Published in final edited form as:

Otol Neurotol. 2013 April ; 34(3): 490–498. doi:10.1097/MAO.0b013e318287c985.

Speech Intonation and Melodic Contour Recognition in Children with Cochlear Implants and with Normal Hearing

Rachel L. See, MA,

Iowa Cochlear Implant Clinical Research Center, University of Iowa Hospitals and Clinics, School of Music, University of Iowa

Virginia D. Driscoll, MA,

Iowa Cochlear Implant Clinical Research Center, University of Iowa Hospitals and Clinics

Kate Gfeller, PhD,

Iowa Cochlear Implant Clinical Research Center, University of Iowa, Hospitals and Clinics, School of Music, Department of Speech Pathology and Audiology, Iowa Cochlear, Implant Clinical Research Center, University of Iowa

Stephanie Kliethermes, MS, and

Dept. of Biostatistics University of Iowa

Jacob Oleson, PhD

Dept. of Biostatistics University of Iowa

Abstract

Background—Cochlear implant (CI) users have difficulty perceiving some intonation cues in speech and melodic contours because of poor frequency selectivity in the cochlear implant signal.

Objectives—To assess perceptual accuracy of normal hearing (NH) children and pediatric CI users on speech intonation (prosody), melodic contour, and pitch ranking, and to determine potential predictors of outcomes.

Hypothesis—Does perceptual accuracy for speech intonation or melodic contour differ as a function of auditory status (NH, CI), perceptual category (falling vs. rising intonation/contour), pitch perception, or individual differences (e.g., age, hearing history)?

Method—NH and CI groups were tested on recognition of falling intonation/contour vs. rising intonation/contour presented in both spoken and melodic (sung) conditions. Pitch ranking was also tested. Outcomes were correlated with variables of age, hearing history, HINT, and CNC scores.

Results—The CI group was significantly less accurate than the NH group in spoken (CI, $M=63.1\%$; NH, $M=82.1\%$) and melodic (CI, $M=61.6\%$; NH, $M=84.2\%$) conditions. The CI group was more accurate in recognizing rising contour in the melodic condition compared with rising intonation in the spoken condition. Pitch ranking was a significant predictor of outcome for both groups in falling intonation and rising melodic contour; age at testing and hearing history variables were not predictive of outcomes.

Conclusions—Children with CIs were less accurate than NH children in perception of speech intonation, melodic contour, and pitch ranking. However, the larger pitch excursions of the melodic condition may assist in recognition of the rising inflection associated with the interrogative form.

Keywords

Cochlear implants; pediatric; pitch perception; melodic contour; speech prosody

Introduction

Speech and music are two pervasive forms of auditory communication made up of discrete elements that have little inherent meaning, but when combined, form structures with great diversity of learned meaning.¹ Both speech and music are made up of complex sound waves with spectral and temporal envelopes that vary in time.^{2–3} However, from a perceptual standpoint, the salience of spectral and temporal features differs for speech and for music. For example, the temporal envelope provides information sufficient for conveying segmental features of speech (e.g., consonants and vowels)⁴ and the rhythmic components of music (tempo, beat, rhythm patterns).⁵ Thus, relatively good word recognition and rhythmic perception is possible with only coarse spectral information.^{4–5}

Greater spectral resolution is required, however, for those acoustic attributes of speech or music for which pitch cues are important. For speech, this includes lexical tones (i.e., tonal languages such as Mandarin Chinese or Cantonese), emotional or affective prosody (which conveys the emotional state of the speaker), as well as intonation that conveys linguistic distinctions, such as interrogative vs. declarative forms of a sentence. While duration and stress (or intensity) also contribute to perception of lexical tones or prosody, voice pitch is an important perceptual cue.⁴

One can argue that for music, in comparison with speech, pitch is a particularly salient component in perceptual accuracy, given that pitch has been described as the most basic organizing structure in most musical cultures.^{6–7} Precise frequency resolution contributes to the perception of exacting intervallic relations among musical pitches within musical scales, which form the basis for musical syntax and formal organization. The exact magnitude of pitch changes within accepted musical scales, as well as direction of pitch change contribute to the formation of melodic patterns; accurate perception of those features is important in music listening tasks such as recognition of melodic contours, familiar melodies, or detection of errors in melody patterns.^{4–5}

Differences in the temporal and spectral parameters of speech and music, along with the different perceptual skills that listening to these sounds entail, have particular implications for persons with who use cochlear implants (CIs), given the device's processing characteristics. Current CI processing strategies, which have been designed primarily to support speech recognition (e.g., word recognition, vowel or consonant recognition), usually remove the temporal fine-structure information in the stimulus waveforms and preserve the temporal envelopes extracted from 6 to 22 frequency bands. This type of signal processing provides only coarse spectral information of the signal, but is adequate to support perception of segmental features (e.g., consonants and vowels) of speech in quiet⁸ as well as rhythmic elements (e.g., duration of notes as represented in musical beats and overall tempo) in music. However, the poor spectral selectivity transmitted by the CI signal is insufficient to support perception of sounds for which pitch cues are especially salient.^{2–3,9–14}

