Published in final edited form as:
Ear Hear. 2015 January ; 36(1): 102–110. doi:10.1097/AUD.0000000000000086.

Melodic Pitch Perception and Lexical Tone Perception in Mandarin-Speaking Cochlear Implant Users

Duoduo Tao[1,2], **Rui Deng**[1], **Ye Jiang**[1], **John J. Galvin III**[2,3], **Qian-Jie Fu**[2,3], and **Bing Chen**[1]

[1]Department of Otology and Skull Base Surgery, Eye Ear Nose and Throat Hospital, Fudan University, Shanghai, China

[2]Division of Communication and Auditory Neuroscience, House Research Institute, Los Angeles, California, USA

[3]Department of Head and Neck Surgery, David Geffen School of Medicine, UCLA, Los Angeles, CA, USA

Abstract

Objectives—To examine the relationship between lexical tone perception and melodic pitch perception in Mandarin-speaking cochlear implant (CI) users, and to investigate the influence of previous acoustic hearing on CI users' speech and music perception.

Design—Lexical tone perception and melodic contour identification (MCI) were measured in 21 prelingual and 11 postlingual young (age: 6–26 years old) Mandarin-speaking CI users. Lexical tone recognition was measured for four tonal patterns: Tone 1 (flat F0), Tone 2 (rising F0), Tone 3 (falling-rising F0), and Tone 4 (falling F0). MCI was measured using 9 five-note melodic patterns that contained changes in pitch contour, as well as different semitone spacing between notes.

Results—Lexical tone recognition was generally good (overall mean = 81% correct), and there was no significant difference between subject groups. MCI performance was generally poor (mean = 23% correct). MCI performance was significantly better for postlingual (mean = 32% correct) than for prelingual CI participants (18% correct). After correcting for outliers, there was no significant correlation between lexical tone recognition and MCI performance for prelingual or post-lingual CI participants. Age at deafness was significantly correlated with MCI performance only for postlingual participants. CI experience was significantly correlated with MCI performance for both prelingual and postlingual participants. Duration of deafness was significantly correlated with tone recognition only for prelingual participants.

Conclusions—Despite the prevalence of pitch cues in Mandarin, the present CI participants had great difficulty perceiving melodic pitch. The availability of amplitude and duration cues in lexical tones most likely compensated for the poor pitch perception observed with these CI listeners. Previous acoustic hearing experience seemed to benefit postlingual CI users' melodic pitch

Address for correspondence: Bing Chen, Department of Otology and Skull Base Surgery, Eye Ear Nose and Throat Hospital, Fudan University, 83 Fenyang Road, Shanghai 200031, China. entdtao@gmail.com.

Conflicts of Interest and Source of Funding

Portions of this article were presented as a poster at the 2013 Conference on Implantable Auditory Prostheses, Lake Tahoe, California, July 14, 2013.

perception. Longer CI experience was associated with better MCI performance for both subject groups, suggesting that CI users' music perception may improve as they gain experience with their device.

Introduction

Cochlear implants (CIs) provide good speech understanding to many profoundly deaf individuals (Wilson et al. 1991; Miyamoto et al. 1996; Niparko 2004; Shannon et al. 2004; Geers and Hayes 2011). Although CIs typically transmit 16 to 22 spectral channels, CI users' functional spectral resolution may be limited to only 8 channels (Friesen et al. 2001), due to channel interaction. Four to eight spectral channels may be adequate for speech understanding in optimal listening conditions, but complex listening tasks such as speech understanding in noise, talker identification, and music perception require more than 32 channels to maintain good performance (Friesen et al. 2001; Smith et al., 2002; Shannon et al. 2004; Vongphoe and Zeng, 2005). Current CI technology does not provide sufficient information for complex pitch perception, which makes music perception and appreciation difficult for CI users (Smith et al. 2002; Mirza et al. 2003; Shannon et al. 2004; Vandali et al. 2005; Looi et al. 2012). CIs also do not provide voice pitch cues important for perception of vocal emotion (Xin et al. 2007), speech prosody (Chatterjee and Peng 2008), and lexical tones (Peng et al. 2004; Morton et al. 2008; Han et al. 2009).

Pitch cues provide lexical meaning in tonal languages such as Mandarin Chinese (Fu et al. 1998). There are four tonal patterns in Mandarin Chinese, which can be characterized by the variation in the fundamental frequency (F0) during voiced speech (Peng et al. 2004). In general, Tone 1 has a flat and high F0 pattern, Tone 2 has a rising F0 pattern, Tone 3 has a falling and then rising F0 pattern, and Tone 4 has a falling F0 pattern. Unlike NH listeners, CI users have limited access to F0 cues due to the coarse spectral resolution. Although F0 is represented in the temporal envelopes within individual frequency channels, CI users are able to extract only some of the temporal pitch information, and only for relatively low F0's (Green et al. 2004). Even with very high stimulation rates (e.g., 2000 pulses per second or greater), CI users' temporal pitch perception generally degrades for F0s greater than 300 Hz (Chatterjee and Ozerbut 2011; Fraser and McKay 2012).

Although the principal phonetic feature for Chinese tones is the F0 contour (Abramson 1978), other acoustic characteristics that co-vary with changes in F0 may also contribute to tone recognition. One such acoustic cue is vowel duration (Fu and Shannon 1998; Fu et al. 2004). The falling-rising tone (Tone 3) is generally longest in vowel duration, whereas the falling tone (Tone 4) is shortest. Another cue is the acoustic amplitude. The falling-rising tone (Tone 3) is generally produced with the lowest peak amplitude, whereas the falling tone (Tone 4) is the highest. Because of limited F0 cues, CI users may rely on duration and amplitude envelope cues to recognize Chinese tones (Fu et al. 1998; Fu et al. 2004). Luo et al. (2008) reported only moderately good lexical tone recognition in a study with Mandarin-speaking CI users. In another study, Luo et al. (2009) measured single- and concurrent-syllable recognition in 4 adolescent and 4 adult CI participants. Results showed mean tone recognition with single syllables was only moderately good (63% correct); scores dropped by 40–60 percentage points with concurrent syllables, reflecting poor access to F0 information.

