JSLHR

Research Article

Effects of Age and Cochlear Implantation on Spectrally Cued Speech Categorization

Mishaela DiNino,[a] Julie G. Arenberg,[b] Ann L. R. Duchen,[c] and Matthew B. Winn[d]

Purpose: Weighting of acoustic cues for perceiving place-of-articulation speech contrasts was measured to determine the separate and interactive effects of age and use of cochlear implants (CIs). It has been found that adults with normal hearing (NH) show reliance on fine-grained spectral information (e.g., formants), whereas adults with CIs show reliance on broad spectral shape (e.g., spectral tilt). In question was whether children with NH and CIs would demonstrate the same patterns as adults, or show differences based on ongoing maturation of hearing and phonetic skills.
Method: Children and adults with NH and with CIs categorized a /b/–/d/ speech contrast based on two orthogonal spectral cues. Among CI users, phonetic cue weights were compared to vowel identification scores and Spectral-Temporally Modulated Ripple Test thresholds.
Results: NH children and adults both relied relatively more on the fine-grained formant cue and less on the broad spectral tilt cue compared to participants with CIs. However, early-implanted children with CIs better utilized the formant cue compared to adult CI users. Formant cue weights correlated with CI participants' vowel recognition and in children, also related to Spectral-Temporally Modulated Ripple Test thresholds. Adults and child CI users with very poor phonetic perception showed additive use of the two cues, whereas those with better and/or more mature cue usage showed a prioritized trading relationship, akin to NH listeners.
Conclusions: Age group and hearing modality can influence phonetic cue-weighting patterns. Results suggest that simple nonlexical categorization tests correlate with more general speech recognition skills of children and adults with CIs.

Accurate perception of the many spectral cues in speech is important for categorization of those sounds. Studies of normal-hearing (NH) listeners have shown that the ability to accurately identify speech deteriorates with systematic degradation of spectral resolution (ter Keurs et al., 1993; Xu et al., 2005). Cochlear implants (CIs) are highly successful auditory prostheses but provide limited spectral information due to the spread of electrical activation within the cochlea (Boëx et al., 2003) and the small number of frequency channels in an implant (Friesen et al., 2001). All CI users thus have limited spectral resolving capabilities, but spectral resolution can be further affected by a number of factors, such as hearing history and the interface between implant electrodes and auditory neurons (Bierer, 2007, 2010). Variation in the ability to resolve frequency components of auditory signals likely contributes to the wide range of speech identification scores observed in adults (e.g., Holden et al., 2013) and children (e.g., Wang et al., 2008) with CIs.

Previous investigations have found disparities in the CI electrode–neuron interface between prelingually deafened, early implanted children and postlingually deafened, late-implanted adults that could differentially affect the spectral resolution of these groups. Evidence from these studies suggests that, relative to adults, children exhibit greater spiral ganglion neuron integrity (Jahn & Arenberg, 2020) and higher levels of tissue growth in the cochlea (Busby et al., 2002; Molisz et al., 2015), which both contribute to lower levels of electrical current required to achieve both auditory perception and a comfortable listening level (DiNino et al., 2019). Yet, prior studies that have compared spectral discrimination of children and adults with CIs have produced mixed results. These investigations also

[a]Department of Psychology, Carnegie Mellon University, Pittsburgh, PA
[b]Massachusetts Eye and Ear, Harvard Medical School Department of Otolaryngology, Boston
[c]Seattle Children's Hospital, WA
[d]Department of Speech-Language-Hearing Sciences, University of Minnesota, Minneapolis

Correspondence to Mishaela DiNino: mdinino@andrew.cmu.edu

Editor-in-Chief: Frederick (Erick) Gallun
Editor: Steve Aiken

Received June 6, 2019
Revision received August 12, 2019
Accepted March 30, 2020
https://doi.org/10.1044/2020_JSLHR-19-00127

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

yielded conflicting findings regarding the relationship between early-implanted children's spectral discrimination abilities and their speech recognition performance (Gifford et al., 2018; Horn et al., 2017; Landsberger et al., 2018).

Furthermore, although previous research in adult CI listeners has shown strong correlations between speech recognition scores and performance on several assessments of spectral resolution, such as spectral ripple discrimination (Henry et al., 2005; Kenway et al., 2015; Supin et al., 1994; Won et al., 2007) and electrode pitch ranking (Nelson et al., 1995), the exact mechanisms underlying these relationships are unclear; the stimuli in these tests do not have spectral properties that are comparable to those found in speech (Saoji et al., 2009; Singh & Theunissen, 2003). Investigation of *functional* spectral resolution as it relates to speech identification is thus necessary to better understand mechanisms of potential differences in spectral resolution between early-implanted children and late-implanted adults with CIs.

Assessments of spectral resolution also diverge from the process of speech recognition in several other important ways. Classic psychophysical tasks measure discrimination of spectral content, whereas speech identification is essentially a process of categorization (Holt & Lotto, 2010; Liberman et al., 1957; Winn et al., 2016). In other words, listeners in everyday life hear utterances and identify them, rather than hearing two utterances and deciding whether they are the same or different. Speech recognition tests themselves can be contaminated by many nonauditory factors, such as lexical knowledge (Ganong, 1980), context effects (Norris et al., 2003; Schertz & Hawthorne, 2018), and working memory capacity (Hadar et al., 2016). These variables likely differ between children and adults with CIs, further complicating comparisons between these two groups. It would be desirable to have an assessment of spectral resolution for child and adult CI users that is comparable to the process of speech perception in terms of auditory processing demands, but that allows for greater control over the processes being tested, while avoiding nonauditory linguistic factors. The current study pursues such a goal.

Spectrally Cued Speech Categorization

To leverage the acoustic structure of speech sounds toward the goal of measuring spectral resolution, Winn and Litovsky (2015) developed a test that depends on weighting of two spectral phonetic cues. A listener is required to categorize /bɑ/ and /dɑ/ stimuli, which are contrasted by place of articulation. Because the difference is conveyed by variation in vocal tract resonant frequencies, perception of this consonant feature is greatly affected by reduced frequency resolution. Accordingly, categorizing /bɑ/ and /dɑ/ is particularly difficult for individuals with hearing loss and with CIs (Munson et al., 2003). Co-occurring acoustic cues exist for any speech contrast, and the value of the test developed by Winn and Litovsky hinges on detecting whether a listener relies on a primary cue carried by fine spectral contrast, or a secondary cue conveyed by coarse spectral information. In this test, naturally spoken /bɑ/ and /dɑ/ stimuli were manipulated to change both formant transitions and spectral tilt, the relative balance of high and low frequency energy in the spectrum. Stimuli with a diffuse-falling spectrum (decreasing spectral tilt and second formant [F2] frequency at the onset of the transition) are perceived acoustically as "ba," whereas stimuli with a diffuse-rising spectrum (increasing spectral tilt and F2 frequency at the onset of the transition) are perceived acoustically as "da" (Stevens & Blumstein, 1978).

In the case of poor use of or poor access to one acoustic–phonetic cue, a listener can typically compensate by utilizing a different cue (Repp, 1982), termed "cue weighting" or "cue trading" (in the case when the use of one cue directly demands less of the other cue). Considering that many acoustic cues are spectral in nature, speech categorization could be useful to determine how an individual with a CI uses acoustic–phonetic information to perform speech recognition tasks. This would contribute to a better understanding of the possible underlying mechanisms of how the spectrum is mapped to a recognizable percept and how it could be utilized to achieve a certain word recognition score. In contrast, a classic test of speech perception gives a person's overall score, but does not investigate the perceptual variables that may influence how that score is achieved.

Resolving formant cues requires a high level of spectral resolution, given the narrow bands of energy in the frequency domain. In contrast, the spectral tilt cue covers a wide range of frequencies and is thus accessible even with poor spectral resolution (Alexander & Kluender, 2009; Winn & Litovsky, 2015). NH adults have been found to utilize the onset frequency of F2 to perceive place of articulation in the /bɑ/–/dɑ/ contrast under normal conditions; yet, these same individuals weighted spectral tilt higher when formant information in the contrast was degraded or missing. Individuals with adequate spectral resolution should be able to utilize a formant cue to categorize /bɑ/ and /dɑ/, whereas those with poor spectral resolution will have limited access to the formant cue and should place greater perceptual weight on spectral tilt.

This hypothesis was validated in the study by Winn and Litovsky (2015), who tested adults with NH and with CIs. Degrading spectral resolution by vocoding the stimuli for NH adults resulted in decreased weighting of the formant cue and increased perceptual weighting for the spectral tilt cue. Cue trading in adult CI listeners was also observed by Winn et al. (2016), who demonstrated that higher relative weighting of the formant cue was a slightly stronger predictor of word recognition scores compared to a nonlinguistic test of spectral ripple discrimination. This supports the notion that measures more reflective of speech recognition might be favorable as explanatory variables. It is important to note that this is a case of acoustic cue categorization predicting speech recognition more broadly, because the categorization task utilizes a cue known to be specifically relevant for speech sounds. Results from these studies and others (e.g., Moberly et al., 2016) indicate that speech categorization tests can be sensitive to spectral resolution as it relates to spectral cues that arise in natural speech.

Phonetic Cue Weighting in Children

This study examined whether children with NH and CIs exhibit the same phonetic cue weighting as previously found in adults with these hearing modalities. NH children exhibit difficulty when categorizing stimuli and seem to require greater salience or acoustic differentiation of cues compared to adults (Hazan & Barrett, 2000; Morrongiello et al., 1984). Prior studies have observed divergent cue weighting patterns for place of articulation between very young children (ages 3–7 years) and adults (Mayo & Turk, 2004, 2005), but even older children (ages 5–11 years) have demonstrated difficulty discriminating consonants based on this spectral feature (e.g., Ohde et al., 1996). Identification of consonants in challenging listening conditions has been found to be immature until age 15 (Johnson, 2000). A number of studies (Nittrouer, 2002, 2005; Nittrouer & Miller, 1997) have suggested that children transition from reliance on harmonic-carried cues to other cues during maturity, but the generality of this finding is unknown, owing to the lack of wide replication in the literature. Thus, the age at which phonetic cue categorization becomes adultlike remains unknown.