What implications do the processing characteristics of CIs have for pediatric CI users with regard to speech perception? While a number of studies show that pediatric CI users can achieve impressive perception of segmental features of speech with extended use,⁴ a growing body of research indicates that even experienced pediatric CI recipients may have greater difficulty in perceiving components of speech for which pitch cues are especially

salient, such as suprasegmental features (e.g., tone, prosody) that convey meaning in part through intonation. For example, in studies of pediatric CI users who use tonal languages, the CI users as a group compared poorly with NH children on recognition of lexical tones in Cantonese¹⁴ and also for Mandarin Chinese, particularly for tone 2, which is made up of a rising inflection.^{9,11,13} For English speaking children, group data from CI users have revealed impaired perception of prosodic cues^{4, 15–19} such as affective prosody (identification of the emotional state of the speaker, i.e., happy or sad)³ and speech intonation (linguistic prosody) used to differentiate between declarative and interrogative sentences. It is important to note, however, that CI users show considerable variability on these tasks. For example, some “star” users may perform similarly to NH peers, while others may perform at levels below chance.^{4,12,19–20}

To better understand this variability, some of these studies have examined perceptual accuracy as a function of age when implanted, age when tested, and duration of CI use. Results are mixed regarding the impact of these variables.^{2–4,9–14} For example, of the aforementioned studies, age of implantation was significantly correlated with tone recognition in only one study⁹ which examined Mandarin Chinese tone recognition. The authors noted, however, that for the population tested, age at implantation was highly correlated with chronological age at time of testing. Duration of CI use was significantly correlated with outcomes in one of these studies. Peng⁴ found duration of CI use to be significantly correlated with speech intonation of English-speaking CI users in the task of identifying interrogative vs. declarative sentences.

Three of the aforementioned studies revealed significant correlations between perceptual outcomes and age at time of testing. These included two studies of Mandarin tone recognition.^{9,11} Peng¹¹ noted that the developmental patterns of tone acquisition were similar to that of NH children (e.g., Tone 1 and 4 acquired earlier than Tones 2 and 3), though the rate of development was slower/delayed. Peng⁴ also found age at testing to be significantly correlated with perceptual accuracy in speech intonation of English-speaking children. Interestingly, these three studies for which age at time of testing was significantly correlated with speech outcomes examined pediatric CI users comprising a broader age range of participants (at least a 6 year age range), and which also included somewhat older participants (at least 12.6 or older), thus greater maturation and greater variability with regard to age at testing may have been influential factors.

A separate body of research has investigated music perception of CI users. While the majority of extant studies regarding pitch-based music tasks have been conducted with adult CI users (for review, see Looi, Gfeller, & Driscoll;⁵ Looi;²¹ McDermott²²), a modest body of literature has compared pediatric CI recipients with normal hearing (NH) children on the pitch-based tasks of pitch pattern discrimination and melody recognition. Most of these studies have examined recognition of familiar children’s songs such as “Twinkle, Twinkle” or “Happy Birthday” (created using MIDI technology) (e.g., Hsiao;²³ Olszewski et al.;²⁴ Stordahl²⁵) or theme songs from children’s TV programs, presented either in the original recorded version or in manipulated versions presenting and contrasting the perceptual response to specific structural cues (e.g., pitch, rhythm, harmony).^{7,26–27} One study²⁷ also examined pitch pattern discrimination (detecting whether two brief pitch patterns are same or different).

While these studies varied somewhat in the stimuli parameters and response tasks (e.g., computer generated vs. naturalistic recordings; open vs. closed set), the outcomes have been relatively consistent: Group data revealed that pediatric CI recipients were significantly less accurate than NH children in recognition of familiar melodies, or pitch pattern discrimination, especially in listening conditions reliant primarily upon pitch cues (e.g., no

cues from rhythm, timbre, or lyrics). While CI recipients as a group tend to perform significantly more poorly than NH peers on pitch-based tasks, some “star” CI recipients have performed as well or nearly as well as some NH children; others have performed at chance level or below. In the studies noted above, only Hsiao²³ and Olszewski et al.,²⁴ examined age at implantation, age at testing, or duration of CI use as possible predictors of performance. Olszewski²⁴ found moderate correlations ($r=.39$, $p<.03$) for length of CI use in prelingually deaf children. Neither age of implantation nor age at time of testing was predictive of outcomes.

While the use of familiar melodies has a certain ecological validity, in that it reflects a musically-relevant perception task, melody recognition does require prior familiarity with the specific repertoire tested. What’s more, some songs have more distinctive and accessible rhythmic or timbral features that can influence perceptual accuracy. With regard to the present status of research regarding pediatric CI users, extant research regarding pediatric CI perception of structural elements of music is relatively modest in contrast with studies of speech recognition. Further research is required to better understand the perceptual accuracy of pediatric CI users for pitch-based tasks such as pitch ranking (ability to determine direction of pitch change) as well as to identify patterns within a musically-relevant context (e.g., melodic contour) that does not require prior song familiarity.

