; The effect of smoothing filter slope and spectral frequency on temporal speech information. *J Acoust Soc Am.* 121(2):1177–1181. [PubMed: 17348538]

- Hillenbrand J, Getty LA, Clark MJ, et al. 1995; Acoustic characteristics of American English vowels. *J Acoust Soc Am.* 97:3099–3111. [PubMed: 7759650]
- Hillenbrand JM, Clark MJ. 2009; The role of f_0 and formant frequencies in distinguishing the voices of men and women. *Atten Percept Psychophys.* 71(5):1150–1166. [PubMed: 19525544]
- Holden LK, Finley CC, Firszt JB, et al. 2013; Factors affecting open-set word recognition in adults with cochlear implants. *Ear Hear.* 34:342–360. [PubMed: 23348845]
- Horn DL, Won JH, Rubinstein JT, et al. 2017a; Spectral ripple discrimination in normal-hearing infants. *Ear Hear.* 38:212–222. [PubMed: 27768611]
- Horn DL, Dudley DJ, Dedhia K, et al. 2017b; Effects of age and hearing mechanism on spectral resolution in normal hearing and cochlear implanted listeners. *J Acoust Soc Am.* 141(1):613–623. [PubMed: 28147578]
- Johnson CE. 2000; Children's phoneme identification in reverberation and noise. *J Speech Lang Hear Res.* 43:144–157. [PubMed: 10668658]
- Jolly C, Spelman FA, Clopton BM. 1996; Quadrupolar stimulation for cochlear prostheses: modeling and experimental data. *IEEE Trans Biomed Eng.* 43(8):857–865. [PubMed: 9216159]
- Kirby BJ, Browning JM, Brennan MA, et al. 2015; Spectro-temporal modulation detection in children. *J Acoust Soc Am.* 138(5):EL465–EL468. [PubMed: 26627815]
- Klatt DH, Klatt LC. 1990; Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J Acoust Soc Am.* 87:820–857. [PubMed: 2137837]
- Kral A, Hartmann R, Mortazavi D, et al. 1998; Spatial resolution of cochlear implants: the electrical field and excitation of auditory afferents. *Hear Res.* 121:11–28. [PubMed: 9682804]
- Landsberger DM, Padilla M, Martinez AS, et al. 2017; Spectral-temporal modulated ripple discrimination by children with cochlear implants. *Ear Hear.* doi: 10.1097/AUD.0000000000000463
- Litvak LM, Spahr AJ, Emadi G. 2007a; Loudness growth observed under partially tripolar stimulation: model and data from cochlear implant listeners. *J Acoust Soc Am.* 122(2):967–981. [PubMed: 17672645]
- Litvak LM, Spahr AJ, Saoji AA, et al. 2007b; Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. *J Acoust Soc Am.* 122(2):982–991. [PubMed: 17672646]
- Loizou PC, Dorman M, Tu Z. 1999; On the number of channels needed to understand speech. *J Acoust Soc Am.* 106(4):2097–2103. [PubMed: 10530032]
- McGettigan C, Rosen S, Scott SK. 2014; Lexico-semantic and acoustic-phonetic processes in the perception of noise-vocoded speech: implications for cochlear implantation. *Front Syst Neurosci.* 8(18):1–11. [PubMed: 24478639]
- Meister H, Fursen K, Streicher B, et al. 2016; The use of voice cues for speaker gender recognition in cochlear implant recipients. *J Speech Lang Hear Res.* 59:546–556. [PubMed: 27135985]
- Miller GA, Nicely PE. 1955; An analysis of the perceptual confusions among some English consonants. *J Acoust Soc Am.* 27(2):338–352.
- Moore DR. 2002; Auditory development and the role of experience. *Br Med Bull.* 63:171–181. [PubMed: 12324392]
- Moore JK, Linthicum FH. 2007; The human auditory system: a timeline of development. *Int J Audiol.* 46:460–478. [PubMed: 17828663]
- Neuman AC, Hochberg I. 1983; Children's perception of speech in reverberation. *J Acoust Soc Am.* 73(6):2145–2149. [PubMed: 6875100]
- Newman R, Chatterjee M. 2013; Toddlers' recognition of noise-vocoded speech. *J Acoust Soc Am.* 133(1):483–494. [PubMed: 23297920]
- Newman RS, Chatterjee M, Morini G, et al. 2015; Toddlers' comprehension of degraded signals: noise-vocoded versus sine-wave analogs. *J Acoust Soc Am.* 138(3):EL311–317. [PubMed: 26428832]
- Nie K, Barco A, Zeng FG. 2006; Spectral and temporal cues in cochlear implant speech perception. *Ear Hear.* 27(2):208–217. [PubMed: 16518146]