In addition to potential age-related effects that could result in distinct patterns of perceptual weights between adults and children, phonetic cue usage may differ between early-implanted children and late-implanted adults with CIs because each group learned language with a different hearing modality. The development of language with acoustic or electric hearing could potentially have impact on what sound structures are learned to associate with different phonetic categories. Late-implanted adults are likely required to change their auditory perceptual strategies after receipt of the implant (e.g., Hedrick & Carney, 1997; Moberly et al., 2014; Winn et al., 2012), but early-implanted children acquire their perceptual strategies with CI input and no prior acoustic scaffolding. As speech perception abilities are crucial for the development of pediatric CI users' language skills, the current study sought to better understand early-implanted children's patterns of phonetic cue perception as compared to those of adults with CIs and of NH children.

Aims of This Study

The primary goal of this study was to examine the influence of age group (child or adult) and hearing modality (acoustic hearing or CI) on use of a fine-resolution formant cue as assessed by the cue-weighting task developed by Winn and Litovsky (2015). We expected to replicate the results of that study and thus predicted that NH adults would use the formant cue more reliably than adults with CIs. However, the cue weights of both groups of children may be affected by developmental effects in auditory perception, and those of children with CIs may be further influenced by language acquisition with the CI. A second goal of this study was to assess the relationship between child and adult CI users' formant cue resolution to their performance on vowel identification, which should also depend on sufficiently good perception of spectral shape to achieve high performance.

A third aim was to compare CI users' perceptual weighting of the formant cue (a measure of one's ability to use fine spectral cues as they arise in speech stimuli) to performance on a standardized spectral ripple task (a nonlinguistic assessment of spectral discrimination) to link with previous studies of spectral resolution in CI users. These tasks have been used extensively to assess spectral discrimination (e.g., DiNino & Arenberg, 2018; Henry & Turner, 2003; Horn et al., 2017; Landsberger et al., 2018) but may have limited explanatory power for CI listeners (Winn & O'Brien, 2019), especially for those who achieve excellent thresholds on this task, and thus the test of speech categorization may be more suitable for assessing functional spectral resolution in children and adults who use CIs.

Bilaterally implanted participants first performed these tasks with each CI separately. This allowed us to confirm that any differences seen in auditory spectral resolution between the two CIs of the same participant did not result from nonauditory variables such as age, cognitive factors, and linguistic experience. Many of these participants then performed the speech categorization task with both ears simultaneously so that bilateral performance could be compared to that of their individual CIs.

Method

Participants

Two groups of CI users participated in this study: 12 prelingually deafened children (mean age at first visit = 13.3, *SD* = 1.7 years) and 15 postlingually deafened adults (mean age at first visit = 67.9, *SD* = 11.1 years). All children received their first implant prior to 5 years of age (mean age at first CI = 2.0, *SD* = 1.1 years). All adults received a CI later in life (mean age at first CI = 59.0, *SD* = 13.3 years). Twelve of the 13 children were bilaterally implanted; each ear of those children was tested separately, for a total of 25 ears in the child CI participant group. Seven of the adults were bilaterally implanted, and therefore 22 adult ears were tested for the adult CI group. All participants with CIs used oral communication and were native speakers of American English. Most used Advanced Bionics HiRes90K devices; five of the adults (SC01, SC03, SC05, SC06, and SC07) and three of the children with CIs (PC02, PC03, PC04) utilized cochlear devices. CI participant demographics are shown in Table 1.

Each CI was initially tested separately to investigate auditory task performance while minimizing the influence of nonauditory factors within a participant. The order of ear tested (first- or second-implanted) and of tasks performed with each CI was randomized. However, Winn and Litovsky (2015) tested bilaterally implanted participants with both CIs simultaneously, and so for the purposes of

Table 1. Demographics of participants with cochlear implants.

Children						Adults					
Subject	Ear	Etiology of deafness:	Age at first testing:	Age at implantation (yrs)	Duration of deafness (yrs)	Subject	Ear	Etiology of deafness	Age at first testing	Age at implantation (yrs)	Duration of deafness (yrs)
P01	R	Unknown	15.7	2.3	1.5	S22	R	Hereditary	76.9	66.7	11.8
	L			12.1	11.3	S23/S36	L	Unknown	72.1	62.0	3.9
P02	R	EVA	11.8	1.1	1.0		R			64.5	6.5
	L			3.1	3.0	S29	L	Unknown	87.8	76.8	29.8
P03	R	Unknown	12.9	1.4	1.1		R			85.7	38.7
	L			5.6	5.3	S39/S30	R	Hereditary	53.4	30.1	9.1
P04	R	Unknown	13.2	1.5	0.8		L			40.1	19.1
	L			4.5	3.8	S43	R	Noise Exposure	72.5	67.9	17.9
P06	R	Unknown	17.2	4.3	2.5	S45	R	Hereditary	65.3	54.0	32.0
	L			11.0	9.1	S46	R	Unknown	69.4	64.2	48.2
P07	R	Unknown	13.3	1.9	0.4	S47/S51	R	Unknown	39.1	36.4	10.3
	L			4.9	3.5		L			38.4	12.4
P09	L	Unknown	13.5	2.6	1.3	S50	R	Measles	76.5	61.1	41.1
	R			3.9	2.7	S52	R	Unknown	70.9	66.0	21.1
P10	L	DFNB1	13.3	1.1	0.9	SC01	R	Unknown	64.1	62.3	37.3
	R			5.1	4.9	SC03	L	Unknown	70.5	64.7	12.2
P11	R	DFNB1	13.3	1.4	1.2		R			66.0	13.5
	L			10.2	10.0	SC05	L	Hereditary	62.0	58.1	1.2
P12	R	DFNB1	13.3	1.7	1.4		R			59.9	3.0
	L			10.2	10.0	SC06	R	Unknown	71.0	66.2	49.4
PC02	R	Connexin	13.2	1.7	0.5	SC07	R	Hereditary	66.7	46.5	30.4
	L			1.7	0.5		L			59.5	43.5
PC03	L	Connexin	10.4	1.1	0.5						
	R			1.1	0.5						
PC04	R	Connexin	11.7	3.9	1.6						

Note. yrs = years; R = right; L = left; EVA = enlarged vestibular aqueduct; DFNB1 and connexin = gene mutations resulting in nonsyndromic hearing loss.

comparing study results and to relate cue weights from individual CIs to those in a bilateral condition, we tested a subset of bilaterally implanted participants (eight children and six adults) on the speech categorization test while they listened with both CIs simultaneously. We were unable to acquire speech categorization data in this condition from all participants due to testing session time constraints.

Seventeen NH adults between the ages of 19 and 74 (M_{age} = 41.8 years, SD = 18.4 years) and 17 NH children between the ages of 8 and 16 (M_{age} = 12.1 years, SD = 2.6 years) were also included in this study. The NH child group age range was chosen to be similar to that of child participants with CIs. The NH adult group consisted of 10 "young adults" between the ages of 19 and 32 (M_{age} = 27.4 years, SD = 3.5 years) and seven "older adults," aged 53–74 years (M_{age} = 62.3 years, SD = 6.4 years) to control for any potential effects of aging on cue-weighting patterns. The older NH adults were closer in age to the adult participants with CIs than were the younger NH adults. All NH participants completed a screening to verify clinically NH thresholds ≤ 20 dB HL at 250, 500, 750, 1000, 2000, and 4000 Hz.

All study procedures were approved by the University of Washington Human Subjects Division. Adults gave written informed consent prior to participation. Children gave written informed assent and a parent or legal guardian provided written consent to their participation in this research.

Assessments

All participants performed the spectral speech cue categorization test, described in detail below. Participants with CIs were additionally tested on vowel identification and spectral ripple discrimination. Stimuli for all assessments were presented through a Crown D75 amplifier and an external A/D device (SIIF USB SoundWave 7.1) and were played through speakers in a double-walled sound-attenuating booth (IAC RE-243). Participants were seated in front of a computer screen with the speaker directly in front. NH participants performed the task with both ears simultaneously. All participants with CIs were first tested monaurally on each task and turned off their CI or hearing aid in the contralateral ear during testing. The ear not being tested was additionally plugged for unilaterally implanted participants to minimize the effects of any residual hearing. Formant cue weights in the monaural condition were statistically compared to a listener's vowel identification and spectral ripple discrimination performance with that same CI. Most bilaterally implanted child and adult CI users additionally performed the categorization test with both CIs concurrently so that cue weights from individual CIs could be compared to those from a condition that was more representative of everyday listening for bilaterally implanted participants.

Test of Spectral Cue Categorization

Stimuli Creation

Full details of stimulus generation can be found in the paper by Winn and Litovsky (2015). Six naturally spoken sounds from a male, native American English speaker were utilized as stimuli. These included the contrast of interest, /bɑ/ and /dɑ/, which emerged as a 40-token matrix contrasting by eight steps of formant transitions and five steps of spectral tilt, as well as two additional two-way contrasts: /sɑ/ and /ʃɑ/, and /rɑ/ and /lɑ/. Each of the contrasts involved a separate manner of articulation, leading to the high likelihood that perceptions would be constrained to the intended contrast (i.e., “sha” is very unlikely to be misperceived as “da”). All stimuli manipulations were performed in Praat (Boersma & Weenink, 2019).

The /bɑ/–/dɑ/ contrast was created by orthogonal manipulation of the tokens on a continuum of (a) second and third formant transitions and (b) spectral tilt at the onset of the syllable. The formant continuum was created by first downsampling the /bɑ/ token to 10000 Hz and estimating 12 linear predictive coding coefficients below 5000 Hz. The sound was then inverse filtered by the linear predictive coding to remove formant peaks, creating a “source” stimulus that allowed for filtering by a different formant structure. As shown in Figure 1, the formant contours of the original /bɑ/ and /dɑ/ tokens were extracted, and six intermediate formant contours were interpolated using the Bark frequency scale, for a total of eight formant steps from /bɑ/ to /dɑ/. Each step of the stimuli was low-pass filtered at 3500 Hz and added to the original /bɑ/ sound that had been high-pass filtered above 3500 Hz, to restore naturalness of the stimuli.