The aforementioned studies have focused on perceptual accuracy of either suprasegmental aspects of speech, or melodic patterns in music, with each study essentially addressing the uniqueness of each communicative form. However, recent research suggests that perception of some aspects of speech and music may rely on shared acoustic and neural mechanisms within the auditory system.^{28–29} Prosodic, melody, and pitch perception rely on common neural systems, and both use low level bottom up processes and higher level, top-down processes.³⁰ In relation to electric hearing (CI use), research by Wang, Zhou, and Xu,¹³ suggest that lexical and musical pitch perception are strongly correlated and may share similar mechanisms. Consequently, direct comparisons of spoken and musical stimuli may yield interesting information regarding perception of complex auditory signals.

Furthermore, studies with NH persons have suggested that extensive music experience, which requires on-going and heightened attention to fine changes in spectrally-complex sounds, may lend perceptual benefits to perception of spectrally complex features of speech. Thus, some have hypothesized that auditory training with musical elements, such as exercises in listening to pitch contours, might generalize to enhanced speech perception by persons with communication deficits (e.g., SLI, learning disabilities).^{28–32} However, most of these studies included normal-hearing listeners. Thus, additional research is required with CI users, whose auditory input does not provide good frequency resolution, in order to better understand potential overlaps in auditory processing of spoken and musical sounds for which pitch cues are salient. This could be an initial step in more critically evaluating the potential of music in rehabilitative protocols with this population.

In summary, although music and speech differ in several important aspects, such as the perceptual salience of various structural categories of sound, there are interesting areas of overlap, particularly with regard to spectrally-complex elements of each communicative form (e.g., suprasegmental features of speech, pitch and timbre in music). This overlap is of particular interest in relation to electrical hearing, which has insufficient spectro-temporal resolution to support pitch perception and components of speech more reliant on voice pitch cues. If musical and voice pitch perception share similar mechanisms in electric hearing, as suggested by Wang, Zhou, and Xu,¹³ then those CI recipients who have better pitch perception may also perform better in suprasegmental aspects of speech. It is also important

to consider other sources of variance, such as chronological age or duration of use, which may influence perceptual accuracy.

This study examined perception of both speech and music, focusing specifically on those aspects for which pitch cues are particularly salient. The following research questions were asked: (1) How do pediatric CI users and NH children compare in perception of rising and falling contours as presented within connected sentences (speech intonation) or music phrases (melodic contour)? Do the larger pitch excursions associated with the structural syntax of music (melodic contour made up of intervallic relations found in Western musical scales), in comparison with speech prosody, enhance identification of rising and falling contours? (2) Is pitch perception an influential variable in perceptual accuracy for speech intonation or melodic contour? (3) How do individual differences (e.g., age, onset of loss, age when implanted, length of implant use, hearing aid use, word and phoneme recognition, mode of communication, gender) relate to performance by CI recipients on the speech intonation or melodic contour identification? (4) Does speech intonation and melodic contour perception of pediatric implant users improve significantly as a result of more extended device use and greater chronological maturity?

Materials and Methods

Participants

Participants included 2 groups of native English speakers: 49 pediatric CI recipients (i.e., categorization as a pediatric patient in our center indicating implantation prior to the age of 18 years) (30 M, 19 F) and a comparison group of 12 (NH) children (5 M, 7 F, age 6–10, $M=7.96$). All participants were recruited in compliance with human subject guidelines. The NH children, who were recruited through advertisements posted within newspapers, completed a hearing screening to determine eligibility (25 dB HL at thresholds of 250, 500, 1000, 2000 and 4000 Hz).

CI recipients used one of the following devices: Contour, Freedom, or CI512, CI22, CI24M, or 90K. Strategies included SPEAK, ACE, or HiRes. All but 3 used a Cochlear device. The CI group included 2 age groups: (a) <15 yrs. ($n=37$, ages 7–14, $M=10.57$ years, $SD=2.34$) and (b) > 15 yrs. when tested ($n=15$, ages 15–25, $M=19.53$ years, $SD=3.04$).

Data from the NH children and the younger age group (<15 yrs.) of CI recipients were used to answer the first three research questions. The dependent variables, perception of speech intonation, melody contour, and pitch ranking were compared for the younger (< 15 years) age CI group and the NH group. Outcomes were examined in relation to individual differences (e.g., chronological age). Hearing history variables were examined as possible predictors of outcomes for the CI users.

Because several prior studies indicated that increasing maturation may contribute to perceptual accuracy,^{4,9,11} the outcomes of the younger CI group (< age 15) were subsequently compared with those of the CI users who were implanted during childhood, but who had reached age 15 or older by time of testing (Research Question 4). This analysis examined outcomes on the dependent variables as a function of longer CI use and older chronological age (greater maturation).

Test Stimuli

Melodic Contour Speech Prosody Test (MCSP)—The MCSP was used to measure recognition of falling and rising contours within melodic (sung) and spoken forms. The 10 sentences used in the MCSP were adopted from prior research by Peng et al.,⁴ which examined the accuracy of English-speaking pediatric CI recipients for speech intonation.^{4,12}

In Peng's research, each sentence (e.g., They all went to the circus) were presented in both interrogative and declarative forms, and participants were asked to judge whether each utterance was a question (someone asking) or a statement (someone telling), depending upon whether the sentence ended with a rising or falling intonation.