- Nilsson M, Soli S, Sullivan J. 1994; Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and noise. *J Acoust Soc Am.* 95:1085–1099. [PubMed: 8132902]
- Niparko JK, Tobey EA, Thal DJ, et al. 2010; Spoken language development in children following cochlear implantation. *JAMA.* 303(15):1498–1506. [PubMed: 20407059]
- Nittrouer S, Lowenstein JH, Packer R. 2009; Children discover the spectral skeletons in their native language before the amplitude envelope. *J Exp Psychol Hum Percept Perform.* 35(4):1245–1253. [PubMed: 19653762]
- Nittrouer S, Lowenstein JH. 2010; Learning to perceptually organize speech signals in native fashion. *J Acoust Soc Am.* 127(3):1624–1635. [PubMed: 20329861]
- Olsho LW. 1985; Infant auditory perception: tonal masking. *Infant Behav Dev.* 8:371–384.
- Peter V, Wong K, Sharma M, et al. 2014; Assessing spectral and temporal processing in children and adult using temporal modulation transfer function (TMTF), iterated ripple noise (IRN) perception, and spectral ripple discrimination (SRD). *J Am Acad Audiol.* 25:210–218. [PubMed: 24828221]
- Peterson GE, Barney HI. 1952; Control methods used in a study of the vowels. *J Acoust Soc Am.* 24:75–184.
- Ponton CW, Eggermont JJ, Kwong B, et al. 2000; Maturation of human central auditory system activity: evidence from multichannel evoked potentials. *Clin Neurophysiol.* 111(2):220–236. [PubMed: 10680557]
- Rayes H, Sheft S, Shafiro S. 2014; Discrimination of static and dynamic spectral patterns by children and young adults in relationship to speech perception in noise. *Audiol Res.* 101(4):28–35.
- Roman AS, Pisoni DB, Kronenberger WG, et al. 2017; Some neurocognitive correlates of noise-vocoded speech perception in children with normal hearing: A replication and extension of Eisenberg et al. (2002). *Ear Hear.* 38:344–356. [PubMed: 28045787]
- Rosen SA. 1992; Temporal information in speech: acoustic, auditory and linguistic aspects. *Phil Trans R Soc Lond.* 336:367–373. [PubMed: 1354376]
- Schneider BA, Trehub SE, Morrongiello BA, et al. 1989; Developmental changes in masked thresholds. *J Acoust Soc Am.* 86:1733–1742. [PubMed: 2808922]
- Schvartz KC, Chatterjee M. 2012; Gender identification in younger and older adults: Use of spectral and temporal cues in noise-vocoded speech. *Ear Hear.* 33(3):411–420. [PubMed: 22237163]
- Shannon RV, Zeng FG, Kamath V, et al. 1995; Speech recognition with primarily temporal cues. *Science.* 270:303–304. [PubMed: 7569981]
- Shannon RV, Jansvold A, Padilla M, et al. 1999; Consonant recordings for speech testing. *J Acoust Soc Am.* 106(6):L71–L74. [PubMed: 10615713]
- Shannon RV, Fu QJ, Galvin J III. 2004; The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. *Acta Otolaryngol. Suppl* 552:50–54.
- Srinivasan AG, Padilla M, Shannon RV, et al. 2013; Improving speech perception in noise with current focusing in cochlear implant users. *Hear Res.* 299:29–36. [PubMed: 23467170]
- Stafford RC, Stafford JW, Wells JD, et al. 2014; Vocoder simulations of highly focused cochlear stimulation with limited dynamic range and discriminable steps. *Ear Hear.* 35:262–270. [PubMed: 24322978]
- Stuart A. 2005; Development of auditory temporal resolution in school-age children revealed by word recognition in continuous and interrupted noise. *Ear Hear.* 26:78–88. [PubMed: 15692306]
- Studebaker GA. 1985; A “rationalized” arcsine transform. *J Speech Hear Res.* 28:455–462. [PubMed: 4046587]
- Talarico M, Abdilla G, Aliferis M, et al. 2007; Effect of age and cognition on childhood speech in noise perception abilities. *Audiol Neurotol.* 12(1):13–19.
- Tinnemore AR, Zion DJ, Kulkarni AM, et al. 2018; Children's recognition of emotional prosody in spectrally degraded speech is predicted by their age and cognitive status. *Ear Hear.* doi: 10.1097/AUD.0000000000000546

- Tyler, R, Preece, J, Tye-Murray, N. The Iowa audio-visual speech perception laser videodisc Laser videodisc and Laboratory Report. Iowa City, IA: Department of Otolaryngology, Head and Neck Surgery, University of Iowa Hospital and Clinics; 1986.
- Vongpaisal T, Trehub SE, Glenn Schellenberg E, et al. 2012; Age-related changes in talker recognition with reduced spectral cues. *J Acoust Soc Am.* 131(1):501–508. [PubMed: 22280611]
- Walton JP. 2010; Timing is everything: temporal processing deficits in the aged auditory brainstem. *Hear Res.* 264:63–69. [PubMed: 20303402]
- Wang MD, Bilger RC. 1973; Consonant confusions in noise: A study of perceptual features. *J Acoust Soc Am.* 54:1248–1266. [PubMed: 4765809]
- Warner-Czyz AD, Houston DM, Hynan LS. 2014; Vowel discrimination by hearing infants as a function of number of spectral channels. *J Acoust Soc Am.* 135(5):3017–3024. [PubMed: 24815281]
- Warren RM, Bashford JA, Lenz PW. 2004; Intelligibility of bandpass filtered speech: steepness of slopes required to eliminate transition band contributions. *J Acoust Soc Am.* 115(3):1292–1295. [PubMed: 15058351]
- Werner LA, Boike K. 2001; Infants' sensitivity to broadband noise. *J Acoust Soc Am.* 109(5):2103–2111. [PubMed: 11386562]
- Xu L, Thompson CS, Pfingst BE. 2005; Relative contributions of spectral and temporal cues for phoneme recognition. *J Acoust Soc Am.* 117(5):3255–3267. [PubMed: 15957791]

Abbreviations/Acronyms

cNH	children with normal hearing
aNH	adults with normal hearing
cCI	children with cochlear implants
aCI	adults with cochlear implants
RAUs	rationalized arcsine units
ANOVA	analysis of variance
SINFA	sequential information analysis