After creation of the /bɑ/-/dɑ/ formant continuum, the spectral tilt (from the first to fourth formant at the syllable onset) of the stimuli was altered on a continuum of five steps within each formant step (see Figure 2). A filter that utilized logarithmic multiplication of the amplitude spectrum either amplified or attenuated frequency energy above 800 Hz (as a function of frequency) to create varying slopes of spectral tilt. The formant continuum stimuli described above were then each multiplied by the five filters of shifted spectral tilt to create a two-dimensional continuum. The filtered stimuli were cross-faded into the vowel nucleus using an 80-ms window, so that the spectral tilt change was a dynamic rather than a static cue (as suggested by the work of Alexander & Kluender, 2008, 2009). A uniform consonant burst (blended between /bɑ/ and /dɑ/) was pre-appended to each syllable to ensure a clear and natural perception of a stop sound.

The formant contrast (F2) between Step 1 and Step 8 was a range of about 1000–1800 Hz, whereas spectral tilt was altered within a frequency range between about 800 and 6000 Hz. Therefore, the formant cue demanded finer spectral resolution. Individuals with poorer spectral resolution would not be able to resolve the fine-resolution formant cue and would be relegated to using the spectral tilt cue.

Testing Procedure

Participants performed a one-interval, six-alternative forced-choice task of speech sound categorization. Subjects were seated in front of a computer screen that contained six boxes, each labeled with “ba,” “da,” “sha,” “sa,” “la,” or “ra.” The additional four phonemes /sɑ/, /ʃɑ/, /rɑ/, and

Figure 1. Formant transition manipulation for /bɑ/–/dɑ/ stimuli, as in Winn and Litovsky (2015). Illustrated are spectral slices of the first 80 ms of the vowel onset. Lighter colored lines are the most like /d/ and black lines are most like /b/.

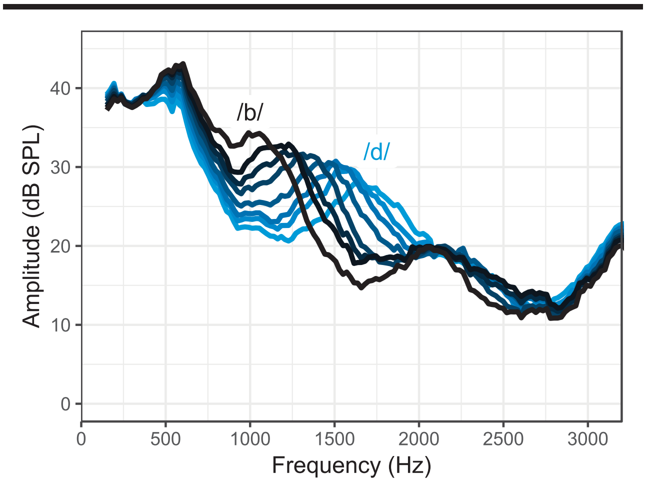
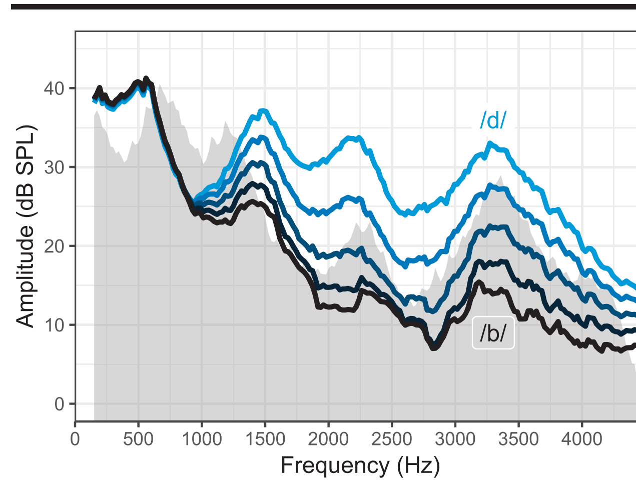


Figure 2. Spectral tilt manipulation for /bɑ/–/dɑ/ stimuli as in Winn and Litovsky (2015). Illustrated are spectral slices of the first 80 ms of the vowel onset. These spectra correspond to five spectral tilts within a single step of the formant continuum. Lighter colored lines are the most like /d/ and black lines are most like /b/. The gray shaded area indicates the spectrum of the vowel following the transition from consonant to /ɑ/.



/lɑ/ were used as fillers for the test to reduce monotony and so that participants would not become unnaturally sensitive to the manipulation of the /bɑ/–/dɑ/ contrast over the testing time. We chose these "filler" stimuli in particular because, as noted by Winn and Litovsky (2015), they are easier for CI listeners to categorize than /bɑ/–/dɑ/ and could therefore help maintain participant confidence and motivation during the task. Stimuli were presented at 65 dBA. After presentation of a sound, participants were asked to select the sound they heard by using a computer mouse to click the appropriate box. Each participant completed one practice run in which only the end point stimuli of all contrasts were presented, to familiarize the participant with the testing procedure. Participants then performed five test runs with one repetition of each unique /bɑ/-/dɑ/ stimulus in the formant and spectral tilt matrix (as well as all of the filler stimuli) per run. Data from the practice run were not included in the analysis. Participants were allowed to repeat the stimuli during all runs but were encouraged to guess from their first impression of the sound instead of repeating it. No feedback was provided during practice or test runs.

Vowel Identification

Children and adults with CIs performed a closed-set test of vowel identification. Ten vowels in /hVd/ context (/i/, /ɪ/, /eɪ/, /ɛ/, /æ/, /ɑ/, /u/, /ʊ/, /o/, /ʌ/) naturally spoken by a female talker were presented at 60 dBA. Stimuli were presented, and participant's responses were recorded through ListPlayer software (Version 2.2.11.52, Advanced Bionics, LLC). Participants were seated in front of a computer screen that contained 10 boxes labeled with each of the responses ("heed," "hid," etc.). Following presentation of a vowel sound, participants were asked to select the box that was labeled with the sound they perceived.

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

The Spectral-Temporally Modulated Ripple Test

Participants with CIs performed the Spectral-Temporally Modulated Ripple Test (SMRT), a test of broadband spectral ripple discrimination, so that performance on speech-based spectral cue categorization could be compared to thresholds from this more commonly used, nonlinguistic test of spectral discrimination. Stimuli were composed of 202 summed pure tones with a sinusoidal spectral shape with a drifting phase, as described by Aronoff and Landsberger (2013). The variable element of each stimulus was the density of spectral peaks, expressed as ripples per octave. Discrimination of a larger number of ripples per octave is typically interpreted as a sign of better spectral resolution. The CI processor does not faithfully transmit the spectral density of ripple stimuli even at moderately low ripple per octave (RPO) values (DiNino & Arenberg, 2018; Lawler et al., 2017; Winn & O'Brien, 2019), but such tests are widely used in studies of CI listeners, and thus was utilized in this study for comparison to related literature.

Stimuli were presented at 65 dBA in a three-alternative forced-choice one-down one-up adaptive procedure with 10 reversals. The interface for this test consisted of three boxes on a computer screen labeled with either "1," "2," or "3." Each box was highlighted in red during the presentation of the corresponding first, second, or third sound in a trial. Participants were asked to select the box of the sound that "sounded different" from the others.

Thresholds for each run were calculated based on the average of the last six reversals, with higher thresholds indicating better spectral discrimination abilities. Each participant completed one practice run and two test runs. Although the practice run was identical to the test runs, these data were not included in the calculation of participants' average SMRT thresholds. Repetition of the sounds was not allowed, and no feedback was provided during practice or test runs. If a participant's thresholds from two test runs differed by more than one RPO, he/she completed a third run. Results from all test runs were averaged to determine the mean SMRT threshold for each subject. SMRT data were not collected from one adult (SC03) due to time constraints.

Statistical Analyses

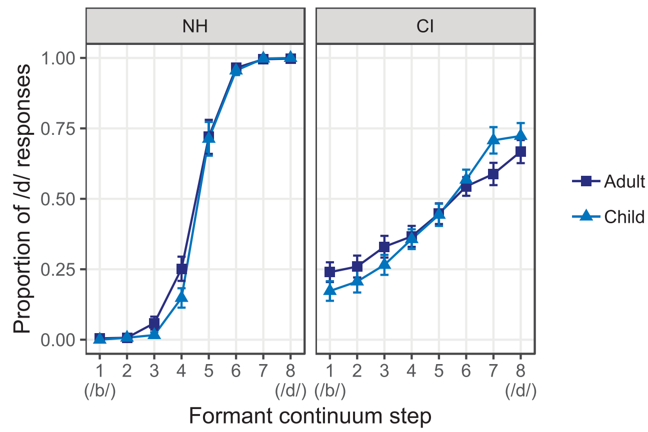
Formant cue weights were the marker of successful/unsuccessful performance in this study, and thus perceptual weighting of the formant cue was the variable of interest. Only responses to the /bɑ/–/dɑ/ continuum were analyzed, excluding the filler stimuli. A generalized logistic mixed-effects regression model (GLMM) was performed in R (R Core Development Team, 2013) using lme4 (Bates et al., 2015) with the Bound Optimization by Quadradic Approximation (bobyqa) optimizer algorithm to compare the perceptual weighting of the formant cue by hearing modality (default group = NH) and age group (default group = adults). The model formula in R syntax was as follows:

```
glmer ( /d/ ~ formant * Hearing * age +
(1+ formant | Listener) + (1 + formant | Hearing) +
(1 + formant | age), control = glmerControl(optimizer =
"bobyqa", optCtrl = list(maxfun = 2e5)), family =
"binomial")
```

Participant's responses (0 or 1, corresponding to "ba" or "da") were set as the dependent variable ("/d/" in the model formula). Fixed effects included the formant step ("formant," coded as the center continuum step) that elicited each response, as well as hearing group ("Hearing," NH or CI) and age group ("age," adult or child). Random slopes and intercepts for formant cue weights per subject ("1 + formant | Listener") in each hearing ("1 + formant | Hearing") and age ("1 + formant | age") group were also specified to account for the repeated measure of data from individual CIs of the same participants. Spectral tilt cue weights were not included as a fixed effect, as they were considered to be an orthogonal source of variability for the target cue of formant frequency and would therefore confound the model results.