Pitch cues are vital to music perception and appreciation. Previous CI studies have shown that music enjoyment declines substantially after implantation (Lassaletta et al. 2007). CI users consistently perform poorer than normal-hearing (NH) listeners in a familiar melody identification task (Kong et al. 2004; McDermott 2004; Kong et al. 2005; Looi et al. 2008). Many studies have measured CI users' musical pitch discrimination. Kang et al. (2009) found a moderate correlation between pitch discrimination and familiar melody recognition. Fujita and Ito (1999) found a wide range (4 – 24 semitones) in CI users' pitch ranking thresholds. Looi et al (2008) found that CI users were unable to reliably rank pitches that were 3 semitones apart (mean = 52% correct), which was not significantly different from the chance level performance (50% correct). Pitch ranking performance was only moderately good even with 6-semitone (mean = 64% correct) or 12-semitone spacing (mean = 68% correct); both were significantly better than chance level performance score (50% correct; $p < 0.05$ in both cases). Sucher and McDermott (2007) found that mean pitch ranking performance was significantly poorer for CI users than for NH listeners, for both 1 semitone (CI = 49% correct; NH = 81% correct; $p < 0.01$) and 6 semitone intervals (CI = 60% correct; NH = 89% correct; $p < 0.01$); chance-level performance was 50% correct. Other studies have used more complex tasks to probe the limits of CI users' music perception. Cooper et al. (2008) used the Montreal battery for evaluation of amusia (MBEA) to identify specific deficits in the pitch-based aspects of music perception in CI users. Results showed that CI users had great difficulty with various tests of melodic pitch perception, with mean scores comparable to those of NH participants listening to acoustic CI simulations with 4 to 6 spectral channels.

Melodic contour identification (MCI) has also been used to measure CI users' melodic pitch perception (Galvin et al. 2007; Galvin et al. 2008; Galvin et al. 2009a 2009b; Zhu et al. 2011). For postlingually deafened adult CI users, MCI performance has been shown to be highly variable (Galvin et al. 2007); with simple stimuli, scores ranged from 14% to 91% correct. MCI performance may also be significantly affected by instrument timbre (Galvin et al. 2008) or by a competing instrument (Galvin et al. 2009a; Zhu et al. 2011). Bandpass filtering the acoustic input has been shown to improve some CI users' MCI performance (Galvin and Fu 2011), and even greater improvements have been observed with MCI training (Galvin et al. 2007; Galvin et al. 2012).

Given the strong contribution of pitch cues to both melodic pitch and lexical tone perception, performance in these tasks would seem to be correlated in Mandarin-speaking CI listeners. Wang et al. (2011) measured lexical tone perception and pitch interval discrimination in 19 adult Mandarin-speaking CI users. Results showed a strong correlation between CI users' tone recognition and melodic pitch discrimination thresholds. Wang et al. (2012) measured lexical tone perception and music perception in adult Mandarin-speaking CI users. Results showed that tone perception was significantly correlated with pitch discrimination, melody discrimination, and instrument identification. These findings suggest that musical and lexical pitch perception might share similar mechanisms in electric hearing.

Most Mandarin-speaking Chinese CI users have been implanted at an early age, and many are prelingually deafened. Prelingual CI users develop pitch patterns exclusively with electric hearing, in which F0 information is poorly transmitted and poorly perceived. For tone recognition, prelingual CI users may weight other cues such as amplitude and duration more strongly than F0 (Peng et al. 2009, 2012). For melodic pitch perception, the poor spectral resolution of the CI may limit or even prevent the development of central pitch patterns for prelingual CI users. This may in turn limit prelingual CI users' music perception. Postlingually deafened CI users adapt to electric hearing in the context of previous acoustic hearing, whether unaided or aided by a hearing aid (HA). Central pitch patterns might be stronger in postlingual CI users, given this previous acoustic hearing experience which would have provided the fine-structure cues needed for complex harmonic pitch. While pitch information would certainly be degraded by CI signal processing, postlingual CI users may be better able to adapt the electric pitch patterns to the previous acoustic patterns.

In this study, lexical tone and music perception were measured in young prelingually or postlingually deafened Mandarin-speaking CI users. Music perception was measured using melodic contour identification (MCI; Galvin et al., 2007; 2008; 2009a b; 2012). In the MCI task, listeners must explicitly use pitch information to identify the melodic contours. Due to their previous experience with acoustic hearing, it was hypothesized that postlingual CI users would outperform prelingual CI users in both tone recognition and music perception, as pitch perception would be expected to contribute strongly to both listening tasks. Demographic factors such as age at profound deafness, CI experience, and duration of deafness were compared to tone and music perception. For both groups, it was hypothesized that tone and music perception would be positively correlated with CI experience and negatively correlated with duration of deafness. For postlingual CI users, it was hypothesized that the age at deafness (roughly equivalent to the amount of acoustic hearing experience before implantation) would be positively correlated with tone and music perception. Finally, it was hypothesized that tone and music perception performance would be correlated for both groups, given the importance of pitch cues to both listening tasks.