Phoneme identification as a function of age and filter slope

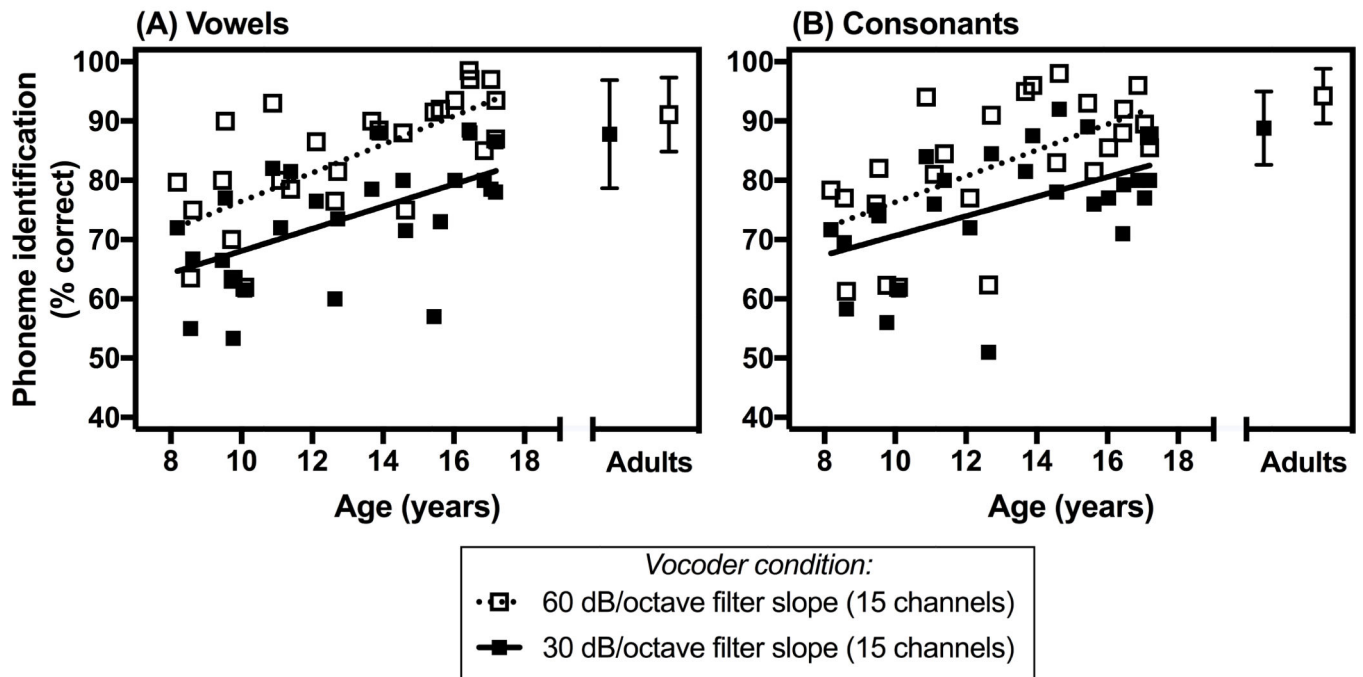


Figure 1. Phoneme identification scores (in percent correct) as a function of chronological age and filter slope for children with normal hearing. Average scores for adults with normal hearing are also presented. A) Vowel identification scores. B) Consonant identification scores. In each panel, scores are shown for two experimental conditions that have the same number of processing channels (15 channels), but different filter slopes (30 dB/octave and 60 dB/octave). For the adult data, error bars represent ± 1 standard deviation.

Phoneme identification as a function of age and number of channels

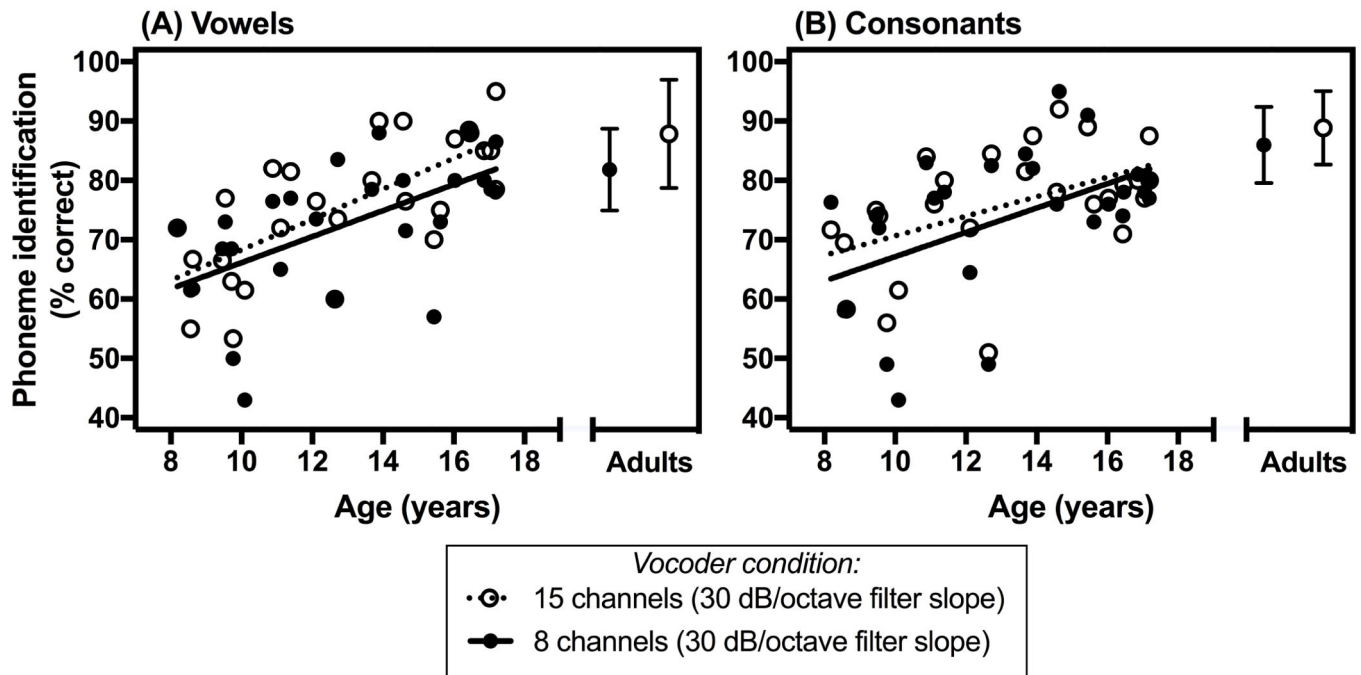


Figure 2.

Phoneme identification scores (in percent correct) as a function of chronological age and number of processing channels for children with normal hearing. Average scores for adults with normal hearing are also presented. (A) Vowel identification scores. (B) Consonant identification scores. In each panel, scores are shown for two experimental conditions that have the same filter slope (30 dB/octave), but different number of processing channels (8 channels and 15 channels). For the adult data, error bars represent ± 1 standard deviation.