To calculate both formant and spectral tilt cue-weighting coefficients for individual CIs and for the bilateral listening condition, an empirical logit transformation in R was used, as described by Winn and Litovsky (2015). Empirical logit analysis applies a fractional adjustment (0.5 in the current analysis) of success and failure to each analysis bin, proportional to the amount of data collected (c.f., Mirman, 2014). The empirical logit transformation addresses the potentially intractable outcome of "perfect separation" (responses at 0% for some continuum steps, rising to 100% for subsequent steps, with no estimable transition rate), which are more likely to occur in individual data sets because of their smaller size compared to group data sets (Barr, 2008). The adjustment scales were inversely proportional with the amount of data collected, such that it has less impact for larger data sets, essentially rewarding larger data sets with more faithful reflection of raw data. As in the basic group-level GLMM, higher empirical logit coefficients indicated greater use of the cue by an individual listener or individual CI, but with restricted range of growth. Cue-weighting coefficients equal to zero indicated no use of the cue. Coefficients less than zero signified cue usage in the opposite of the expected direction. The resulting empirical logit coefficients were used for further statistical comparisons.

Figure 3. Psychometric functions of formant cue usage for each listener group. Proportion of /d/ responses as a function of the formant cue continuum step. Steeper psychometric functions indicate greater perceptual weighting of the cue. Adult data are represented by dark blue squares and child data by light blue triangles. Error bars represent ± 1 standard error of the mean. NH = normal hearing; CI = cochlear implant.



Results

Figure 3 shows the psychometric functions for NH listeners and from the monaural listening condition for CI users, relating the formant continuum to their perception of the /b/–/d/ contrast. Psychometric function slopes were similar within a hearing modality regardless of age group. NH children and adults exhibited high perceptual reliance on the formant cue, indicated by steeply sloping psychometric functions. In contrast, children and adults with CIs demonstrated shallower psychometric functions for formant cue weighting, demonstrating low reliance on the cue. These results are in coherence with those in adults found by Winn and Litovsky (2015) and extend to children with the same hearing modalities.

The full table of GLMM results are shown in Table 2. Interactions with intercept effects were all nonsignificant, implying that all listener groups were statistically equivalent in their bias toward /b/ or /d/ perception. The essential effects of interest were the main and interactive effects of formant, described in detail below.

Consistent with observation of formant cue psychometric functions, the GLMM analyses for statistical group comparisons revealed that formant cue slopes for individual CIs and NH participants were significantly shallower for adults with CIs relative to NH adults (see Table 2: [6] $\beta = -2.09$, $z = -13.29$, $p < .001$). A significant interaction was also found between formant slope and age group (see Table 2: [7] $\beta = .49$, $z = 2.42$, $p = .02$), indicating that NH children demonstrated statistically steeper formant cue weighting slopes compared to NH adults, although this effect was relatively small. Comparison of formant psychometric functions between age groups for CI users showed that the sum of the individual effects of being (a) a CI user and (b) a child on perceptual weighting of the formant cue is equivalent to the superposition of the component main effects, as shown by a nonsignificant three-way interaction between CI hearing modality, age group and formant (see Table 2: [8] $z = -1.35$, $p = .18$). This result demonstrates an increase in formant slopes by children with CIs compared to adults with CIs. There were not statistically detectable differences in perceptual bias among any of the groups (see Table 2 [2, 3, 4]). A separate GLMM of formant cue slopes with CI user data from the bilateral condition (instead of individual CIs) showed this same pattern of results.

Although no prior evidence points to adult age as a confounding factor on the speech contrasts tested in this study, a t test was performed between formant cue weights of the "younger" and "older" NH adults to confirm the absence of adult age effects on perceptual weighting of this cue. No significant differences were found between these two NH adult groups, $t(13.7) = 1.04$, $p = .32$, suggesting that normal aging does not significantly influence use of the formant cue in this speech categorization paradigm. A sensitivity analysis was also performed by performing the main regression analysis without the younger NH adults' data. The pattern of results was unchanged.

Table 2. Results of generalized logistic mixed-effects regression models describing perception of /d/ as a function of stimulus parameters, participant age, and hearing status.

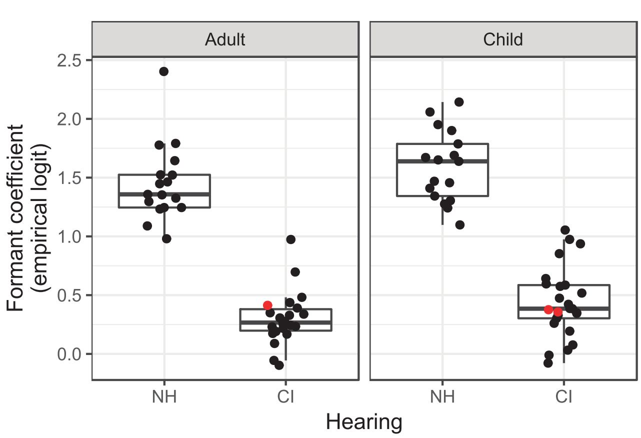
Model term	Estimate (β)	*SE*	*z*	*p*	
Intercept (bias) effects:					
Intercept (default : NH adults) [1]	–0.05	0.234	–0.215	.83	
Intercept : CI [2]	–0.327	0.306	–1.068	.285	
Intercept : age (child) [3]	–0.454	0.333	–1.364	.173	
Intercept : CI : age (child) [4]	0.388	0.429	0.905	.366	
Slope effects:					
Formant (slope) [5]	2.435	0.134	18.112	< .001	***
Formant (slope) : CI [6]	–2.09	0.157	–13.292	< .001	***
Formant (slope) : age (child) [7]	0.488	0.202	2.42	.016	*
Formant (slope) : CI : age (child) [8]	–0.311	0.231	–1.347	.178	

Note. NH = normal hearing; CI = cochlear implant.

Acoustic–Phonetic Categorization: Individual Differences

Individual variability in the use of acoustic–phonetic cues was explored using empirical logit cue weights. This allowed the direct comparison of the use of the formant cue relative to the use of spectral tilt among individuals in each hearing and age group. Figure 4 shows the range of individual scores (NH listeners and individual CIs) for the use of the formant cue, as estimated by the empirical logit functions. There is a stark separation between the NH and CI listener groups but no apparent difference in the distributions of scores between age groups within a hearing modality.

Figure 4. Empirical logit coefficients for the formant cue among the four listener groups. Box plots depict the average coefficient values within a group. Lower and upper ends of the boxes denote the 25th and 75th percentiles, respectively. Whiskers extend from the third quartile to the highest value that is +1.5 * the interquartile range and from the first quartile to the lowest value that is –1.5 * the interquartile range. The middle line of each boxplot indicates the median. Each point represents data from one individual NH listener or one CI. Red points are for CI listeners who showed a detectable bias toward the same response (/bɑ/ or /dɑ/) greater than 80% of the time while performing the task with their individual CIs. NH = normal hearing; CI = cochlear implant.



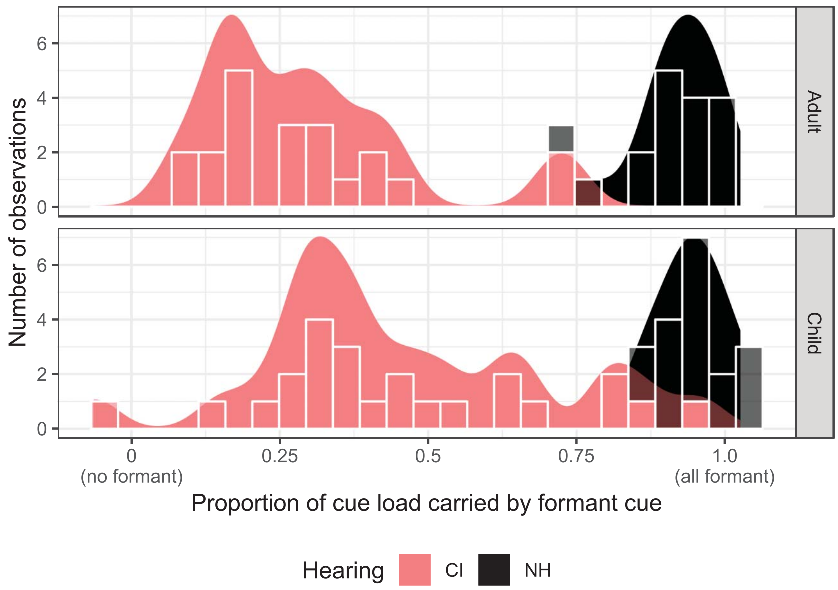
An important caveat to the interpretation of a high or low formant empirical logit coefficient is the magnitude of that coefficient relative to the sum of the formant and spectral tilt coefficients (i.e., total cue load) for the individual. For example, Listener A might have a formant weight of 0.8 and a spectral tilt weight of 0.2; if Listener B has a formant weight of 1.2 but a spectral tilt weight of 0.8, it would indicate that listener B has overall more reliable use of each cue, but that Listener A has greater reliance on the formant cue (80% of the total load) compared to Listener B (only 60% of the load). Proportionalizing the cue usage in this way allows us to examine cue reliance independent of the overall precision in using the cues. Figure 5 shows the proportion of cue weighting load carried by the formant cue for each of the four groups, demonstrating the similarity in formant cue perceptual weights between NH children and NH adults, the detriment of formant cue use in CI compared to NH listeners, and the advantage that children with CIs had over adults with CIs in use of the fine-resolution formant cue.