For the present study, the sentences used in Peng's⁴ study were recorded in two conditions: spoken and musical. The spoken stimuli (10 statements with falling intonation; 10 questions with rising intonation) were recorded by 2 female speakers using detachable PC-compatible microphones as well as built-in laptop microphones. Voice recordings occurred in closed-door environments with no background noise. All voice samples were normalized to the same RMS value and edited to include .3 seconds of silence at the beginning using Adobe Audition[®] software. The categorization of the spoken stimuli as statements (falling contour) or questions (rising contour) was validated by a panel of experts in speech-language pathology and linguistics. Because the participants in the study were 6 years of age and older, and in order to remain more consistent with the speech tokens used in Peng's research on speech prosody,⁴ we refrained from the use of the exaggerated pitch excursions of motherese more typically associated with communication with infants and toddlers.

Three experts in music, speech-language pathology, and linguistics advised on preparation of the melodic "analogs" of the spoken sentences. To create the melodic "analogs," the spoken voice samples used for this study were captured via microphone and analyzed to determine levels of frequency, intensity, and duration with Multi-Speech[®] software, a PC-based speech analysis program (KayPENTAX[®]). From this analysis, the overall contour and duration of the segmental and suprasegmental features of the spoken statements were presented in visual form.

The melodic versions included the same sentences and imitated the basic contours of the spoken sentences, but utilized the specific intervallic relations from one pitch to the next consistent with the diatonic (7-note octave) musical scales of Western culture. Because music from the Western tradition (e.g., scales and harmonic principles associated with the European tradition) is typically composed of intervals (sequential changes in fundamental frequency) that are a semitone (6% change) or larger, the magnitude of pitch changes in music is often likely to be larger than pitch excursions found in the prosody of conversational speech (as opposed to the larger pitch excursions heard in motherese with infants and toddlers). Consequently, the melodic "analogs," while containing the same text as the spoken items, included somewhat larger pitch excursions and overall contours (thus spanning a wider range of frequencies) than in the spoken versions, which reflected pitch excursions associated with normal conversational speech. The spoken items were also created to be consistent with speech tokens used in prior research by Peng et al..⁴ Melodic versions of the sentences were recorded in a simple and clear singing voice with minimal vibrato and stylistic elements that could obscure word recognition.

For purposes of this paper, the four dependent variables derived from this test will be described as spoken statements (SS), spoken questions (SQ), melodic statements (MS), and melodic questions (MQ). The statements ended with falling speech intonation or melodic contours. The questions ended with rising speech intonation or melodic contours. The spoken and sung items were evaluated by experts in music, speech-language pathology, and linguistics for suitability and overall intelligibility.

Pilot testing with NH children confirmed equivalent intelligibility of the spoken and sung sentences prior to formal testing for children as young as age 6. Pilot testing also confirmed that the children understood the task. The MCSP was administered via computer, with items presented in random order. Two practice items were administered to confirm that the

children understood the test instructions. Pictures corresponding to the topic of the stimulus appeared on the computer screen in conjunction with each sentence. After listening to each sentence, a screen prompt asked the participant to indicate if the sentence was a question or statement (2AFC). The scores were recorded as percent correct (out of 10) for each condition.

Pitch Ranking Test. (PRT)—In order to examine speech intonation and melodic contour recognition as a function of pitch perception, the PRT was administered to the CI and NH groups. This adaptive test is described in detail in Gfeller et al.³³ This test examined how accurately the participant could determine the direction of pitch change (was the second pitch higher or lower than the first?). The stimuli in this task were complex tones presented 2AFC in a range of 36 semitones (three octaves) with fundamental frequencies ranging from E (73 Hz) up to the C# (553 Hz).

The initial interval presented was set to 12 semitones. The initial tone of the pair presented was chosen randomly from the 36 pitches presented within the test. An adaptive algorithm, a variation of the staircase model, was used to find the participant's minimum pitch resolution. The maximum possible resolution was 24 semitones (2 octaves) while the smallest possible was 1 semitone, which is the smallest interval size on standard pianos and in the scale of traditional Western music.

Measures of Speech Perception—Tests of word and phoneme recognition included the HINT in quiet and in noise³⁴ and CNC scores (words, phonemes).³⁵ These scores, collected in relatively close proximity to the MCSP testing (that is, during the child's annual visit, during the previous visit), were available through the patient files and were included in the analyses in order to examine perceptual accuracy of suprasegmental speech features and melodic contour in relation to segmental elements of speech.

Individual characteristics—Demographic information (age at time of testing, gender) was recorded for all participants. Individual characteristics related to hearing history were recorded for the CI recipients. Hearing history variables included onset of loss, age when implanted, length of implant use, hearing aid use, and mode of communication.

Procedure

Test stimuli were delivered via computer in a sound-treated room with stimuli presented at 65 dB SPL; CI sensitivity was set at each CI participant's regular listening preference. Each participant completed the PRT (procedure described in Gfeller et al.³³) and then the MCSP. The instructions for the MCSP were similar to those used in the Peng et al., study.⁴ Participants were asked to determine if the spoken or sung sentence was a question or a statement, depending upon whether the sentence ended with a rising or falling intonation/contour. Before the MCSP was administered, an explanation of the difference between a statement and a question was given along with 2 practice items. No feedback was given after the practice stimuli were completed. The MCSP test lasted approximately 10 minutes.