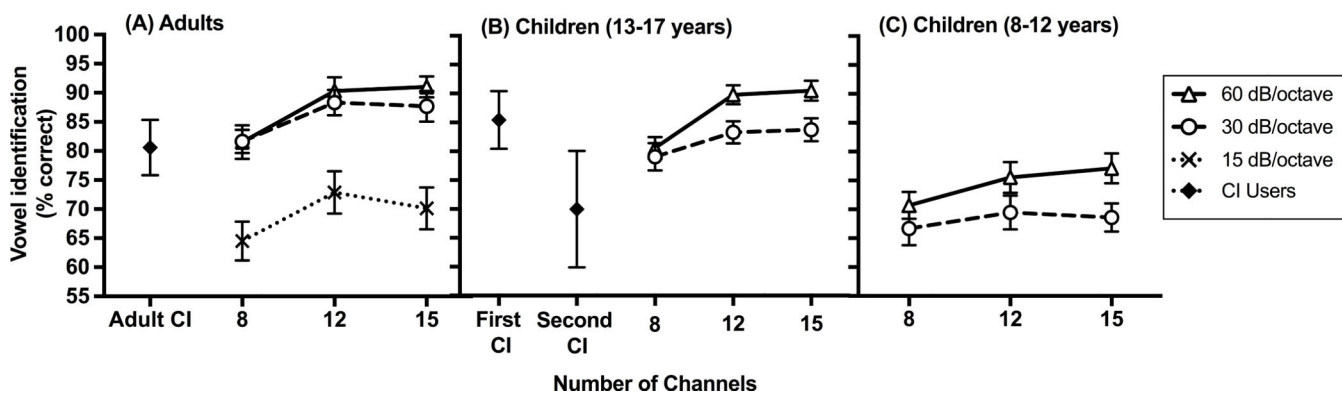


Figure 3. Vowel identification scores (in percent correct) across all filter slope and processing channel conditions for (A) Adults with normal hearing (NH) and adults with cochlear implants (CIs), (B) Children with NH and children with CIs age 13-17 years, and (C) Children with NH age 8-12 years. Error bars represent ± 1 standard error of the mean.

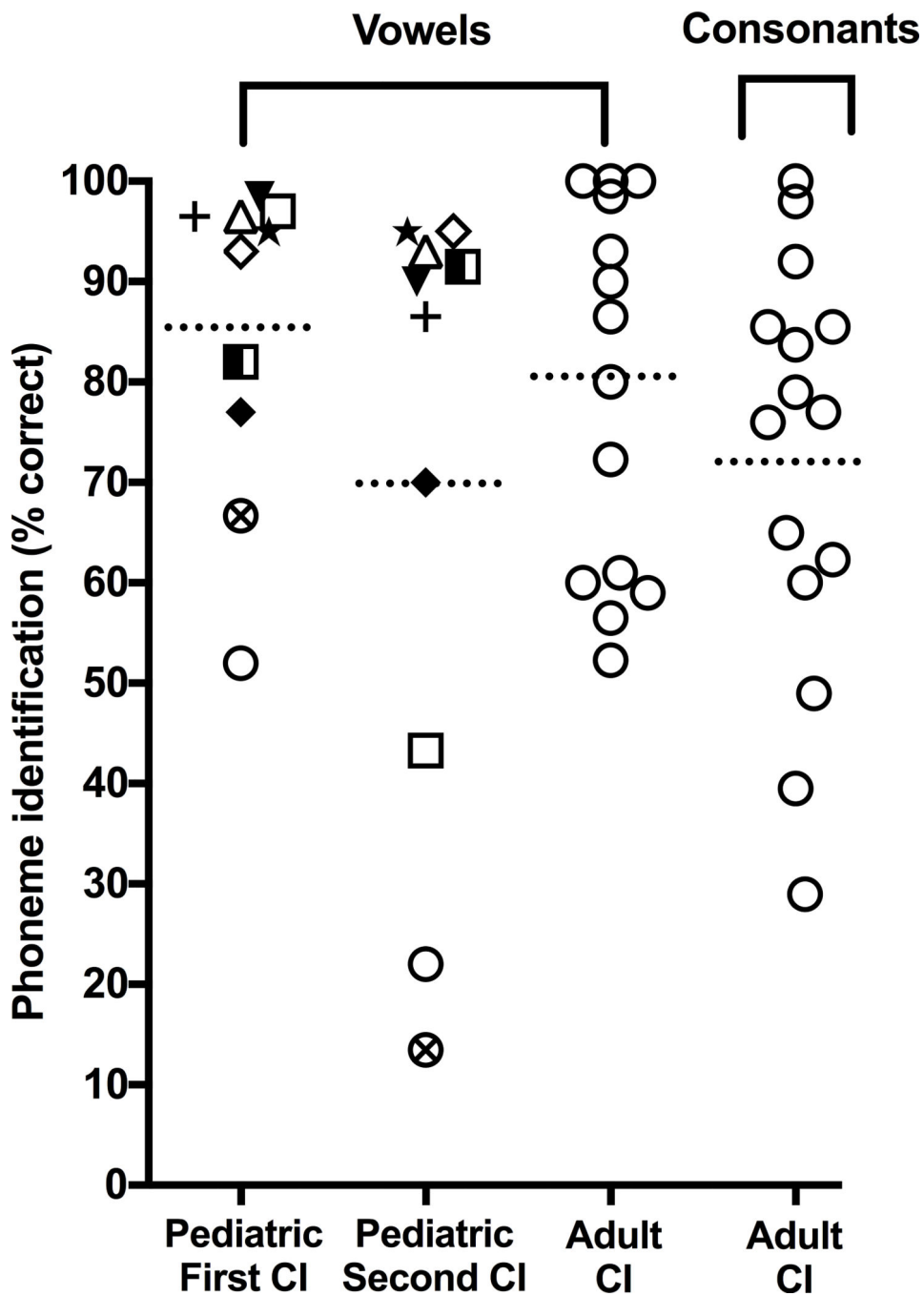


Figure 4. Phoneme identification scores (in percent correct) for individual children and adults with cochlear implants (CIs). Data from each pediatric CI user is presented with a unique symbol for comparison between ears within an individual.