Visual inspection of individual formant cue psychometric slopes revealed that a few CI listeners were biased toward the same response (/bɑ/ or /dɑ/) greater than 80% of the time while performing the task with their individual CIs. A sensitivity analysis was performed by removing these participants' data (P02R, P03R, S29R) from the main GLMM. The overall pattern of results was unchanged.

Spectral Cue Trading

Winn and Litovsky (2015) had observed significant negative correlations between formant and spectral tilt cue weighting for adults with NH and with CIs, indicating a trading relationship; when participants exhibited relatively low use of one cue, they tended to increase use of the other cue. Consistent with results from that study, NH adults in the current study who relied heavily on the formant cue tended to rely less on the spectral tilt cue, shown by a significant negative correlation ($R = -.59$, $r^2 = .35$, $p = .01$). This trading relationship was also evident in the data from

Figure 5. Proportion of cue load carried by the formant cue for each listener group. Histograms are overlaid on probability density functions to specify the number of NH listeners or individual CIs that fall into each bin of formant cue use. Adult data are shown in the top panel, and child data are in the bottom panel. Data from CI listeners are in light red, and data from NH listeners are in gray (but appear darker when overlapping probability density functions). NH = normal hearing; CI = cochlear implant.



NH children (see Figure 6) but did not reach statistical significance ($R = -.25$, $r^2 = .06$, $p = .33$).

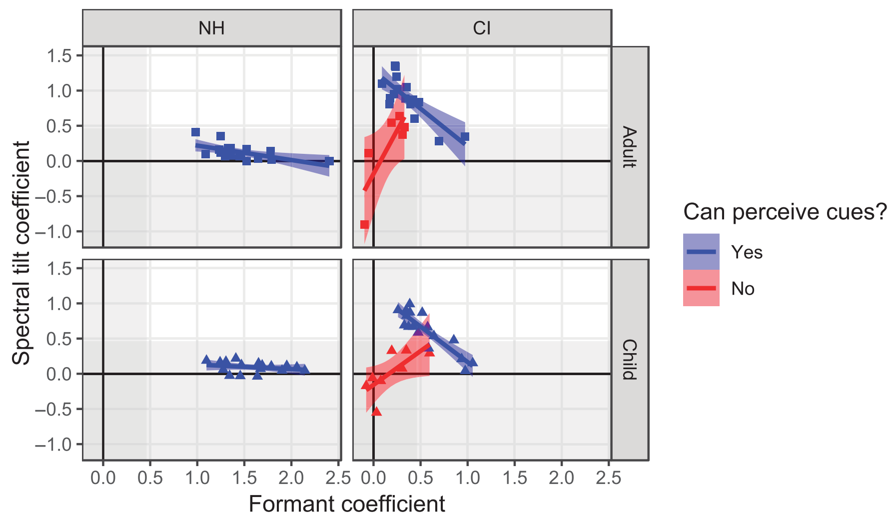
Cue-trading analysis for CI listeners in the current study turned out to be more complicated than the analysis used by Winn and Litovsky (2015) because there were some listeners who did not demonstrate use of any cues at all. At the group level, children and adults with CIs demonstrated positive, nonsignificant correlations between formant and spectral tilt cue coefficients when listening in the monaural condition. Visual inspection of individual CI cue-trading relations revealed that those CI listeners who showed better perception of the formant cue did show the expected trading relationship. However, for those who had poorer formant perception, the relationship unfolded in the opposite way; the cues were consolidated/combined rather than anticorrelated.

To handle the two distinctly different types of CI listeners, a follow-up analysis of cue weighting was conducted in which participants were separated intro groups of cue perceivers and nonperceivers, defined by the summed magnitude of their formant and spectral tilt empirical logit coefficients for each individual CI. Because outright guessing would result in coefficients of zero, this metric was used as a numerical proxy of a listener's ability to use acoustic cues to categorize the speech sounds. Data from CIs that fell below the 25th percentile of this metric were considered to result from inability to perceive the acoustic cues. Data from above and below this cutoff were analyzed separately, and correlation analyses between formant and spectral tilt cue weighting were run again. Figure 6 illustrates the cue-trading relationships for children and adults from both hearing modalities, divided into "cue perceiver" and "non-cue perceiver" groups. These analyses revealed a striking pattern in both groups of CI users that to our knowledge has not been identified previously in the literature. Children and adults with CIs with insufficient (< 25th percentile) use of the formant cue showed a positive relationship between the use of formants and spectral tilt, indicating that they combined or consolidated (rather than prioritized) cues in order to attain categorization. Conversely, those CI users with sufficient (> 25th percentile) use of the formant cue showed the typical trading relationship; they were able to selectively down-weight the tilt cue because the superior formant cue was usable. Statistically, this pattern emerged as a significant negative relationship between formant and spectral tilt cue-weighting for CI users above the 25th percentile (children: $R = -.89$, $r^2 = .79$, $p < .001$; adults: $R = -.74$, $r^2 = .54$, $p = .002$), and positive, nonsignificant relationships between these cues for those below that criterion (children: $R = .71$, $r^2 = .50$, $p = .05$; adults: $R = .56$, $r^2 = .32$, $p = .19$).

Bilateral CIs

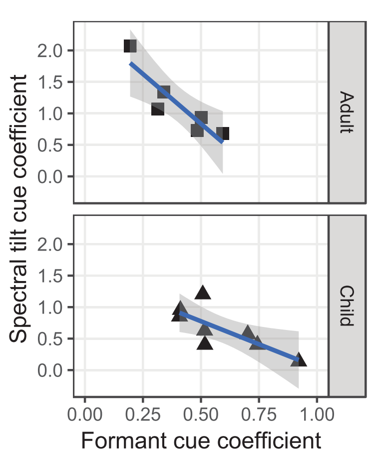
Although some CI users could not successfully utilize either cue when listening with one CI alone, all participants in the bilateral condition were able to adequately perform the task while listening with both CIs simultaneously. There were significant cue-trading relationships among children ($R = -.77$, $r^2 = .59$, $p = .027$) and adults

Figure 6. Relation between perceptual weighting of formant and spectral tilt cues for each listener group, separated by "cue perceivers" and "non-cue perceivers." Formant cue coefficients plotted against spectral tilt coefficients of participants in all four groups. Each point represents data from one individual NH listener (left panels) or one CI (right panels). Data from "cue perceivers" are in blue, and data from participants who had difficulty perceiving either cue (those with cue coefficients below the 25th percentile cutoff of 0.93) are in red. Adult data are in the top panels and are represented by squares. Child data are in the bottom panels and are represented by triangles. The lines on each plot indicate the regression lines for the relationship between formant and spectral tilt coefficients, separated by "cue perceiver" status. The shaded area around each regression line represents the 95% confidence level interval for predictions from the linear model. NH = normal hearing; CI = cochlear implant.



($R = -.90$, $r^2 = .81$, $p = .015$) in the bilateral condition. Figure 7 shows these trading relationships in the bilateral listening condition for each group of CI users.

Figure 7. Relation between perceptual weighting of formant and spectral tilt cues for cochlear implant (CI) users in the bilateral condition. Formant cue coefficients plotted against spectral tilt coefficients of CI participants. Each point represents data from one bilaterally implanted listener performing the task with both CIs simultaneously. Adult data (top panel) are represented by squares and child data (bottom panel) by triangles. The line on each plot indicates the regression line for the relationship between formant and spectral tilt cue coefficients. The shaded area around each regression line represents the 95% confidence level interval for predictions from the linear model.



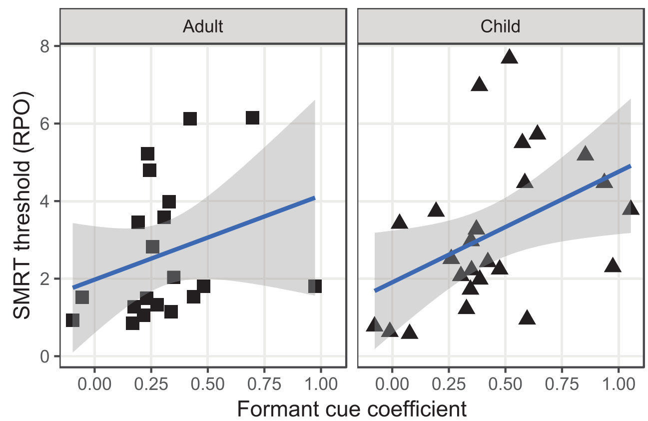
SMRT and Vowel Identification Performance

SMRT thresholds for children with CIs ranged between 0.58 and 7.68 RPO ($M = 3.15$; $Mdn = 2.51$; $SD = 1.95$) and between 0.85 and 6.15 RPO for adults with CIs ($M = 2.64$, $Mdn = 1.8$, $SD = 1.77$). Vowel identification scores in quiet ranged from 12.0 to 123.0 RAU ($M = 83.5$; $Mdn = 88.7$; $SD = 30.2$) in children and from 34.2 to 123.0 RAU in adults ($M = 87.1$, $Mdn = 92.8$, $SD = 20.2$). Results of independent-samples t tests indicated no significant differences between age groups for SMRT thresholds, $t(42.3) = 0.91$, $p = .37$, or vowel identification scores, $t(42.1) = -0.50$, $p = .62$.

Linear mixed-effects models (LMMs) using lme4 in R were performed to determine whether CI listeners' use of the fine-resolution formant cue was related to performance on vowel identification scores, a speech identification task that demands spectral processing, and the SMRT, a nonlinguistic test thought to quantify spectrotemporal resolution. SMRT thresholds and vowel identification scores were found to positively correlate with each other for children ($r^2 = .22$, $p = .02$) and adults ($r^2 = .33$, $p = .008$), and therefore, separate regression models were performed with these dependent variables. Formant cue coefficients were set as the independent variable, and a random slope effect of either vowel identification score or SMRT threshold per subject was included in all models to account for data points from two CIs of bilaterally implanted participants. Formant cue weights successfully accounted for variation in vowel identification scores for both groups (children: $\beta = 46.4$, *SE* = 17.39, *p* = .015; adults: β = 41.3, *SE* = 17.55, *p* = .029; see Figure 8).