Results

The first set of analyses compared the data from the younger (<age 15) CI recipients and the NH children. Three individuals were excluded due to missing PRT values, resulting in 49 participants for first analyses. Table 1 provides summary statistics and univariate comparisons between the CI (<age 15) and NH groups. On average, NH children were younger than CI users, with average ages of approximately 8 years and 11 years old, respectively ($p=0.0012$); however, subsequent analyses revealed no significant differences

in the dependent variables as a function of age. Table 1 presents the mean score and standard deviations on the PRT scores (smaller scores=greater accuracy). The CI group was less accurate than the NH group ($p = 0.0029$). Table 2 provides additional summary statistics on additional covariates that were relevant for CI users only.

Our first research question asked if pediatric CI users differ from NH children in perceptual accuracy of rising and falling intonation/contours in spoken and in melodic (sung) sentences. We were interested in whether the larger pitch excursions that are typically found in music based upon conventional interval changes found in Western musical scales would be more readily accessible than the more subtle pitch cues in rising and falling speech intonation of declarative and interrogative sentence forms in conversational speech. To assess potential subject response bias toward statements or questions, Wilcoxon signed rank tests were used to compare incorrect responses in regards to statement or questions for both spoken and melodic sentences. No significant difference was detected for the NH group in either the spoken or melodic categories ($p = 0.1875$ and 0.1250 , respectively). CI users were more likely to incorrectly respond “statement” to a question as opposed to “question” with a statement ($p < 0.001$). There was no significant difference between melodic statements and melodic questions ($p = 0.5171$).

Univariate comparisons for the dependent variables (SS, SQ, MS, MQ) showed significant differences in spoken and melodic questions ($p = 0.0163$ and 0.0047) (see Table 1). The NH group was more accurate than the CI group on all four dependent variables (SS, SQ, MS, MQ), with larger mean differences between the CI and NH groups for the two question conditions (SQ, MQ) than the two statement conditions. The CI group was less accurate on the combined question items (SQ, MQ) than on the combined statements (SS, MS). The two groups differed significantly on the combined spoken category (SS and SQ), the combined melodic category (MS and MQ) and the category with all four dependent variables combined ($p = 0.0225$, 0.0012 and 0.0036 , respectively). In short, NH children were able to perceive speech intonation and melodic contours more accurately than CI users, and CI users tended to have more difficulty recognizing the rising inflection in the question items.

Our second research question examined whether pitch perception was an influential variable in perceptual accuracy on any of the stimulus categories (SS, SQ, MS, MQ). A separate logistic regression analysis was conducted on each dependent variable (SS, SQ, MS, MQ) to test for a difference in percent of sentences correctly identified between the NH group and CI group. In order to account for potentially influential variables in addition to the PRT in the primary dependent variables (SS, SQ, MS, MQ), the following covariates of interest were examined: group (NH, CI), age when tested, gender, and all possible two-way interactions. Age was not significant in all models and excluded from consideration. After conducting goodness of fit analyses, all other variables remained in each model with some interactions present.

Results for all variables in each logistic regression analysis are provided in Table 3. Odds ratios and 95% confidence intervals are given for significant variables. In all models, the PRT was significantly associated with the dependent variable; however, the relationship between PRT and 3 of the dependent variables (SQ, MS, MQ) differed depending on group (NH or CI). Due to the presence of interactions in each model, odds ratios for variables included in an interaction were interpreted at the level of the interaction. For example, the odds of correctly identifying a spoken question for a NH individual with a given PRT were 1.20 times the odds of correctly identifying a spoken question for a NH individual with a PRT score that is one unit higher (less accurate), after adjusting for other covariates ($p < 0.0001$). Similar conclusions can be drawn for melodic statements and melodic questions.

Thus, these data suggest a strong relationship between pitch perception and accuracy on both speech prosody and melodic contour, regardless of group membership (CI, NH).

Figure 1 shows the relationship between PRT and the dependent variables for NH and CI groups. In all cases, as PRT increases, the percent correct decreases, on average. However, the relationship differed depending on whether an individual was NH or had CIs among all outcomes except spoken statement.

The third research question examined how individual differences in the CI users (< age 15) (e.g., age, age of onset, age when implanted, length of implant use, hearing aid use, perception on segmental speech tasks, mode of communication, gender) related to performance on the dependent variables. Potentially influential variables regarding hearing history and speech perception for CI users were obtained during an annual visit in close proximity (e.g., typically the same day, or within one year) to the administration of the MCSP and PRT. Continuous predictors of interest included HINT noise, HINT quiet, CNC-words, CNC-phonemes and months of use (MOU) scores as well as the child's age at implantation. The categorical predictors include a yes/no variable for HA, the type of CI, and gender. Secondary analyses were then conducted to determine which variables were significantly related to each of the dependent variables for the CI group. Due to the high correlation among the covariates, regression models are difficult to create and interpret. Therefore, pairwise Spearman correlations between the dependent variables and the continuous covariates were calculated and Kruskal-Wallis tests for categorical covariates were conducted to measure association with each of the dependent variables.