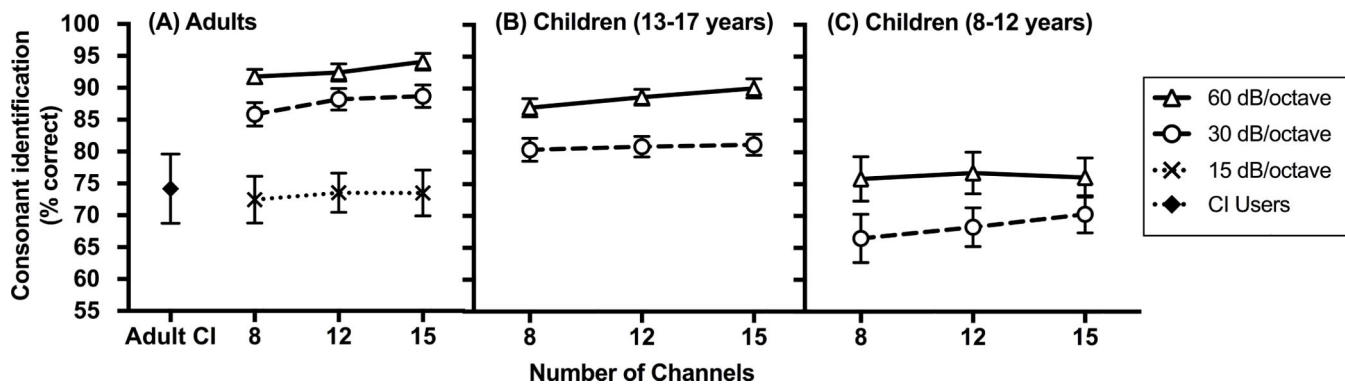


Figure 5. Consonant identification scores (in percent correct) across all filter slope and processing channel conditions (A) Adults with normal hearing (NH) and adults with cochlear implants (CIs), (B) Children with NH and children with CIs age 13-17 years, and (C) Children with NH age 8-12 years. Error bars represent ± 1 standard error of the mean.

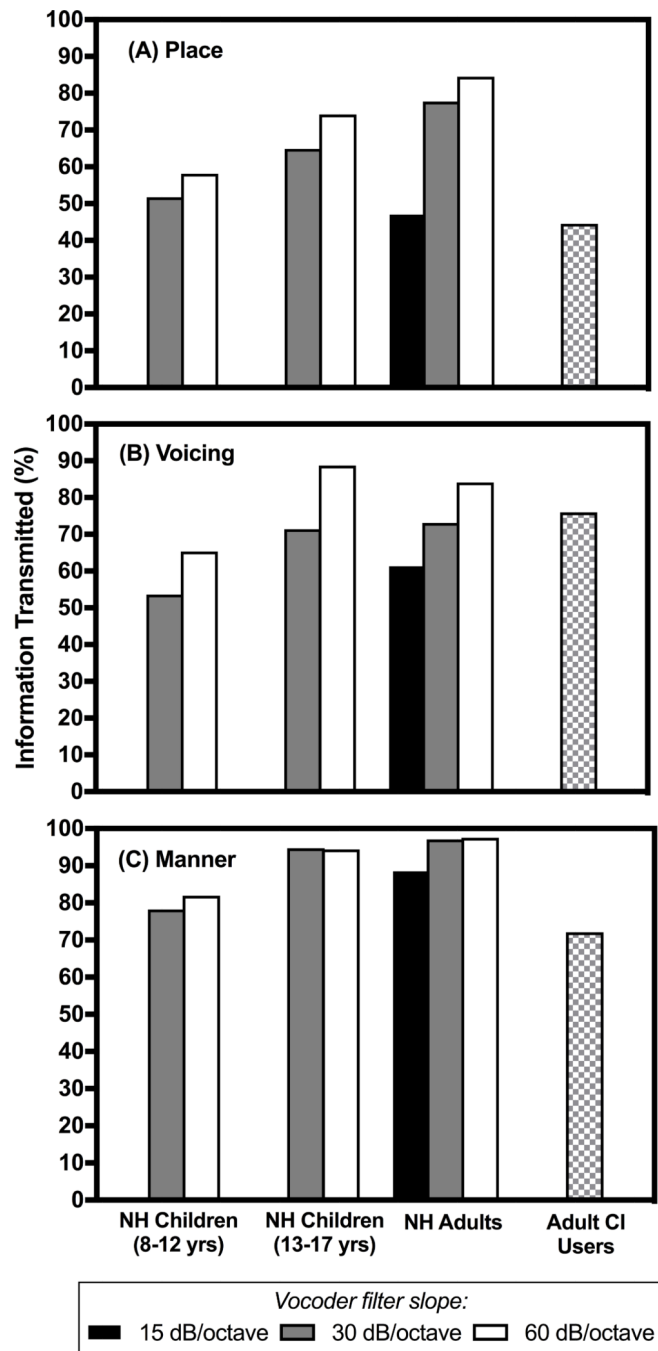


Figure 6. Results of the sequential information analysis (SINFA), showing consonant acoustic-phonetic feature transmission (in percent) for all filter slope conditions, collapsed across processing channels for (A) place of articulation, (B) voicing, and (C) manner of articulation. Data for each normal hearing age group and the adults with cochlear implants are shown.