Use of the formant cue was found to significantly relate to SMRT thresholds only in children (β = 2.84, *SE* = .98, *p* = .009; see Figure 9). This relationship in adults was not significant (*p* = .22), contrary to what one might expect based on correlation between formant cue perception and discrimination of static spectral ripples (Winn et al., 2016).

Figure 9. Relation between cochlear implant (CI) user formant cue coefficients and SMRT thresholds. Formant cue coefficients are plotted against SMRT thresholds for adult (left panel) and child (right panel) CI users. Each point represents data from one CI. Adult data are shown as squares and child data are shown as triangles. The line on each plot indicates the regression line for the relationship between variables. The shaded area around each regression line represents the 95% confidence level interval for predictions from the linear model. SMRT = Spectral-Temporally Modulated Ripple Test; RPO = ripple per octave.

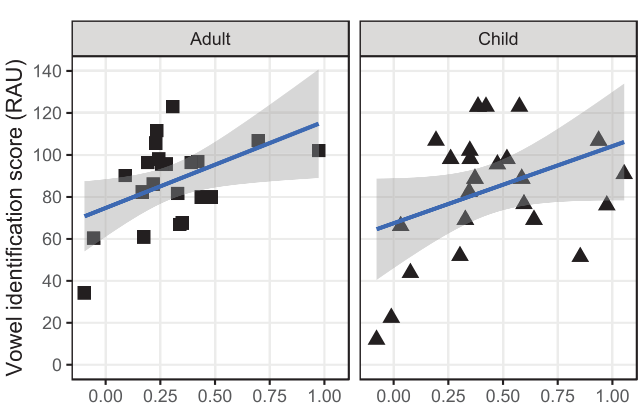


Discussion

Consistent with our hypotheses, group analyses showed that both child and adult CI users used the fine-grained spectral cue (formant transitions) less than NH listeners did. These results are in line with findings from the previous study in adults that utilized this test (Winn & Litovsky, 2015) and expanded the finding to children. To discriminate changes in spectral tilt, a listener merely needs to perform a comparison between large spectral regions and therefore does not need fine frequency resolution. It is thus sensible that, when listening with a CI, the spectral tilt cue is relatively more accessible than formant transitions.

Both groups of children showed a slight advantage in the use of formant cues compared to adults with the same hearing modality (see Figure 3). While these statistical results alone suggest that children utilize the formant cue to a higher extent than do adults regardless of hearing modality, both groups of NH listeners demonstrated high perceptual weighting of the formant cue, and small differences between age groups at that high level may not be meaningful. In addition, when variation in participant cue weight magnitude was accounted for, the distribution of formant cue load was narrow for both groups of NH listeners and essentially analogous between NH children and adults (see Figure 5), indicating that both groups used the formant cue to a similar (high) extent. In contrast, children with CIs exhibited a broader distribution of formant cue load with a peak shifted toward higher values relative to adults with CIs, demonstrating substantial differences in formant cue weighting between early-implanted children and late-implanted adults.

Figure 8. Relation between CI user formant cue coefficients and vowel identification scores. Formant cue coefficients are plotted against vowel identification scores in rationalized arcsine units (RAU) for adult (left panel) and child (right panel) CI users. Each point represents data from one CI. Adult data are shown as squares and child data are shown as triangles. The line on each plot indicates the regression line for the relationship between variables. The shaded area around each regression line represents the 95% confidence level interval for predictions from the linear model. CI = cochlear implant.



In CI users, this supports the hypothesis that hearing history differentially affects the use of perceptual cues. The children with CIs who participated in this study lost their hearing and received at least one implant at a young age. All of these children therefore learned language with the implant, and the CI has been the sole auditory experience for the majority of these children. In contrast, the adults in this study learned language with acoustic hearing and lost their hearing later in life. Receipt of a CI after decades of acoustic hearing can result in shifts in reliance on particular phonetic cues to categorize speech sounds (c.f., discussion in Winn et al., 2012). There could be a meaningful difference between *learning* to use a degraded formant cue and *re-learning* to use that same cue after years of experience perceiving it acoustically. Still, additional work is necessary before definitive conclusions can be made regarding phonetic cue acquisition in CI users.

Based on prior findings of poorer spectral pattern discrimination in NH children compared to adults (Allen & Wightman, 1992; Rayes et al., 2014), it was expected that NH children would show *less* effective use of the fine-resolution formant cue relative to adults. The findings that NH children demonstrate slightly higher average formant cue coefficients (see Figure 4) and a similar proportion of total cue load carried by the formant cue relative to NH adults (see Figure 5) may be the result of the age group of children we tested: The children in our study (ages 8–17 years), while comparable in age to the CI users in this investigation, were older than many of those in prior experiments that have found poor categorization of phonetic contrasts in children relative to adults. For example, Giezen et al. (2010) found shallower phonetic classification slopes in children aged 5–6 years relative to adult listeners, and Hazan and Barrett (2000) observed poorer and more variable phonemic boundary categorization in children aged 6–12 years compared to adults. Most of the children tested in this study were older than these ages, which might explain the generally mature formant and spectral tilt cue weighting among most of them. A study testing a larger number of children from a broader age range may identify a developmental timeline of formant and spectral tilt cue weighting for this place of articulation contrast.

Additive Versus Trading Relationships in Phonetic Cue Perception

No cue-trading relation could emerge for children and adults with CIs who could not hear either cue when listening with one CI. Interestingly, for CI listeners who had very limited use of either cue, the cues were consolidated—instead of trading one cue over the other, these participants relied on both cues to the extent that they could perceive them. This pattern of results suggests a possibility that there are two stages of phonetic cue weighting; first, where any cue is used to drive perception (as evidenced by the CI users who could not use either cue effectively, and instead relied equivalently on both) and, second, when cues can be prioritized according to their reliability and/or accordance with typical NH patterns (as evidenced by the better-performing CI listeners, who exhibited cue trading). Whether these two distinct types of cue use occur sequentially or simply idiosyncratically (based on spectral resolution) remains unknown.

The cue-trading results from this study also provide evidence for the benefits of bilateral implantation: Cue-trading relationships were only evident for individual CIs that demonstrated sufficient use of either cue to enable any perceptual weighting (see Figure 6), while all CI users in the bilateral condition demonstrated cue-trading relationships (see Figure 7). Furthermore, all listeners in the bilateral condition were able to perceive at least one of the cues, while many of those same listeners (P01, P03, P04, S23/36, S47/51) fell into the lowest 25th percentile range when tested unilaterally and therefore would have been labeled as a "non–cue perceiver" if only tested with one ear.

Although cue-trading results demonstrated that better performing listeners upweighted use of the formant cue, while downweighting use of spectral tilt, this does not consider the potential interaction between formant and spectral tilt cues. The exact step of spectral tilt could potentially influence perception of the formant cue, and thus we explored the interaction between these two cues by examining formant cue psychometric functions for each step of the spectral tilt cue continuum. Spectral tilt step was not observed to affect formant cue weighting for children or adults with NH, but the end point steps of spectral tilt (most /b/-like and most /d/-like) weakened the influence of the formant cue for CI users, particularly for adults. This finding is in line with previous work on cue trading and in agreement with the general trend of results from this study: The spectral tilt cue is more influential for CI listeners, and thus when it is informative, it should diminish the effect of formant cues. This interactive effect was not explicitly modeled statistically because, like many other cue-trading effects, it is nonmonotonic, emerging differently at continuum end points than in the continuum center.

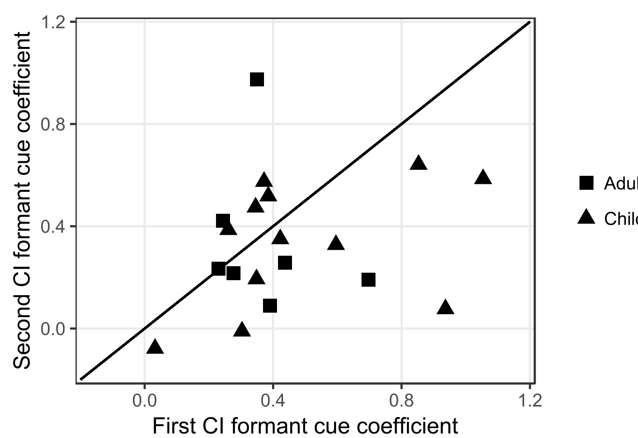
Correlation Between Tasks

Previous studies that utilized this cue-weighting task found that adult CI users with larger perceptual weights for a formant cue were those with higher word recognition scores (Winn & Litovsky, 2015; Winn et al., 2016). Vowel identification in particular is dependent on perception of formant cues, as the contrasts between vowels are primarily spectral. It thus follows that a relationship between perceptual weighting of the formant cue and vowel identification performance should exist. Consistent with this reasoning and with results from prior studies, both child and adult CI users in the current study demonstrated a significant relationship between formant cue weighting and vowel identification scores (see Figure 8). This significant, positive correlation was similar to that found between SMRT thresholds and vowel recognition performance. These findings provide additional evidence that outcomes for categorizing a single spectral speech cue are related to overall CI user speech recognition performance, while producing information about perceptual strategies used for speech perception that nonlinguistic tests of spectral resolution cannot provide.