Materials and methods

Participants

Thirty-two CI users (21 prelingually deafened and 11 postlingually deafened) participated in the study. Prelingual status was defined as having been diagnosed with profound hearing loss at 5 years old or younger. All subjects were recruited from Shanghai Eye, Ear, Nose and Throat Hospital. None were musicians before becoming deaf. Detailed participant demographic information is shown in Table 1. All participants used Cochlear Corp. devices (device type for each participant is listed in Table 1), and all used the same speech-processing strategy (Advanced Combination Encoder, or ACE). The mean age at testing was 10.8 yrs (range: 6 – 16 yrs) for prelingual participants and was 17.1 yrs (range: 9 – 26 yrs) for postlingual participants. The mean duration of deafness (i.e., the interim between diagnosis of profound hearing loss and implantation) was 4 yrs (range: 2 – 12 yrs) for prelingual participants and was 2.1 yrs (range: 1 – 5 yrs) for postlingual participants. The mean age at implantation was 4.3 yrs (range: 2 – 12 yrs) for prelingual participants and was 14.2 yrs (range: 7 – 25 yrs) for postlingual participants. The mean age at deafness (i.e., the age when diagnosed with profound hearing loss) was 0.3 yrs (range: 0 – 2 yrs) for prelingual participants and 12.2 yrs (range: 6 – 22 yrs) for postlingual participants. The mean CI experience was 6.5 yrs (range: 2 – 11 yrs) for prelingual participants and 2.9 yrs (range: 0.3 – 6 yrs) for postlingual participants. One-way analyses of variance (ANOVAs) showed significant difference in age at deafness between prelingual and postlingual participants ($p < 0.001$); there was no significant difference between prelingual and postlingual participants for CI experience or duration of deafness. Participants were tested using their everyday clinical processors and settings, which were not changed during testing. Fifteen CI subjects wore HAs on the contralateral ear during their everyday listening experience. During testing, the HA was turned off, and these subjects were tested using the CI only; note that the HA ear was not plugged during testing. Unaided pure tone average thresholds across audiometric frequencies 125, 250, and 500 Hz for the non-implanted ear are listed in Table 1. Each participant provided written informed consent before participating in the experiment in compliance with the Institutional Review Board protocol of Shanghai Eye, Ear, Nose and Throat Hospital, Fudan University, China.

Lexical tone stimuli—Lexical tone recognition was measured for four tonal patterns: Tone 1 (flat F0), Tone 2 (rising F0), Tone 3 (falling-rising F0), and Tone 4 (falling F0). Stimuli were taken from the Standard Chinese Database (Wang 1993). Four Mandarin Chinese monosyllables (b/a/, b/o/, b/u/, b/i/ in Pinyin) were produced by two males and two females for each of the lexical tones, resulting in a total of 64 stimuli. These vowels were selected because their productions were relatively stable and they were located at the corners of the Chinese single-vowel space (based on the first and second formant frequencies). The original recording were used (i.e., F0, duration and amplitude contour cues were all preserved across stimuli). Stimuli were normalized to have the same long-term root-mean-square amplitude (65 dB SPL). The mean duration across talkers was 273 ms for Tone 1, 340 ms for Tone 2, 410 ms for Tone 3, and 213 ms for Tone 4. Figure 1 shows F0 contours and duration for each lexical tone and for each talker.

MCI stimuli—Melodic contour identification was using stimuli and methods similar to previous studies by Galvin et al. (2007, 2008, 2009a). Stimuli consisted of 9 five-note pitch contours: Rising, Flat, Falling, Flat-Rising, Falling-Rising, Rising-Flat, Falling-Flat, Rising-Falling, and Flat-Falling. A piano sample similar to Galvin et al. (2008) was used as the instrument. Each note was 300 ms in duration. The lowest note in any contour was A3 (220 Hz). The pitch interval between successive notes in the contour was varied to be 1, 2, 3, 4, 5, or 6 semitones. Thus, the MCI stimulus set consisted of 54 melodic contours (9 contours x 6 semitone spacings).

Figure 2 shows electrodograms for four tonal patterns for the syllable /ba/ produced by male talker 1 (left four panels) and for two Rising melodic contours with 1-semitone or 6-semitone spacing (right-most panels). The electrodograms were produced using the clinical default stimulation parameters for the ACE strategy: 900 pulse-per-second (pps) stimulation rate, frequency allocation (input frequency range 188–7938 Hz), 8 spectral maxima, etc. The stimulation patterns for the four lexical tones were quite different. For the melodic contour with 1-semitone spacing, the stimulation pattern is quite similar across the 5 notes for many electrodes (e.g., electrodes 22, 21, 20, 18, and 11). Listeners would have to attend to very subtle changes in the spectral patterns or to changes in F0 encoded in the temporal envelope to perceive this contour. For the 6-semitone melodic contour, the stimulation patterns are quite different across the 5 notes, especially for the most apical electrodes. Listeners may be able to use only changes in the spectral pattern to perceive this contour. Notice also the difference in the time course of the pattern changes between the lexical tones and the melodic contour. For lexical tones, the changes in stimulation pattern and F0 occur over a much shorter time frame (180 to 450 ms), while the changes for the melodic contour occur over 1500 ms.

Test procedures—For all subjects, lexical tone recognition was measured before MCI. Participants were seated in sound-treated booth and listened to stimuli presented in sound field over a single loudspeaker at a comfortably loud level. For tone recognition, a stimulus was randomly selected from the stimulus set (without replacement) and presented to participant who responded by clicking on one of the 4 response boxes, labeled “Tone 1”, “Tone 2”, “Tone 3”, and “Tone 4.” Each participant completed two runs and the recognition scores were averaged across both runs. No preview or feedback was provided. For MCI, participants were instructed to listen to the stimuli one time to familiarize them with the test procedure and with the pitch changes in each contour before formal testing. During testing, a contour would be randomly selected (without replacement) from the stimulus set and presented to the participant, who responded by clicking on one of the nine pictographically labeled response buttons that best matched the stimulus. No trial-by-trial feedback was provided. Detailed description of testing materials and procedures for the MCI task can be found in Galvin et al. (2007).