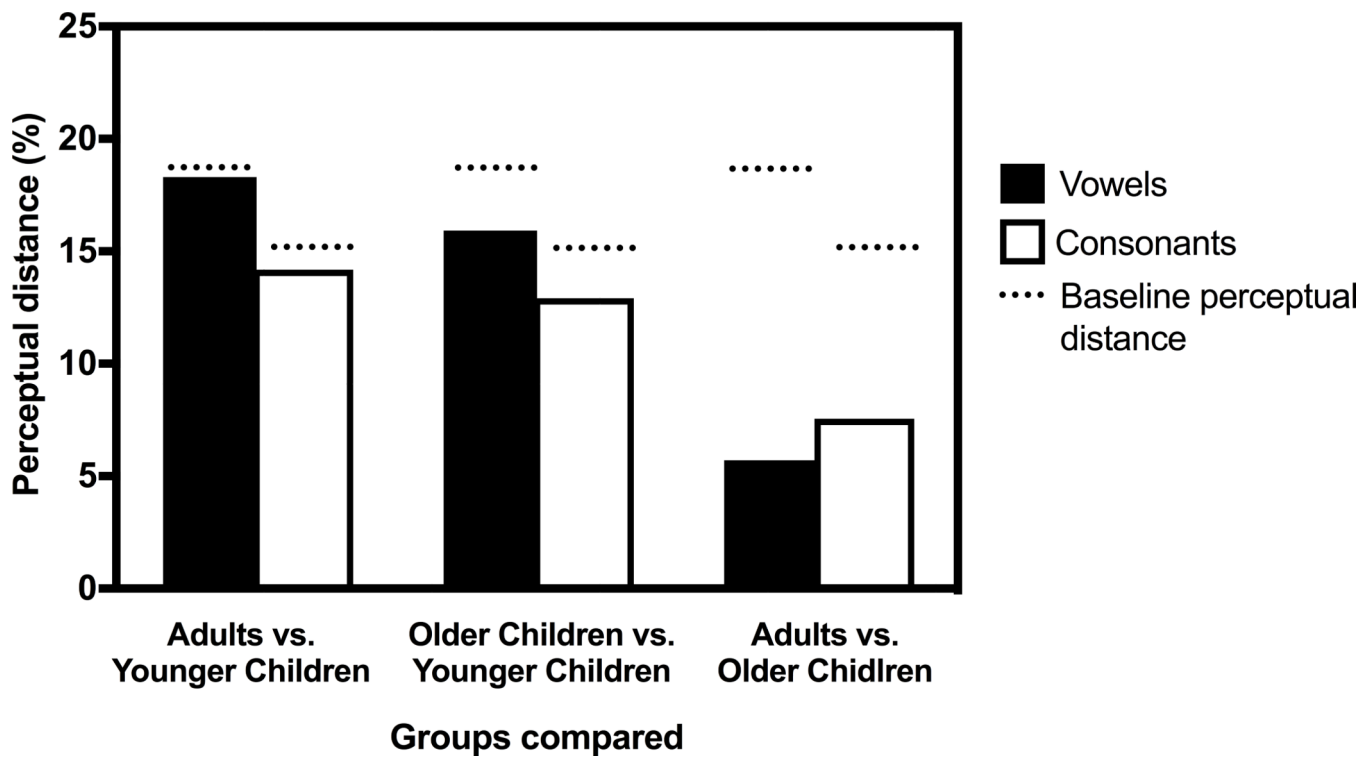


Figure 7. Perceptual distance (in percent) between the vowel and consonant confusion matrices in the 15-channel, 30 dB/octave filter slope condition. Comparisons are made between each normal hearing age group. Dashed lines indicate the baseline (within-subject) perceptual distances.

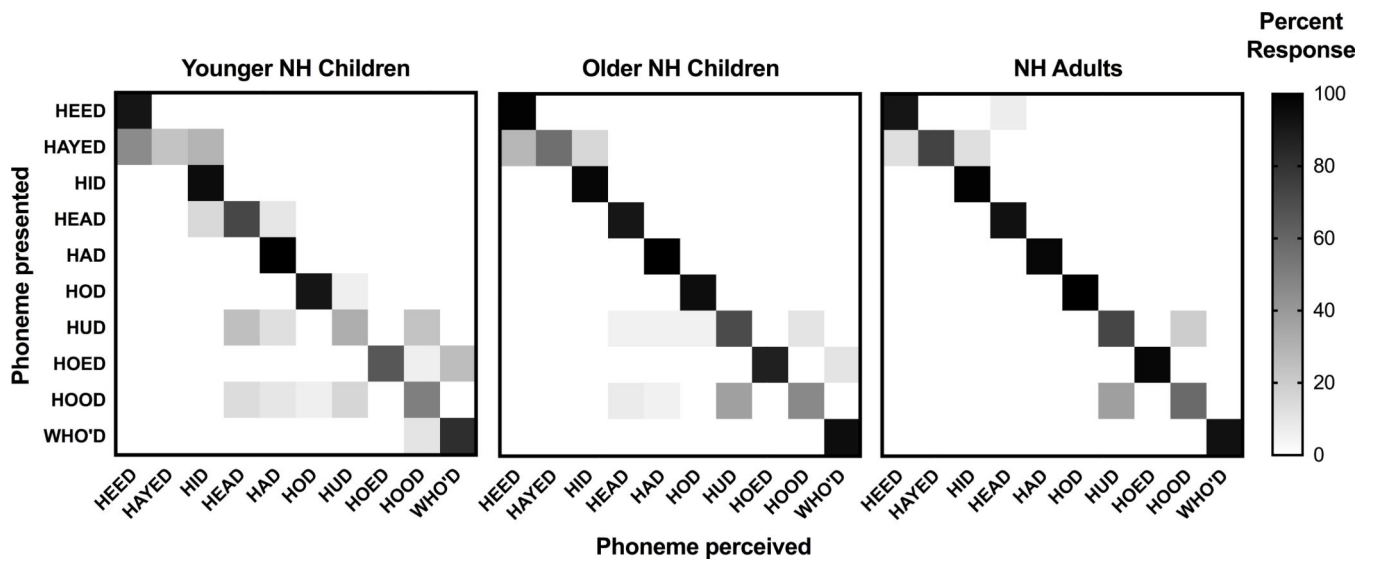


Figure 8. Vowel confusion matrix for the 15-channel, 30 dB/octave filter slope condition for younger children with normal hearing (NH), older children with NH, and adults with NH. Percent response for each cell is indicated via a color gradient, with black representing 100% responses and white representing 0% responses. Adjacent vowels have similar formant frequencies and vowels on opposite ends of the matrix have similar first formant frequencies.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