Perceptual weighting of the formant cue was significantly related to performance on the SMRT in children with CIs, but not in adults with CIs (see Figure 9). These results in children are similar to those from Winn et al. (2016), who observed a significant relationship between adult CI users' performance on a traditional spectral ripple test and use of the formant cue. However, stimuli in traditional spectral ripple assessments and in the SMRT exhibit higher spectral density than do speech stimuli (Saoji et al., 2009). In addition, unlike the spectral ripple task used by Winn et al. (2016), the SMRT uses spectral ripples with drifting modulation phases, which introduces a temporal component into the signal that could dissociate its score from one that depends primarily on spectral resolution. It is possible that the SMRT stimuli did not contain spectrotemporal modulations that were representative of speech formant transitions, resulting in the null relationship between performance on these tasks for adult CI users, the age group with less SMRT threshold variability.

Interestingly, the child and adult CI user groups in this study were found to differ in formant cue usage but not SMRT thresholds. This suggests that the test of speech-based spectral resolution might be more sensitive to one's spectral resolving capabilities than traditional, nonlinguistic tests of spectral discrimination. In addition, the CI processer aliases and distorts the spectral and temporal components of spectral ripple stimuli at moderate and high RPO levels (Lawler et al., 2017; Winn & O'Brien, 2019). SMRT stimuli thus do not increase monotonically in any one dimension with increasing RPO level for CI users. It is, therefore, unclear whether participants with CIs who achieve high thresholds on the SMRT do indeed exhibit superior spectral discrimination abilities, or are using some alternative perceptual strategy to perform well. The stimuli in spectral discrimination tasks also do not represent the important modulations of speech (Singh & Theunissen, 2003; Winn & O'Brien, 2019), and thus the perceptual processes linking speech identification and spectral ripple discrimination are uncertain. The results from the current study suggest that speech-based tests of spectral resolution may more accurately account for speech recognition capabilities of CI users compared to the SMRT and other nonlinguistic spectral discrimination tasks.

Figure 10. Relation between formant cue coefficients from each CI of bilaterally implanted participants. Each point represents the formant cue coefficient from a participant's first-implanted CI (*x*-axis) plotted relative to the coefficient from their second-implanted CI (*y*-axis). Adult data are indicated by squares and child data by triangles. The diagonal line in the center of the plot represents perfect correlation between formant cue weights from first- and second-implanted ears. Data points that fall on this line signify similar, if not equal, formant cue coefficients for both CIs of the same participant. Data points further away from the line indicate distinct formant cue coefficients between a participant's CIs. CI = cochlear implant.



Exploring Individual CI Formant Cue Weights From Bilaterally Implanted Participants

Testing individual CIs of bilaterally implanted participants in this study allowed for assessment of phonetic cue weighting while minimizing between-subjects factors such as idiosyncratic cue-weighting tendency, intelligence quotient, cognitive abilities, and reading performance. Use of phonetic cues was observed to differ between the two CIs of most participants. In Figure 10, each bilaterally implanted participant's formant empirical logit coefficient from their first-implanted ear is plotted against the coefficient from their second-implanted ear. As only one data point falls along the center line of the plot (representing zero separation in formant coefficients between ears), this figure illustrates that formant cue weights tended to be distinct between the two CIs of the same participant. A correlation analysis also indicated no significant correlation for formant coefficients between first- and second-implanted CIs (R^2 = .05, R = .22, p = .36). These results demonstrate that at least some factors that contribute to cue-weighting patterns are related to individual CIs, rather than global perceptual strategies. Future directions of this work include investigation of peripheral and central determinants related to individual CIs that may influence perceptual cue weighting of CI listeners.

Conclusions

Phonetic cue-weighting patterns are sensitive to hearing (acoustic vs. electric/CI), age, and differences between individual implants of the same CI listener. Individuals who use CIs tend to show less use of fine spectral cues such as formant transitions and rely more on global spectral properties such as spectral tilt. This study demonstrated this pattern in adults and extended those findings to children. On average, children showed a slight advantage in using fine spectral cues relative to adults with the same hearing modality. However, when controlling for overall differences in cue use magnitudes, formant cue reliance was quite similar between NH children and adults, but diverged markedly between children and adults with CIs. This finding demonstrates a considerable difference between early-implanted children and late-implanted adults in use of a fine-resolution phonetic cue. Surprisingly, poor performing CI listeners who showed very little use of either the fine- or broad-resolution cue demonstrated an additive rather than weighted use of the cues. Use of formant cues and SMRT thresholds were related to vowel recognition scores in children and adults with CIs; yet, these two groups of CI users were differentiated in formant cue usage, but not in SMRT performance.

The phoneme categorization test showed differences between ears of the same listeners, suggesting that it could prove useful in distinguishing auditory perception resulting from different CI program processing strategies within the same person. Furthermore, phonetic cue usage supports speech perception but does not necessarily involve the potential higher order confounds of traditional word- and sentence-recognition tests; cue-weighting assessments may provide more targeted assessment of auditory abilities than traditional tests of speech identification, in which intelligibility scores could result from either typical or atypical patterns of phonetic perception.

Acknowledgments

This research was supported by National Institute on Deafness and Other Communication Disorders Grants T32 DC 005361 (awarded to M. D., PI: Perkel), R01 DC 012142 (awarded to J. G. A.), and R03 DC 014309 (awarded to M. B. W.). The authors wish to thank David Horn, Susan Norton, Margaret Meredith, and Wendy Parkinson for assistance with subject recruitment. The authors would also like to acknowledge the participants in this study for their time and effort.

References

Alexander, J. M., & Kluender, K. R. (2008). Spectral tilt change in stop consonant perception. *The Journal of the Acoustical Society of America, 123*(1), 386–396. https://doi.org/10.1121/1.2817617

Alexander, J. M., & Kluender, K. R. (2009). Spectral tilt change in stop consonant perception by listeners with hearing impairment. *Journal of Speech, Language, and Hearing Research, 52*(3), 653–670. https://doi.org/10.1044/1092-4388(2008/08-0038)

Allen, P., & Wightman, F. (1992). Spectral pattern discrimination by children. *Journal of Speech and Hearing Research, 35*(1), 222–233. https://doi.org/10.1044/jshr.3501.222

Aronoff, J. M., & Landsberger, D. M. (2013). The development of a modified spectral ripple test. *The Journal of the Acoustical Society of America, 134*(2), EL217–EL222. https://doi.org/10.1121/1.4813802

Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59*(4), 457–474. https://doi.org/10.1016/j.jml.2007.09.002

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bierer, J. A. (2007). Threshold and channel interaction in cochlear implant users: Evaluation of the tripolar electrode configuration. *The Journal of the Acoustical Society of America, 121*(3), 1642–1653. https://doi.org/10.1121/1.2436712

Bierer, J. A. (2010). Probing the electrode-neuron interface with focused cochlear implant stimulation. *Trends in Hearing, 14*(2), 84–95. https://doi.org/10.1177/1084713810375249

Boersma, P. & Weenink, D. (2019). *Praat: Doing phonetics by computer* (Version 5.3.56) [Computer program]. http://www.praat.org/

Boëx, C., Kós, M.-I., & Pelizzone, M. (2003). Forward masking in different cochlear implant systems. *The Journal of the Acoustical Society of America, 114*(4), 2058–2065. https://doi.org/10.1121/1.1610452

Busby, P. A., Plant, K. L., & Whitford, L. A. (2002). Electrode impedance in adults and children using the Nucleus 24 cochlear implant system. *Cochlear Implants International, 3*(2), 87–103. https://doi.org/10.1179/cim.2002.3.2.87

DiNino, M., & Arenberg, J. G. (2018). Age-related performance on vowel identification and the Spectral-Temporally Modulated Ripple Test in children with normal hearing and with cochlear implants. *Trends in Hearing, 22*. https://doi.org/10.1177/2331216518770959

DiNino, M., O'Brien, G., Bierer, S. M., Jahn, K. N., & Arenberg, J. G. (2019). The estimated electrode-neuron interface in cochlear implant listeners is different for early-implanted children and late-implanted adults. *Journal of the Association for Research in Otolaryngology, 20*(3), 291–303. https://doi.org/10.1007/s10162-019-00716-4

Friesen, L. M., Shannon, R. V., Baskent, D., & Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *The Journal of the Acoustical Society of America, 110*(2), 1150–1163. https://doi.org/10.1121/1.1381538

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance, 6*(1), 110–125. https://doi.org/10.1037/0096-1523.6.1.110

Giezen, M. R., Escudero, P., & Baker, A. (2010). Use of acoustic cues by children with cochlear implants. *Journal of Speech, Language, and Hearing Research, 53*(6), 1440–1457. https://doi.org/10.1044/1092-4388(2010/09-0252)

Gifford, R. H., Noble, J. H., Camarata, S. M., Sunderhaus, L. W., Dwyer, R. T., Dawant, B. M., Dietrich, M. S., & Labadie, R. F. (2018). The relationship between spectral modulation detection and speech recognition: Adult versus pediatric cochlear implant recipients. *Trends in Hearing, 22,* 2331216518771176. https://doi.org/10.1177/2331216518771176

Hadar, B., Skrzypek, J. E., Wingfield, A., & Ben-David, B. M. (2016). Working memory load affects processing time in spoken word recognition: Evidence from eye-movements. *Frontiers in Neuroscience, 10,* 221. https://doi.org/10.3389/fnins.2016.00221

Hazan, V., & Barrett, S. (2000). The development of phonemic categorization in children aged 6–12. *Journal of Phonetics, 28*(4), 377–396. https://doi.org/10.1006/jpho.2000.0121

Hedrick, M. S., & Carney, A. E. (1997). Effect of relative amplitude and formant transitions on perception of place of articulation by adult listeners with cochlear implants. *Journal of Speech, Language, and Hearing Research, 40*(6), 1445–1457. https://doi.org/10.1044/jslhr.4006.1445

Henry, B. A., & Turner, C. W. (2003). The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. *The Journal of the Acoustical Society of America, 113*(5), 2861–2873. https://doi.org/10.1121/1.1561900

Henry, B. A., Turner, C. W., & Behrens, A. (2005). Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners. *The Journal of the Acoustical Society of America, 118*(2), 1111–1121. https://doi.org/10.1121/1.1944567