Results

Figure 3 shows boxplots of tone recognition performance for prelingually and postlingually deafened CI participants, as a function of lexical tone. The distribution of overall lexical tone recognition scores across all participants and lexical tones was significantly different from normal (Kurtosis = 2.4, Skewness = −1.6, $p = 0.02$). Overall lexical tone performance across all participants and lexical tones was 81.0% correct (SE = 2.9). Mean performance was 80.9% (SE = 3.7) for prelingually deafened participants and 81.1% (SE = 5.1) for postlingually deafened participants. One-sample *t* tests showed that mean recognition performance for each of the lexical tones was significantly better than chance level (25% correct; $p < 0.001$ in all cases). A split-plot repeated-measures analysis of variance (RM ANOVA) with lexical tone (Tone 1, 2, 3, or 4) as the within-subject factor and CI status (prelingual or postlingual) as the between-subject factor, showed a significant effect for lexical tone [$F(3,90) = 12.9$, $p < 0.001$], but not for CI status [$F(1,30)= 0.001$, $p = 0.97$]; there was no significant interaction [$F(3,90) = 0.54$, $p = 0.65$]. Post hoc Bonferroni *t* tests revealed that performance with Tone 2 was significantly poorer than with Tones 1 ($p =$ 0.008), 3 ($p = 0.006$), or 4 ($p < 0.001$); performance with Tone 1 was significantly poorer

NIH-PA Author Manuscript
NIH-PA Author Manuscript

than with Tone 4 ($p = 0.037$). There were no significant differences in performance among the remaining tones.

Figure 4 shows boxplots of MCI performance for prelingually and postlingually deafened CI participants, as a function of semitone spacing. The distribution of overall MCI scores across all participants and semitone spacings was significantly different from normal (Kurtosis = −0.2, Skewness = 0.8, $p = 0.002$). Mean MCI performance was 18.5% (SE = 2.5) for prelingual participants and 32.3% (SE = 3.5) for postlingual participants. One-sample *t* tests were used to compare MCI performance to chance-level performance (11.1% correct). For prelingual participants, mean performance for all semitone spacing conditions except for 1 semitone was significantly better than chance level (p = 0.032, 0.020, 0.036, 0.010 and 0.032 for 2, 3, 4, 5, and 6 semitone spacings, respectively). For postlingual participants, mean performance for each of the semitone spacing conditions was significantly better than chance level (p = 0.003, 0.007, 0.009, 0.003, 0.002, and 0.001 for 1, 2, 3, 4, 5, and 6 semitone spacings, respectively). A split-plot RM ANOVA, with semitone spacing (1, 2, 3, 4, 5, or 6) as the within-subject factor and CI status (prelingual or postlingual) as the between-subject factor, showed significant effects for semitone spacing [$F(5,150) = 2.56$, $p = 0.03$] and CI status [$F(1,30) = 10.6$, $p = 0.003$]; there was no significant interaction [$F(5,150) = 0.99$, $p = 0.42$]. Post-hoc Bonferroni *t* tests showed that performance with the 1-semitone spacing was significantly poorer than with the 6-semitone spacing ($p = 0.005$). There were no significant differences among the remaining semitone spacings.

As noted above, MCI performance was significantly better for postlingual than for prelingual participants. Ten out of eleven postlingual participants used a HA in combination with their CI for everyday listening; 5 out of 21 prelingual CI participants also wore a HA for everyday listening. For all participants, the HA was removed but the non-implanted ear was not plugged. It is possible that residual low-frequency hearing may have contributed to performance. Unaided pure-tone average (PTA) thresholds were calculated across audiometric frequencies 125, 250, and 500 Hz for all participants (Table 1). Note that lower frequencies were used for the present PTA than in Wang et al. (2011), who used PTAs calculated across 500, 1000, and 2000 Hz. The lower frequencies were used for the present PTAs because the F0 range was 76 – 299 Hz for the tone stimuli and 220 – 880 Hz for the MCI stimuli. The mean PTA was 96 dB HL for prelingual participants and 80.5 dB HL for postlingual participants. However, across all participants, there was no significant correlation between PTAs and tone recognition ($p = 0.76$) or MCI performance ($p = 0.10$). For prelingual participants, there was no significant correlation between PTAs and tone recognition ($p = 0.63$) or MCI performance ($p = 0.82$). For postlingual participants, there was no significant correlation between PTAs and tone recognition ($p = 0.93$) or MCI performance ($p = 0.67$). While mean PTAs were better for postlingual than for prelingual participants, unaided PTAs were quite high for both groups. It seems unlikely that residual acoustic hearing played a role in the pattern of results.

Figure 5 shows individual participants' overall MCI performance as a function of overall lexical tone recognition performance for prelingually deafened (left panel) and postlingually deafened participants (right panel). There was no significant correlation between tone recognition and MCI performance in prelingual participants ($r^2 = 0.05$, $p = 0.32$). A significant correlation was initially observed in postlingual participants (r^2 = 0.43, p = 0.03; power = 0.60). Note that tone recognition performance for one participant (S28, indicated by the white circle in Fig. 5) was much lower than that for the other postlingual participants. When S28 was removed from the analysis, there was no significant correlation between tone recognition and MCI performance (r^2 = 0.16, p = 0.25); note that after reducing the number of participants in the analysis (n = 10), power was very low (0.21). Thus, it appears that there was no meaningful correlation between tone recognition and MCI performance for postlingual participants.