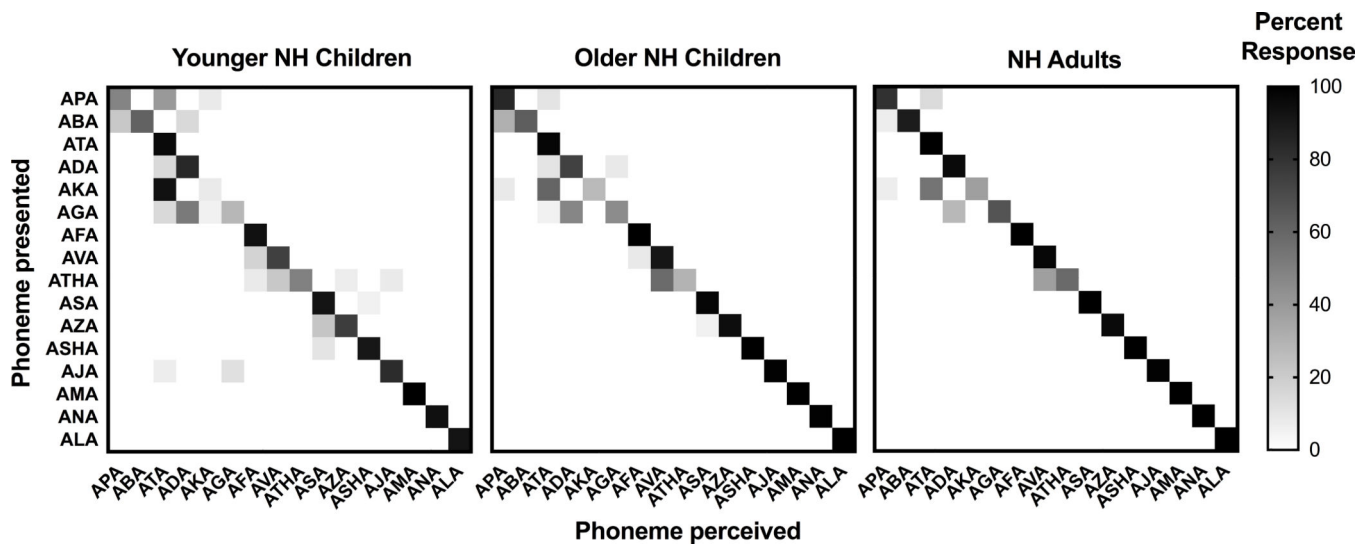


Figure 9. Consonant confusion matrix for the 15-channel, 30 dB/octave filter slope condition for younger children with normal hearing (NH), older children with NH, and adults with NH. Percent response for each cell is indicated via a color gradient, with black representing 100% responses and white representing 0% responses. Consonants are ordered by manner of articulation.

Table 1

Demographic information for all pediatric CI subjects including: ears implanted, chronological age of implantation for each ear (in years), chronological age at time of testing (in years), etiology (if known), and vowel identification score for each implanted ear. DFNB1: genetic non-syndromic hearing loss.

<i>Pediatric CI Demographic Information</i>										
ID	First CI	Age Implanted	Second CI	Age Implanted	Age	Etiology	Vowel Score (%)			
							First CI	Second CI		
P01	R	2.3	L	12.1	15.7	Unknown	97	43		
P03	R	1.4	L	5.6	12.9	Unknown	95	95		
P04	R	1.5	L	4.5	13.2	Unknown	77	70		
P05	R	4.1	L	13.9	17.7	DFNB1	97	87		
P06	R	4.3	L	11.0	17.2	Unknown	99	90		
P07	R	1.9	L	4.9	13.3	Unknown	97	93		
P09	L	2.6	R	3.9	13.5	Unknown	82	92		
P10	L	1.1	R	5.1	13.3	DFNB1	93	95		
P11	R	1.4	L	10.2	13.3	DFNB1	52	22		
P12	R	1.7	L	10.2	13.3	DFNB1	67	14		
<i>Mean</i>		<i>2.23</i>		<i>8.14</i>	<i>14.34</i>		<i>85.60</i>	<i>70.10</i>		
<i>SD</i>		<i>1.13</i>		<i>3.70</i>	<i>1.82</i>		<i>15.90</i>	<i>31.83</i>		

Demographic information for adult CI subjects including: ear implanted, chronological age at time of testing, age diagnosed with a profound hearing loss (if known), age at implantation, etiology (if known), vowel identification score, and consonant identification score. EVA: enlarged vestibular aqueduct.

Table 2

Adult CI Demographic Information

ID	Ear	Age	Age at Profound HL	Age Implanted	Etiology	Vowel Score (%)	Consonant Score (%)
S22	R	74	55	66	Suspected Genetic	61	65
S23*	L	70	59	62	Unknown	100	98
S23*	R	70	59	65	Unknown	80	86
S28	R	76	62	70	Autoimmune	72	60
S29	L	84	47	77	Noise exposure	93	84
S30	L	51	23	40	Genetic	100	92
S38	L	50	45	46	Otosclerosis	59	49
S40	L	52	4	50	EVA	57	29
S41	L	49	42	43	Maternal rubella	100	76
S42	R	64	50	50	Idiopathic	99	100
S43	R	69	50	67	Noise exposure	60	77
S44	R	52	Unknown	51	Antibiotics	90	62
S46	R	66	14	62	Unknown	52	40
S47	R	37	28	37	Unknown	100	86
S48	R	59	36	58	Autoimmune	87	79
<i>Mean</i>		<i>61.53</i>	<i>41.00</i>	<i>56.27</i>		<i>80.67</i>	<i>72.20</i>
<i>SD</i>		<i>12.78</i>	<i>17.79</i>	<i>11.94</i>		<i>18.62</i>	<i>21.01</i>

* denotes bilateral CI user.

Table 3

Per-channel frequency allocations (in Hz) for the 8-, 12-, and 15-channel noise-band conditions.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
15 channels	250-420	421-504	505-606	607-729	730-876	877-1052	1053-1265	1266-1520	1521-1826	1827-2195	2196-2637	2638-3169	3170-3808	3809-4676	4577-8700
12 channels	250-336	337-452	452-607	608-816	817-1097	1098-1475	1476-1982	1983-2665	2666-3582	3583-4815	4816-6472	6473-8700			
8 channels	250-390	391-607	608-946	947-1475	1476-2298	2299-3582	3583-5582	5583-8700							

Table 4

Classification of consonant acoustic-phonetic features. Voicing coding: 0 = voiceless, 1 = voiced. Manner coding: 1 = plosive, 2 = fricative, 3 = affricate, 4 = nasal. Place coding: 1 = labial, 2 = dental, 3 = alveolar, 4 = palatal, 5 = back.

	b	d	f	g	ɟʒ	k	l	m	n	p	s	ʃ	t	θ	v	z
Voicing	1	1	0	1	1	0	1	1	1	0	0	0	0	0	1	1
Manner	1	1	2	1	3	1	5	4	4	1	2	2	1	2	2	2
Place	1	3	1	5	4	5	3	1	3	1	3	4	3	2	1	3

Table 5

Descriptive statistics (means and standard deviations) for the NH participants' phoneme identification scores in each vocoder condition. Data are presented in terms of percent correct.