Holden, L. K., Finley, C. C., Firszt, J. B., Holden, T. A., Brenner, C., Potts, L. G., Gotter, B. D., Vanderhoof, S. S., Mispagel, K., Heydebrand, G., & Skinner, M. W. (2013). Factors affecting open-set word recognition in adults with cochlear implants. *Ear and Hearing, 34*(3), 342–360. https://doi.org/10.1097/AUD.0b013e3182741aa7

Holt, L. L., & Lotto, A. J. (2010). Speech perception as categorization. *Attention, Perception, & Psychophysics, 72*(5), 1218–1227. https://doi.org/10.3758/APP.72.5.1218

Horn, D. L., Dudley, D. J., Dedhia, K., Nie, K., Drennan, W. R., Won, J. H., Rubinstein, J. T., & Werner, L. A. (2017). Effects of age and hearing mechanism on spectral resolution in normal hearing and cochlear-implanted listeners. *The Journal of the Acoustical Society of America, 141*(1), 613–623. https://doi.org/10.1121/1.4974203

Jahn, K. N., & Arenberg, J. G. (2020). Identifying cochlear implant channels with relatively poor electrode-neuron interfaces using the electrically evoked compound action potential. *Ear and Hearing*. Advance online publication. https://doi.org/10.1097/AUD.0000000000000844

Johnson, C. E. (2000). Children's phoneme identification in reverberation and noise. *Journal of Speech, Language, and Hearing Research, 43*(1), 144–157. https://doi.org/10.1044/jslhr.4301.144

Kenway, B., Tam, Y. C., Vanat, Z., Harris, F., Gray, R., Birchall, J., Carlyon, R., & Axon, P. (2015). Pitch discrimination: An independent factor in cochlear implant performance outcomes.

Otology & Neurotology, 36(9), 1472–1479. https://doi.org/10.1097/MAO.0000000000000845
Landsberger, D. M., Padilla, M., Martinez, A. S., & Eisenberg, L. S. (2018). Spectral-temporal modulated ripple discrimination by children with cochlear implants. *Ear and Hearing, 39*(1), 60–68. https://doi.org/10.1097/AUD.0000000000000463
Lawler, M., Yu, J., & Aronoff, J. M. (2017). Comparison of the Spectral-Temporally Modulated Ripple Test with the Arizona Biomedical Institute Sentence Test in cochlear implant users. *Ear and Hearing, 38*(6), 760–766. https://doi.org/10.1097/AUD.0000000000000496
Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54*(5), 358–368. https://doi.org/10.1037/h0044417
Mayo, C., & Turk, A. (2004). Adult–child differences in acoustic cue weighting are influenced by segmental context: Children are not always perceptually biased toward transitions. *The Journal of the Acoustical Society of America, 115*(6), 3184–3194. https://doi.org/10.1121/1.1738838
Mayo, C., & Turk, A. (2005). The influence of spectral distinctiveness on acoustic cue weighting in children's and adults' speech perception. *The Journal of the Acoustical Society of America, 118*(3), 1730–1741. https://doi.org/10.1121/1.1979451
Mirman, D. (2014). *Growth curve analysis and visualization using R*. CRC Press.
Moberly, A. C., Lowenstein, J. H., & Nittrouer, S. (2016). Word recognition variability with cochlear implants: "Perceptual attention" versus "auditory sensitivity." *Ear and Hearing, 37*(1), 14–26. https://doi.org/10.1097/AUD.0000000000000204
Moberly, A. C., Lowenstein, J. H., Tarr, E., Caldwell-Tarr, A., Welling, D. B., Shahin, A. J., & Nittrouer, S. (2014). Do adults with cochlear implants rely on different acoustic cues for phoneme perception than adults with normal hearing? *Journal of Speech, Language, and Hearing Research, 57*(2), 566–582. https://doi.org/10.1044/2014_JSLHR-H-12-0323
Molisz, A., Zarowski, A., Vermeiren, A., Theunen, T., De Coninck, L., Siebert, J., & Offeciers, E. F. (2015). Postimplantation changes of electrophysiological parameters in patients with cochlear implants. *Audiology and Neurotology, 20*(4), 222–228. https://doi.org/10.1159/000377615
Morrongiello, B. A., Robson, R. C., Best, C. T., & Clifton, R. K. (1984). Trading relations in the perception of speech by 5-year-old children. *Journal of Experimental Child Psychology, 37*(2), 231–250. https://doi.org/10.1016/0022-0965(84)90002-X
Munson, B., Donaldson, G. S., Allen, S. L., Collison, E. A., & Nelson, D. A. (2003). Patterns of phoneme perception errors by listeners with cochlear implants as a function of overall speech perception ability. *The Journal of the Acoustical Society of America, 113*(2), 925–935. https://doi.org/10.1121/1.1536630
Nelson, D. A., Van Tasell, D. J., Schroder, A. C., Soli, S., & Levine, S. (1995). Electrode ranking of "place pitch" and speech recognition in electrical hearing. *The Journal of the Acoustical Society of America, 98*(4), 1987–1999. https://doi.org/10.1121/1.413317
Nittrouer, S. (2002). Learning to perceive speech: How fricative perception changes, and how it stays the same. *The Journal of the Acoustical Society of America, 112*(2), 711–719. https://doi.org/10.1121/1.1496082
Nittrouer, S. (2005). Age-related differences in weighting and masking of two cues to word-final stop voicing in noise. *The Journal of the Acoustical Society of America, 118*(2), 1072–1088. https://doi.org/10.1121/1.1940508
Nittrouer, S., & Miller, M. E. (1997). Predicting developmental shifts in perceptual weighting schemes. *The Journal of the Acoustical Society of America, 101*(4), 2253–2266. https://doi.org/10.1121/1.418207
Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47*(2), 204–238. https://doi.org/10.1016/s0010-0285(03)00006-9
Ohde, R. N., Haley, K. L., & McMahon, C. W. (1996). A developmental study of vowel perception from brief synthetic consonant–vowel syllables. *The Journal of the Acoustical Society of America, 100*(6), 3813–3824. https://doi.org/10.1121/1.417338
Rayes, H., Sheft, S., & Shafiro, V. (2014). Discrimination of static and dynamic spectral patterns by children and young adults in relationship to speech perception in noise. *Audiology Research, 4*(1). https://doi.org/10.4081/audiores.2014.101
R Core Development Team. (2013). R: A language and environment for statistical computing [Computer program]. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/
Repp, B. H. (1982). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin, 92*(1), 81–110. https://doi.org/10.1037/0033-2909.92.1.81
Saoji, A. A., Litvak, L., Spahr, A. J., & Eddins, D. A. (2009). Spectral modulation detection and vowel and consonant identifications in cochlear implant listeners. *The Journal of the Acoustical Society of America, 126*(3), 955–958. https://doi.org/10.1121/1.3179670
Schertz, J., & Hawthorne, K. (2018). The effect of sentential context on phonetic categorization is modulated by talker accent and exposure. *The Journal of the Acoustical Society of America, 143*(3), EL231. https://doi.org/10.1121/1.5027512
Singh, N. C., & Theunissen, F. E. (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. *The Journal of the Acoustical Society of America, 114*(6, Pt. 1), 3394–3411. https://doi.org/10.1121/1.1624067
Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *The Journal of the Acoustical Society of America, 64*(5), 1358–1368. https://doi.org/10.1121/1.382102
Studebaker, G. A. (1985). A "rationalized" arcsine transform. *Journal of Speech and Hearing Research, 28*(3), 455–462. https://doi.org/10.1044/jshr.2803.455
Supin, A. Y., Popov, V. V., Milekhina, O. N., & Tarakanov, M. B. (1994). Frequency resolving power measured by rippled noise. *Hearing Research, 78*(1), 31–40. https://doi.org/10.1016/0378-5955(94)90041-8
ter Keurs, M., Festen, J. M., & Plomp, R. (1993). Effect of spectral envelope smearing on speech reception. *II. The Journal of the Acoustical Society of America, 93*(3), 1547–1552. https://doi.org/10.1121/1.406813
Wang, N.-Y., Eisenberg, L. S., Johnson, K. C., Fink, N. E., Tobey, E. A., Quittner, A. L., Niparko, J. K., & The CDaCI Investigative Team. (2008). Tracking development of speech recognition: Longitudinal data from hierarchical assessments in the Childhood Development after Cochlear Implantation Study. *Otology & Neurotology, 29*(2), 240–245. https://doi.org/10.1097/MAO.0b013e3181627a37
Winn, M. B., Chatterjee, M., & Idsardi, W. J. (2012). The use of acoustic cues for phonetic identification: Effects of spectral degradation and electric hearing. *The Journal of the Acoustical Society of America, 131*(2), 1465–1479. https://doi.org/10.1121/1.3672705

Winn, M. B., & Litovsky, R. Y. (2015). Using speech sounds to test functional spectral resolution in listeners with cochlear implants. *The Journal of the Acoustical Society of America, 137*(3), 1430–1442. https://doi.org/10.1121/1.4908308

Winn, M. B., & O'Brien, G. (2019). *Flaws in spectral ripple stimuli for listeners with cochlear implants*. PsyArXiv. https://doi.org/10.31234/osf.io/cbwgh

Winn, M. B., Won, J. H., & Moon, I. J. (2016). Assessment of spectral and temporal resolution in cochlear implant users using psychoacoustic discrimination and speech cue categorization. *Ear and Hearing, 37*(6), e377–e390. https://doi.org/10.1097/AUD.0000000000000328

Won, J. H., Drennan, W. R., & Rubinstein, J. T. (2007). Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. *Journal of the Association for Research in Otolaryngology, 8*(3), 384–392. https://doi.org/10.1007/s10162-007-0085-8

Xu, L., Thompson, C. S., & Pfingst, B. E. (2005). Relative contributions of spectral and temporal cues for phoneme recognition. *The Journal of the Acoustical Society of America, 117*(5), 3255–3267. https://doi.org/10.1121/1.1886405