Results of the secondary analyses were mixed. Those variables, which yielded statistically significant correlations, were as follows: The HINT-quiet ($r=0.498$, $p=0.004$) and PRT ($r=-0.573$, $p<0.001$) were statistically correlated with spoken statements. The HINT-quiet ($r=0.609$; $p<0.001$), PRT ($r=-0.567$; $p<0.001$), CNC words ($r=0.425$; $p=0.015$), and CNC phonemes ($r=0.416$; $p=0.018$) were statistically correlated with melodic questions. None of the above variables were significantly correlated with spoken questions and melodic statements. The following variables were not significantly related to any of the dependent variables: HA use, gender, age, age of onset, age when implanted, length of implant use, device type, processing strategy, and communication mode (TC, oral, ASL). Thus, very few of the variables that have often been associated with greater perception or production of segmental features of speech were associated with better outcome on recognition of speech intonation and melodic contour.

Our fourth research question asked if speech intonation and melodic contour perception improved as a function of longer CI use and greater chronological maturity? We compared the younger CI group (< 15 yrs.) to those implanted during childhood who had reached age 15 or older by the time of testing. Wilcoxin rank sum tests conducted on all measures revealed no statistically significant differences (p values > 0.05) on any measure as a function of CI age group (< age 15 yrs.; > age 15 yrs.).

Discussion

These findings are consistent with prior studies indicating that pediatric CI users have difficulty in recognition of speech intonation and melodic contours, both perceptual tasks for which F0 is a salient cue. Better pitch ranking was predictive of perceptual accuracy for NH as well as CI groups in spoken (SS) as well as melodic conditions (MQ), suggesting shared mechanisms for some aspects of speech and music in persons whose auditory system has developed through electric as well as those receiving auditory stimulation. It is important to acknowledge, however, that F0 may have not been the only cue for recognizing the

difference between questions and statements. Intensity and duration patterns, both of which are effectively transmitted through electrical hearing, could have also been a factor in recognition.

Interestingly, the difficulty demonstrated by the CI group in identifying a rising contour in the SQ and MQ conditions is consistent with other studies with CI users noting problems in perception or production of spoken questions,^{4,20} as well as lexical tones with a rising contour (e.g., tone 2, in Mandarin Chinese).^{11,13} Although the melodic condition, which presented larger pitch excursions than the spoken sentences, did not improve identification of falling contour (statements), the more exaggerated contour in the melodic question condition resulted in better recognition of rising contour (questions) for both the CI and NH groups. This suggests that perhaps some music-based training, such as exercises in listening to pitch contours, which has generalized to speech perception for some types of communication disorders (e.g., SLI), might also be useful in habilitation of speech intonation for some children with CIs.

Consistent with several prior studies,^{4,11} age at implantation, device type, length of CI use, and mode of communication were not predictive of results on the prosody and contour variables. This may relate to how poorly spectrally-complex aspects of speech and music are represented through the CI signal. This degraded signal may interfere with the sort of adaptation/learning that occurs over time for segmental features of speech. Admittedly, due to the variety of devices and small subject sample, it is difficult to determine with certainty whether or not the type of CI device, processor, or strategy was influential in the outcomes.

Age at time of testing has been associated with greater results in some studies.⁴ In Peng's 2004 study, children with CIs showed progress with the use of rising intonation from 0–7 years post-implantation; however, there was no further improvement and declination observed following 7 years of device experience.¹¹ In our study, age at time of testing and length of CI use were not predictive of any of the 4 listening conditions, nor was there a significant difference on the MCSP results between the two CI age groups (children, adolescent CI users). Because studies of adult CI recipients have identified sequential cognitive processing as a predictor of pitch and melodic pattern discrimination,^{36,37} future studies should also examine the possible impact of cognitive processes in relation to perceptual accuracy of pediatric CI recipients for speech features for which pitch cues are salient.

Acknowledgments

This study was supported by grant 2 P50 DC00242 from the NIDCD, NIH; grant RR00059 from the General Clinical Research Centers Program, NCRR, NIH; and the Iowa Lions Foundation. Thanks are due to the participants in this study, Christina Scheperle for assistance in data collection, and Professors Richard Hurtig, Karla McGregor and Jean Gordon for assistance in developing the test stimuli.

Abbreviations

CI	cochlear implant
NH	normal hearing
SS	spoken statement
SQ	spoken question
MS	melodic statement
MQ	melodic question

MCSP	Melodic Contour Speech Prosody test
HINT	Hearing in Noise Test
CNC	Consonant Nucleus Consonant
HA	Hearing Aid