Phoneme Identification (in percent correct) M (SD)												
	15 dB/octave slope			30 dB/octave slope			60 dB/octave slope					
	8	12	15	8	12	15	8	12	15			
Vowels												
Adults	64.49 (11.57)	72.86 (12.61)	70.14 (12.55)	81.65 (6.88)	88.33 (7.61)	87.68 (9.13)	81.54 (10.10)	90.29 (8.22)	91.00 (6.23)			
Children (13-17 yrs)				79.04 (8.61)	83.31 (7.00)	83.73 (7.18)	80.58 (6.78)	89.77 (5.96)	90.50 (6.21)			
Children (8-12 yrs)		--		66.67 (10.89)	69.46 (10.96)	68.61 (9.23)	70.70 (8.79)	75.54 (10.04)	77.11 (9.65)			
Consonants												
Adults	72.47 (12.79)	73.53 (10.72)	73.49 (12.49)	85.83 (6.40)	88.22 (5.85)	88.69 (6.18)	91.78 (3.89)	92.39 (4.77)	94.11 (4.60)			
Children (13-17 yrs)				80.42 (6.47)	80.88 (5.85)	81.21 (6.06)	87.00 (5.16)	88.67 (4.58)	90.04 (5.37)			
Children (8-12 yrs)		--		68.24 (11.80)	69.44 (9.34)	70.15 (10.74)	77.59 (10.66)	77.41 (11.08)	77.22 (10.25)			

Table 6

Statistical data for the repeated measures ANOVA main effects for NH participants' phoneme identification. Statistically significant *p*-values are indicated by bold font.

	Filter Slope (dB/octave)				Number of Channels				Slope*Channel Interaction			
	<i>F</i>	(<i>df</i>)	<i>p</i>	partial η^2	<i>F</i>	(<i>df</i>)	<i>p</i>	partial η^2	<i>F</i>	(<i>df</i>)	<i>p</i>	partial η^2
Vowels												
Adults	158.35	(2, 22)	<0.001	0.94	1160.48	(2, 22)	<0.001	0.76	1.21	(4, 44)	0.320	--
Children (13-17 yrs)	13.80	(1, 12)	0.003	0.54	47.02	(2, 24)	<0.001	0.80	5.05	(2, 24)	0.015	0.30
Children (8-12 yrs)	20.58	(1, 13)	0.001	0.61	6.21	(1.52, 19.79)	0.012	0.32	3.23	(2, 26)	0.056	--
Consonants												
Adults	114.73	(1.23, 13.50)	<0.001	0.91	7.88	(2, 22)	0.003	0.42	2.04	(4, 44)	0.105	--
Children (13-17 yrs)	84.38	(1, 12)	<0.001	0.88	3.69	(2, 24)	0.040	0.24	3.48	(2, 24)	0.047	0.23
Children (8-12 yrs)	102.53	(1, 12)	<0.001	0.90	0.30	(1.41, 16.96)	0.671	--	0.87	(2, 24)	0.433	--

Mean difference in phoneme identification scores (in percent correct) between filter slope conditions for 8, 12 and 15 channels. Statistically significant differences are indicated by bold font.

Table 7

	Difference (in percent correct)								
	30 dB/octave slope – 15 dB/octave slope			60 dB/octave slope – 30 dB/octave slope					
	Channels			Channels			Channels		
	8	12	15	8	15	8	12	15	
Vowels									
Adults	17.17	15.47	17.54	-0.11	1.96	3.32			
Children (13-17 yrs)				1.54	6.46	6.77			
Children (8-12 yrs)		--		4.04	6.07	8.50			
Consonants									
Adults	13.36	14.69	15.21	5.95	4.17	5.42			
Children (13-17 yrs)				6.58	7.79	8.83			
Children (8-12 yrs)		--		9.31	8.52	5.79			

Table 8

Mean difference in phoneme identification scores (in percent correct) between processing channel conditions for 15, 30 and 60 dB/octave filter slopes. Statistically significant differences are indicated by bold font.

		Difference (in percent correct)								
		12 channels – 8 channels			15 channels – 8 channels			15 channels – 12 channels		
		Filter slope (dB/octave)			Filter slope (dB/octave)			Filter slope (dB/octave)		
		15	30	60	15	30	60	15	30	60
Vowels										
Adults		8.38	6.68	8.75	5.65	6.03	9.46	-2.72	-0.65	0.71
Children (13-17 yrs)		--	4.27	9.19	--	4.69	9.92	--	0.42	0.73
Children (8-12 yrs)		--	2.80	4.83	--	1.94	6.40	--	-0.86	1.57
Consonants										
Adults		1.06	2.39	0.61	1.01	2.86	2.33	-0.04	0.47	1.72
Children (13-17 yrs)		--	0.46	1.67	--	0.79	3.04	--	0.33	1.37
Children (8-12 yrs)		--	1.61	0.87	--	3.49	0.23	--	1.88	-0.65

Sequential information analysis (SINFA) results. Percent of information transmitted for the consonant acoustic-phonetic features of manner of articulation, place of articulation, and voicing as a function of filter slope are displayed. Data from the normal-hearing participants and the adults with cochlear implants are presented.

Table 9

	Information Transmitted (%)								
	15 dB/octave slope			30 dB/octave slope			60 dB/octave slope		
	Place	Voicing	Manner	Place	Voicing	Manner	Place	Voicing	Manner
Adults	47.0	61.3	88.5	77.7	73.1	97.1	84.5	84.1	97.5
Children (13-17 yrs)	--	--	--	64.9	71.4	94.7	74.2	88.7	94.4
Children (8-12 yrs)	--	--	--	51.7	53.6	78.2	58.1	65.3	81.9

Adults with cochlear implants: 44.5% (place), 76.0% (voicing), 72.1% (manner)