Demographic factors were compared to lexical tone recognition and MCI performance. As a first approximation, age at deafness (which also represents acoustic hearing experience) was compared to tone and MCI performance; participants were pooled across subject groups for this comparison. Results showed no significant correlation between age at deafness and tone recognition (p = 0.315) or MCI performance (p = 0.285), most likely because there was little variability in prelingual age at deafness, and because pre- and post-lingual deafness may represent different speech pattern development with electric or acoustic hearing. Demographic factors were compared to tone and MCI performance for each subject group independently. The results are shown in Table 2; Bonferroni correction for family-wise error resulted in a significance level of $p < 0.017$. Age at deafness was significantly correlated with MCI performance only for postlingual participants (p = 0.006). CI experience was significantly correlated with MCI performance for prelingual (p = 0.006) and postlingual participants (p = 0.014). Duration of deafness was significantly correlated with tone recognition only for prelingual participants (p = 0.008).

Discussion

In this study, there was no significant difference in tone recognition between the prelingual and postlingual CI participants. This finding remained true even when participant S28 (the poorest performing postlingual participant) was removed from the analysis. The present data do not support the hypothesis that postlingual participants would perform better in tone recognition, given potentially better central pitch patterns developed during previous acoustic hearing. The present results are similar to findings by Wang et al. (2011), but different from Wei et al. (2007) who found better tone recognition for postlingual than for prelingual participants. Lexical tone recognition performance amongst CI participants was highly variable (range: 31 – 97% correct). Tone 2 (rising) was most poorly recognized by the CI participants.

Mean tone recognition (81% correct) in the present study was better than that reported by Wang et al. (2011) or Wang et al. (2012), possibly due to differences in stimuli and subject groups. Unlike the previous studies, duration cues were not controlled across the present tone stimuli. Compared to the 4 vowels used in the present study, previous studies used many more vowel tokens when measuring tone recognition. In this study, tone stimuli were presented in a c/V context, while in Wang et al (2011), some of the tone stimuli included nasal codas (c/V/n or c/V/ng context). Reduced amplitudes at the nasal codas may result in perceptual confusion, especially for falling F0 contours. In the present study, there was a greater number of prelingual participants (n = 21). In Wang et al. (2011), only 5 out of 19

NIH-PA Author Manuscript
NIH-PA Author Manuscript

participants were prelingually deafened; in Wang et al. (2012), all participants were postlingually deafened. The present subjects were also much younger and were implanted at an earlier age than those in the Wang et al. (2011) and Wang et al. (2012). For the present postlingual group, the mean duration of deafness was much shorter than that for the postlingual participants in Wang et al. (2011) and Wang et al. (2012).

In this study, duration of deafness was significantly correlated with tone recognition for prelingual participants, but not for postlingual participants (Table 2). Note that there was generally low power for the tone recognition correlations for postlingual participants, most likely due to the limited number of subjects (n = 11). The mean duration of deafness was nearly twice as long for prelingual (4 yrs) than for postlingual participants (2.1 yrs). This may explain the difference in correlations between subject groups. In Wang et al. (2011), duration of deafness was negatively correlated with tone recognition in postlingual CI participants. Postlingual participants' mean tone recognition was much better in the present study (81% correct) than in Wang et al. (2011: 58% correct), possibly due to reasons described above. It is also possible that prelingual and postlingual participants utilized different listening strategies to perceive F0 cues, with prelingual perceiving patterns developed with electric hearing and postlingual perceiving electric stimulation patterns developed with previous acoustic hearing.

Different from tone recognition, MCI performance was significantly better for postlingual (mean = 32% correct) than for prelingual participants (18% correct). While 10 of the 11 postlingual participants wore HAs in combination with their CIs for everyday listening, they were tested with the HA removed. Although the non-implanted ear was not plugged, residual acoustic hearing did not appear to play a role in MCI performance, as there was no correlation between PTAs (125, 250, and 500 Hz) and MCI performance within or across subject groups. The results are in agreement with the hypothesis that postlingual participants would outperform prelingual participants in the MCI task, possibly due to more robust pitch representations developed during previous acoustic hearing. In the present study, duration of previous acoustic hearing for postlingual participants (mean = 12.2 yrs) was significantly longer than prelingual participants (mean = 0.3 yrs). Interestingly, the largest differences in MCI performance between prelingual and postlingual participants were for the larger semitone spacings (Fig. 4). With six-semitone spacing, mean postlingual MCI performance was 21.9 percentage points better than prelingual performance. Also, postlingual performance generally improved as the semitone spacing increased, suggesting that these participants were better able to hear melodic pitch cues as they became available. Mean prelingual performance was generally unchanged as the semitone spacing increased, suggesting they were unable to perceive even large changes in pitch across the contours. It is possible that prelingual participants needed even larger semitone spacings to perceive the different contours. As shown in the electrodograms in the right panel of Figure 2, with six-semitone spacing, the stimulation pattern upwardly shifts by 1 electrode at the 3^{rd} and 5^{th} notes of the contour. Wang et al. (2011) found a mean threshold of 5.7 semitones for a pitch interval adjustment task using familiar melodies. Wang et al. (2012) found that the mean pitch discrimination threshold (across the piano and flute instruments) was 4.3 semitones. The pitch range for the present 6-semitone rising contour was a maximum of 24 semitones. Thus, while prelingual CI users may have had sufficient resolution to discriminate somewhat large differences in pitch, they did not seem to be able to use this pitch information to identify the melodic contours.