REFERENCES

1. Patel, AD. Music, Language and the Brain. Oxford University Press; 2008. p. 1-10.
2. Chatterjee M, Peng S. Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition. *Hear Res.* 2008; 235:143–156. [PubMed: 18093766]
3. Hopyan-Misakyan TM, Gordon KA, Dennis M, Papsin B. Recognition of affective speech prosody and facial affect in deaf children with unilateral right cochlear implants. *Child Neuropsychol.* 2009; 15:136–146. [PubMed: 18828045]
4. Peng S, Tomblin JB, Turner CW. Production and perception of speech intonation in pediatric cochlear implant recipients and individuals with normal hearing. *Ear Hear.* 2008; 29:336–351. [PubMed: 18344873]
5. Looi V, Gfeller K, Driscoll V. Music appreciation and training for cochlear implant recipients: A review. *Semin Hear.* 2012; 33:307–334. [PubMed: 23459244]
6. Patel AD, Peretz I, Tramo M, Labreque R. Processing prosodic and musical patterns: A neuropsychological investigations. *Brain Lang.* 1998; 61:123–144. [PubMed: 9448936]
7. Nakata T, Trehub SE, Mitani C, Kanda Y. Pitch and timing in the songs of deaf children with cochlear implants. *Music Percept.* 2006; 24:147–154.
8. Wilson, B. Cochlear implant technology. In: Niparko, JK.; Kirk, KI.; Mellon, NK.; Robbins, AM.; Tucci, DL.; Wilson, BS., editors. *Cochlear implants: Principles and practices.* New York: Lippincott, Williams & Wilkins; 2000. p. 109-118.
9. Han D, Liu B, Zhou N, et al. Lexical tone perception with HiResolution and HiResolution 120 sound-processing strategies in pediatric mandarin-speaking cochlear implant users. *Ear Hear.* 2009; 30:169–177. [PubMed: 19194297]
10. Lenden JM, Flipsen P. Prosody and voice characteristics of children with cochlear implants. *J Commun Disord.* 2007; 40:66–81. [PubMed: 16765979]
11. Peng S, Tomblin JB, Cheung H, Lin YS, Wang L. Perception and production of mandarin tones in prelingually deaf children with cochlear implants. *Ear Hear.* 2004; 25:251–264. [PubMed: 15179116]
12. Peng S, Tomblin JB, Spencer LJ, Hurtig RR. Imitative production of rising speech intonation in pediatric cochlear implant recipients. *J Speech Lang Hear Res.* 2007; 50:1210–1227. [PubMed: 17905907]
13. Wang W, Zhou N, Xu L. Musical pitch and lexical tone perception with cochlear implants. *Int J Audiol.* 2011; 50:270–278. [PubMed: 21190394]
14. Wong A, Wong L. Tone perception of cantonese-speaking prelingually hearing-impaired children with cochlear implants. *Otolaryngol Head Neck Surg.* 2004; 130:751–758. [PubMed: 15195063]
15. Adi-Bensaid L, Most T. The acquisition of the prosodic word by children with hearing loss using a cochlear implant. *The Volta Review.* 2009; 109:5–31.
16. Meister H, Landwehr M, Psychny V, Walger M, von Wedel H. The perception of prosody and speaker gender in normal-hearing listeners and cochlear implant recipients. *Int J Audiol.* 2009; 48:38–48. [PubMed: 19173112]
17. Most T, Peled M. Perception of suprasegmental features of speech by children with cochlear implants and children with hearing aids. *J Deaf Stud Deaf Ed.* 2007; 12:350–361.
18. Snow D, Ertmer D. The development of intonation in young children with cochlear implants: a preliminary study of the influence of age at implantation and length of implant experience. *Clin Ling Phon.* 2009; 23:665–679.

19. Peng, S. [Doctoral]. The University of Iowa; 2005. Perception and production of speech intonation in pediatric cochlear implant recipients and children with normal hearing.
20. Phan, J.; Bergeson-Dana, T.; Sayed, Z.; Chin, S. Speech intelligibility and prosody in children with cochlear implants. Presented at the 13th Symposium on Cochlear Implants in Children; July 14–16, 2011; Chicago, Illinois.
21. Looi V. The effect of cochlear implantation on music perception: A review. *Otorinolaryngol.* 2008; 58:169–190.
22. McDermott H. Music perception with cochlear implants: A review. *Trends Amplif.* 2004; 8:49–81. 2004. [PubMed: 15497033]
23. Hsiao F. Mandarin melody recognition by pediatric cochlear implant patients. *J Mus Ther.* 2008; 45:390–404.
24. Olszewski C, Gfeller K, Froman R, Stordahl J, Tomblin B. Familiar melody recognition by children and adults using cochlear implants and normal hearing children. *Cochlear Implant Int.* 2005; 6:123–140.
25. Stordahl J. Song recognition and appraisal: A comparison of children who use cochlear implants and normally hearing children. *J Mus Ther.* 2002; 39:2–19.
26. Trehub SE, Vongpaisal T, Nakata T. Music in the lives of deaf children with cochlear implants. *Ann NY Acad Sci.* 2009; 1169:534–542. [PubMed: 19673836]
27. Vongpaisal T, Trehub SE, Schellenberg EG. Song recognition by children and adolescents with cochlear implants. *J Speech Lang Hear Res.* 2006; 49:1091–1103. [PubMed: 17077217]
28. Shahin A. Neurophysiological influence of musical training on speech perception. *Frontiers in Psychology.* 2011; 2(0)
29. Strait DL, Kraus N, Skoe E, Ashley R. Musical experience promotes subcortical efficiency in processing emotional vocal sounds. *Ann N Y Acad Sci.* 2009; 1169:209–213. [PubMed: 19673783]
30. Chermak G. Music and auditory training. *Hear J.* 2010; 63:57–58.
31. Patel A. Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Front Psychol.* 2011; 2
32. Schon D, Magne C, Besson M. The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology.* 2004; 41:341–349. [PubMed: 15102118]
33. Gfeller K, Turner C, Mehr M, Woodworth G, Fearn R, Knutson JF, Witt S, Stordahl J. Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. *Cochlear Implant Int.* 2002; 3:29–53.
34. Nilsson M, Soli SD, Sullivan JA. Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am.* 1994; 95:1085–1099. [PubMed: 8132902]
35. Peterson GE, Lehiste I. Revised CNC lists for auditory tests. *J Speech Hear Disord.* 1962; 27:62–70. [PubMed: 14485785]
36. Gfeller K, Oleson J, Knutson JF, Driscoll V, Olszewski C. Multivariate predictors of music perception and appraisal by adult cochlear implant users. *J Am Acad Audiol.* 2008; 19:120–134. [PubMed: 18669126]
37. Gfeller K, Jiang D, Oleson J, Driscoll V, Knutson JF. Temporal stability of music perception and appraisal scores of adult cochlear implant recipients. *J Am Acad Audiol.* 2010; 21:28–34. [PubMed: 20085197]