MCI performance was positively correlated with CI experience for both subject groups. This is consistent with Leal et al. (2003) and Gfeller et al. (2007) who found that pitch ranking was significantly correlated with CI experience and in agreement with our hypothesis. Interestingly, MCI performance was negatively correlated with postlingual participants' age at deafness ($r = -0.77$, $p = 0.006$). This finding was not in agreement with our hypothesis. The mean duration of deafness for postlingual CI participants was 2.1 yrs, (range = 1 – 5 yrs), while the mean age at deafness was 12.2 yrs. (range = 6 – 22 yrs). However there was no significant correlation between age at deafness and duration of deafness ($r = 0.27$, $p = 0.42$). The negative correlation between age at deafness and MCI performance suggests that additional acoustic hearing experience (more than 6 years) did not benefit postlingual participants. It could be that the postlingual participants had adequate acoustic hearing experience to develop central pitch representations, but were differently able to use these representations when listening to melodic pitch with electric hearing. Other factors that may explain the negative correlation were not captured in the present study (e.g., everyday music listening habits, the quality of acoustic hearing before implantation, etc.).

There was no significant correlation between tone recognition and MCI performance for prelingual participants, or for postlingual participants after excluding a very poor performer (S28) from the analysis. Note that power was below the desired level of 0.80 for all correlations, whether due to great variability in performance and/or to the limited number of participants, especially for postlingual participants. Thus, the present results are somewhat in contrast to those of Wang et al. (2011) or Wang et al. (2012), who found significant correlations between lexical tone recognition and music perception. As noted above, the present music perception test may have been more challenging than in the previous studies. Also, in Wang et al. (2011) and Wang et al. (2012), duration cues were fixed for the tone stimuli. In this study, natural utterances were used, in which amplitude and duration cues that co-vary with F0 were preserved. It could be that when all speech cues are available, the relationship between speech and music perception is less clear. It is possible that the relationship between lexical tone recognition and music perception would emerge if the temporal cues of Mandarin tones were controlled. Another factor to consider is the novel listening experience with the MCI stimuli for the present CI participants. MCI training has been shown to greatly improve MCI performance in postlingual adult English-speaking CI users, with mean post-training performance improving by 28.3 percent points in Galvin et al. (2007) and 10.3 percentage points in Galvin et al. (2012). More recently, Fu et al. (2014) reported large gains in MCI performance in pediatric, prelingually deafened Mandarin-speaking CI users after a moderate amount of home training. In that study, mean baseline performance MCI performance with the piano was 32.8% correct, somewhat higher than the present prelingual performance. After 4 weeks of home training with different stimuli, mean MCI performance with the piano improved by 46 percentage points. Such training may also have improved MCI performance in the present participants. If so, a relationship between MCI and tone recognition may yet emerge after training.

Conclusions

We evaluated melodic contour identification (MCI) and lexical tone perception in 32 young Mandarin-speaking CI users. Major findings include:

1. MCI performance was significantly better for postlingual than for prelingual CI participants, suggesting that previous acoustic hearing experience may benefit CI users' melodic pitch perception.
2. Consistent with previous studies, the present Chinese CI listeners had great difficulty perceiving melodic pitch, and there was great inter-subject variability in performance.
3. For lexical tone recognition, there was no significant difference between prelingual and postlingual CI participants.
4. There was no significant correlation between lexical tone recognition and MCI performance in prelingual participants, or for postlingual participants after adjusting for outlier performance.
5. Age at deafness was significantly correlated with MCI performance in postlingual participants. CI experience was significantly correlated with MCI performance for both prelingual and postlingual participants, suggesting that melodic pitch perception may improve with CI experience. Duration of deafness was significantly correlated with tone recognition only for prelingual participants, perhaps due to a much longer period of auditory deprivation than experienced by postlingual participants.

Acknowledgments

We are grateful to the participants who participated in this research. This work was funded in part by the Special Fund for Projects of Ministry of Health of China (#201202005), the National Nature Science Foundation of China (#81371087), Shanghai Science and Technology Commission Medical Guide Project (#134119a1800), and by NIH grant R01-DC004993.

This work was funded in part by the Special Fund for Projects of Ministry of Health of China (#201202005), the National Nature Science Foundation of China (#81371087), Shanghai Science and Technology Commission Medical Guide Project (#134119a1800), and by NIH grant R01-DC004993.

References

Abramson AS. Static and dynamic acoustic cues in distinctive tones. Lang Speech. 1978; 21:319–325. [PubMed: 750791]

Chatterjee M, Peng SC. Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition. Hear Res. 2008; 235:143–156. [PubMed: 18093766]

Chatterjee M, Ozerbut C. Detection and rate discrimination of amplitude modulation in electrical hearing. J Acoust Soc Am. 2011; 130:1567–1580. [PubMed: 21895095]

Cooper WB, Tobey E, Loizou PC. Music perception by cochlear implant and normal hearing listeners as measured by the Montreal Battery for Evaluation of Amusia. Ear Hear. 2008; 29:618–626. [PubMed: 18469714]

Fraser M, McKay CM. Temporal modulation transfer functions in cochlear implantees using a method that limits overall loudness cues. Hear Res. 2012; 283:59–69. [PubMed: 22146425]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Friesen LM, Shannon RV, Baskent D, et al. Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. J Acoust Soc Am. 2001; 110:1150–1163. [PubMed: 11519582]

Fu QJ, Hsu CJ, Horng MJ. Effects of speech processing strategy on Chinese tone recognition by Nucleus-24 cochlear implant users. Ear Hear. 2004; 25:501–508. [PubMed: 15599196]

Fu QJ, Shannon RV. Effects of amplitude nonlinearity on phoneme recognition by cochlear implant users and normal-hearing listeners. J Acoust Soc Am. 1998; 104:2570–2577. [PubMed: 9821336]

Fu QJ, Zeng FG, Shannon RV, et al. Importance of tonal envelope cues in Chinese speech recognition. J Acoust Soc Am. 1998; 104:505–510. [PubMed: 9670541]

Fu QJ, Wu JL, Wang X, Galvin JJ 3rd. Benefits of music training in Mandarin-speaking pediatric cochlear implant users. J Speech Lang Hear Res. 2014 Apr. submitted.

Fujita S, Ito J. Ability of nucleus cochlear implantees to recognize music. Ann Otol Rhinol Laryngol. 1999; 108:634–640. [PubMed: 10435919]

Galvin JJ 3rd, Fu QJ, Nogaki G. Melodic contour identification by cochlear implant listeners. Ear Hear. 2007; 28:302–319. [PubMed: 17485980]

Galvin JJ 3rd, Fu QJ, Oba SI. Effect of instrument timbre on melodic contour identification by cochlear implant users. J Acoust Soc Am. 2008; 124:EL189–EL195. [PubMed: 19062785]

Galvin JJ 3rd, Fu QJ, Oba SI. Effect of a competing instrument on melodic contour identification by cochlear implant users. J Acoust Soc Am. 2009a; 125:EL98–EL103. [PubMed: 19275282]

Galvin JJ 3rd, Fu QJ, Shannon RV. Melodic contour identification and music perception by cochlear implant users. Ann N Y Acad Sci. 2009b; 1169:518–533. [PubMed: 19673835]

Galvin JJ, Eskridge E, Oba SI, Fu QJ. Melodic contour identification training in cochlear implant users with and without a competing instrument. Semin Hear. 2012; 33:399–409.

Geers AE, Hayes H. Reading, writing, and phonological processing skills of adolescents with 10 or more years of cochlear implant experience. Ear Hear. 2011; 32:49S–59S. [PubMed: 21258612]

Gfeller K, Christ A, Knutson JF, et al. Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients. J Am Acad Audiol. 2000; 11:390–406. [PubMed: 10976500]

Gfeller K, Turner C, Oleson J, Zhang X, Gantz B, et al. Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise. Ear Hear. 2007; 28:412–423. [PubMed: 17485990]

Green T, Faulkner A, Rosen S. Enhancing temporal cues to voice pitch in continuous interleaved sampling cochlear implants. J Acoust Soc Am. 2004; 116:2298–2310. [PubMed: 15532661]

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]

Kang R, Nimmons GL, Drennan W, et al. Development and validation of the University of Washington Clinical Assessment of Music Perception test. Ear Hear. 2009; 30:411–418. [PubMed: 19474735]

Kong YY, Cruz R, Jones JA, et al. Music perception with temporal cues in acoustic and electric hearing. Ear Hear. 2004; 25:173–185. [PubMed: 15064662]

Kong YY, Stickney GS, Zeng FG. Speech and melody recognition in binaurally combined acoustic and electric hearing. J Acoust Soc Am. 2005; 117:1351–1361. [PubMed: 15807023]

Lassaletta L, Castro A, Bastarrica M, et al. Does music perception have an impact on quality of life following cochlear implantation? Acta Otolaryngol. 2007; 127:682–686. [PubMed: 17573562]

Leal MC, Shin YJ, Laborde ML, Calmels MN, Verges S, et al. Music perception in adult cochlear implant recipients. Acta Otolaryngol. 2003; 123:826–835. [PubMed: 14575398]

Looi V, McDermott H, McKay C, et al. Music perception of cochlear implant users compared with that of hearing aid users. Ear Hear. 2008; 29:421–434. [PubMed: 18344870]

Luo X, Fu QJ, Wei CG, et al. Speech recognition and temporal amplitude modulation processing by Mandarin-speaking cochlear implant users. Ear Hear. 2008; 29:957–970. [PubMed: 18818548]

Luo X, Fu QJ, Wu HP, et al. Concurrent-vowel and tone recognition by Mandarin-speaking cochlear implant users. Hear Res. 2009; 256:75–84. [PubMed: 19595753]

Mirza S, Douglas SA, Lindsey P, et al. Appreciation of music in adult patients with cochlear implants: a patient questionnaire. Cochlear Implants Int. 2003; 4:85–95. [PubMed: 18792140]

Miyamoto RT, Kirk KI, Robbins AM, et al. Speech perception and speech production skills of children with multichannel cochlear implants. Acta oto-laryngologica. 1996; 116:240–243. [PubMed: 8725523]

Morton KD, Torrione PA Jr, Throckmorton CS, et al. Mandarin Chinese tone identification in cochlear implants: predictions from acoustic models. Hear Res. 2008; 244:66–76. [PubMed: 18706497]

Niparko JK. Speech, language, and reading skills after early cochlear implantation. JAMA. 2004; 291:2378–2380. [PubMed: 15151121]

Peng SC, Tomblin JB, Cheung H, et al. Perception and production of mandarin tones in prelingually deaf children with cochlear implants. Ear Hear. 2004; 25:251–264. [PubMed: 15179116]

Peng SC, Lu N, Chatterjee M. Effects of cooperating and conflicting cues on speech intonation recognition by cochlear implant users and normal hearing listeners. Audiol Neurootol. 2009; 14:327–337. [PubMed: 19372651]

Peng SC, Chatterjee M, Lu N. Acoustic cue integration in speech intonation recognition with cochlear implants. Trends Amplif. 2012; 16:67–82. [PubMed: 22790392]

Shannon RV, Fu QJ, Galvin J 3rd. The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. Acta Otolaryngol Suppl. 2004:50–54. [PubMed: 15219048]

Smith ZM, Delgutte B, Oxenham AJ. Chimaeric sounds reveal dichotomies in auditory perception. Nature. 2002; 416:87–90. [PubMed: 11882898]

Sucher CM, McDermott HJ. Pitch ranking of complex tones by normally hearing subjects and cochlear implant users. Hear Res. 2007; 230:80–87. [PubMed: 17604582]

Vandali AE, Sucher C, Tsang DJ, et al. Pitch ranking ability of cochlear implant recipients: A comparison of sound-processing strategies. The Journal of the Acoustical Society of America. 2005; 117:3126. [PubMed: 15957780]