Figure 1.
Interaction between PRT and Group for 4 dependent variables

Table 1

Summary Statistics for CI and NH patients

	Cochlear Implants N=37 ¹	Normal Hearing N=12	p-value ²
Variables	Count, % or Mean (SD)	Count, % or Mean (SD)	
Covariates			
Age (y)	10.57 (2.34)	7.96 (1.41)	0.0012
PRT (exact score, min.=1)	14.12 (9.15)	7.33(8.67)	0.0029
Gender			
Male	21 (61.76%)	5(41.67%)	0.3140
Female	13 (38.24%)	7 (58.33%)	
Dependent Variables			
Spoken Statement	77.0%(25.3)	88.3% (20.4)	0.1120
Spoken Question	49.2%(32.5)	75.8% (34.2)	0.0163
Melodic Statement	58.6% (33.1)	75.0% (32.1)	0.1054
Melodic Question	64.6% (33.8)	93.3% (14.9)	0.0047
Dependent Variables Combined			
Spoken	63.1% (18.2)	82.1% (24.5)	0.0225
Melodic	61.6% (19.9)	84.2% (17.9)	0.0012
All	62.4% (17.1)	83.1% (20.5)	0.0036

¹Summary statistics exclude the 3 individuals (all cochlear implant patients) removed from logistic regression analysis due to missing PRT scores

²P-values reflect Wilcoxon Rank Sum Tests for continuous variables and fisher's exact tests for categorical variables to assess differences between the two groups

Table 2

Additional Covariates for CI Users

	N	Cochlear Implants
Covariates		Count, %, or Mean (SD)
Age at Implant (months)	37	34.43 (23.75)
Months CI use (MOU)	37	92.32 (35.36)
HINT noise	23	73.61% (22.70)
HINT quiet	31	84.06% (17.44)
CNC words	32	74.44% (17.04)
CNC phonemes	32	88.38% (10.33)
Device Code	37	
Contour, Freedom, CI512		30 (81.08%)
CI22		0 (0%)
CI24M		5 (13.51%)
90K		2 (5.41%)
HA Y/N	37	
No		33 (89.19%)
Yes		4 (10.81%)

Table 3

Logistic Regression Analyses for dependent outcomes

Outcome	Predictor	Estimate	χ^2 Test Statistic	Odds Ratios (95% Confidence Interval)	p-value
Spoken Statement	Group^I	-0.8171	3.2289	--	0.0724
	PRT	-0.1258	51.6583	1.134 (1.095, 1.174)	<.0001
	Gender²	-0.4268	2.0648	--	0.1507
	Group*Gender	1.9479	6.1943		0.0128
	<i>NH*Gender</i>			4.577 (1.121, 18.686)	0.0340
	<i>CI*Gender</i>			0.653 (0.365, 1.168)	0.1507
Spoken Question	Group	3.0500	33.4035	--	<.0001
	PRT	-0.00282	0.0521	--	0.8195
	Gender	0.7168	11.2834	2.048 (1.348, 3.111)	0.0008
	PRT*Group	-0.1754	26.2100		<.0001
	<i>NH*PRT</i>			1.195 (1.124, 1.272)	<.0001
	<i>CI*PRT</i>			1.003 (0.978, 1.028)	0.8195
Melodic Statement	Group	2.0098	20.4295	--	<.0001
	PRT	-0.0040	0.1026	--	0.7487
	Gender	0.4418	4.5419	--	0.0336
	PRT*Group	-0.1239	16.8529		<.0001
	<i>NH*PRT</i>			1.136 (1.078, 1.199)	<.0001
	<i>CI*PRT</i>			1.004 (0.979, 1.029)	0.7487
Melodic Question	Group	8.1319	3.9378	--	0.0472
	PRT	-0.0964	44.6335	--	<.0001
	Gender	-0.5764	5.3428	--	0.0208
	PRT*Group	-0.3477	3.0926		0.0786
	<i>NH*PRT</i>			1.560 (1.059, 2.294)	0.0242
	<i>CI*PRT</i>			1.110 (1.071, 1.133)	<.0001

^I. Reference category = CI group.

²Reference category = females