Vongphoe M, Zeng FG. Speaker recognition with temporal cues in acoustic and electric hearing. J Acoust Soc Am. 2005; 118:1055–1061. [PubMed: 16158660]

Wang RH. The standard Chinese database. University of Science of Technology of China, internal materials. 1993

Wang W, Zhou N, Xu L. Musical pitch and lexical tone perception with cochlear implants. Int J Audiol. 2011; 50:270–278. [PubMed: 21190394]

Wang S, Liu B, Dong R, et al. Music and lexical tone perception in Chinese adult cochlear implant users. Laryngoscope. 2012; 122:1353–1360. [PubMed: 22362607]

Wei C, Cao K, Jin X, et al. Psychophysical performance and Mandarin tone recognition in noise by cochlear implant users. Ear Hear. 2007; 28:62S–65S. [PubMed: 17496650]

Wilson BS, Finley CC, Lawson DT, et al. Better speech recognition with cochlear implants. Nature. 1991; 352:236–238. [PubMed: 1857418]

Xin L, Fu QJ, Galvin JJ 3rd. Vocal emotion recognition by normal-hearing listeners and cochlear implant users. Trends Amplif. 2007; 11:301–315. [PubMed: 18003871]

Zhu M, Chen B, Galvin JJ 3rd, et al. Influence of pitch, timbre and timing cues on melodic contour identification with a competing masker (L). J Acoust Soc Am. 2011; 130:3562–3565. [PubMed: 22225012]

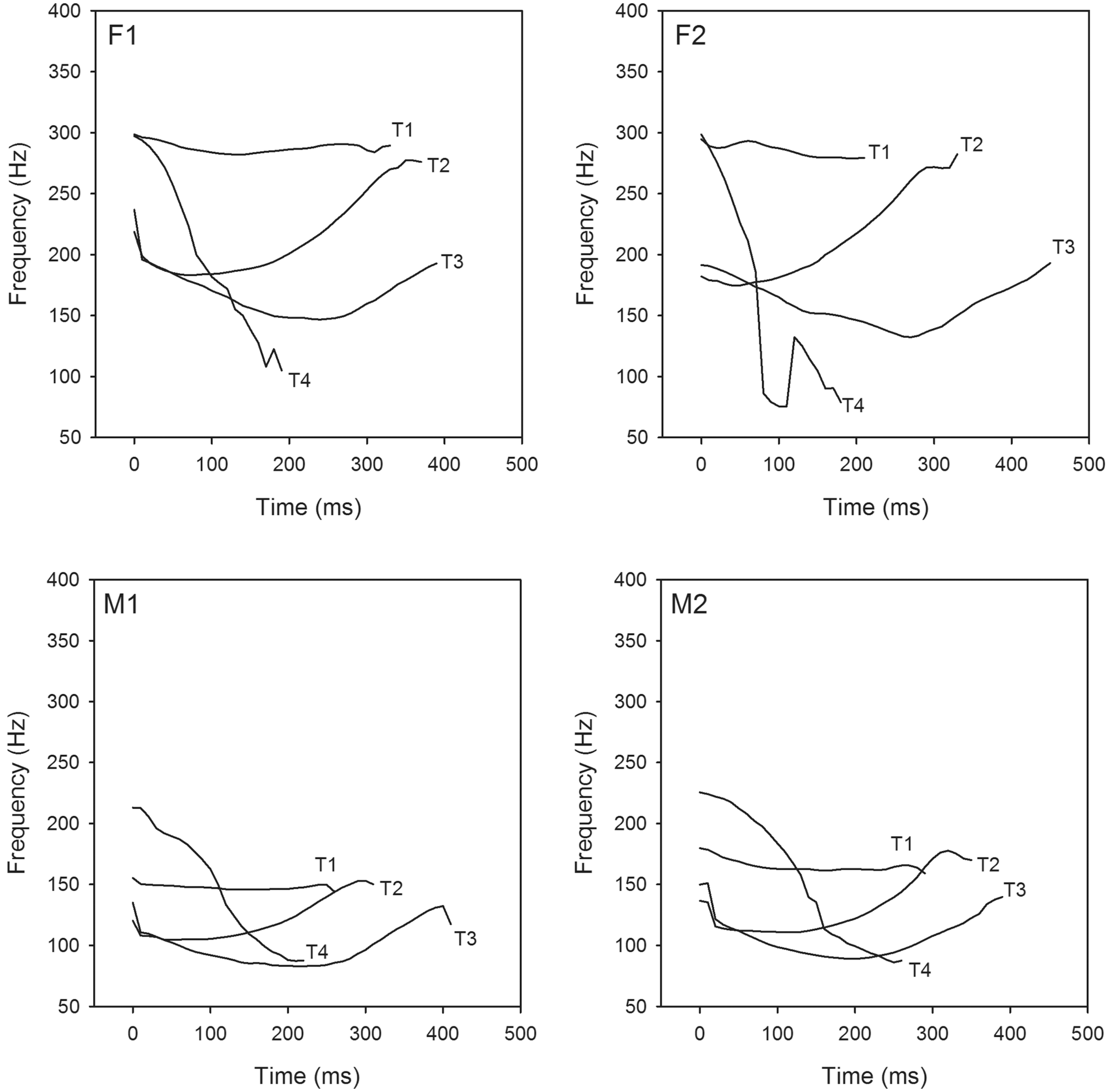

Figure 1.

F0 contours for 4 lexical tones, for each talker. The x-axis shows time, and the y-axis shows frequency. Each lexical tone is labeled within each plot. T1 = Tone 1 (flat); T2 = Tone 2 (rising); T3 = Tone 3 (falling-rising); T4 = Tone 4 (falling).

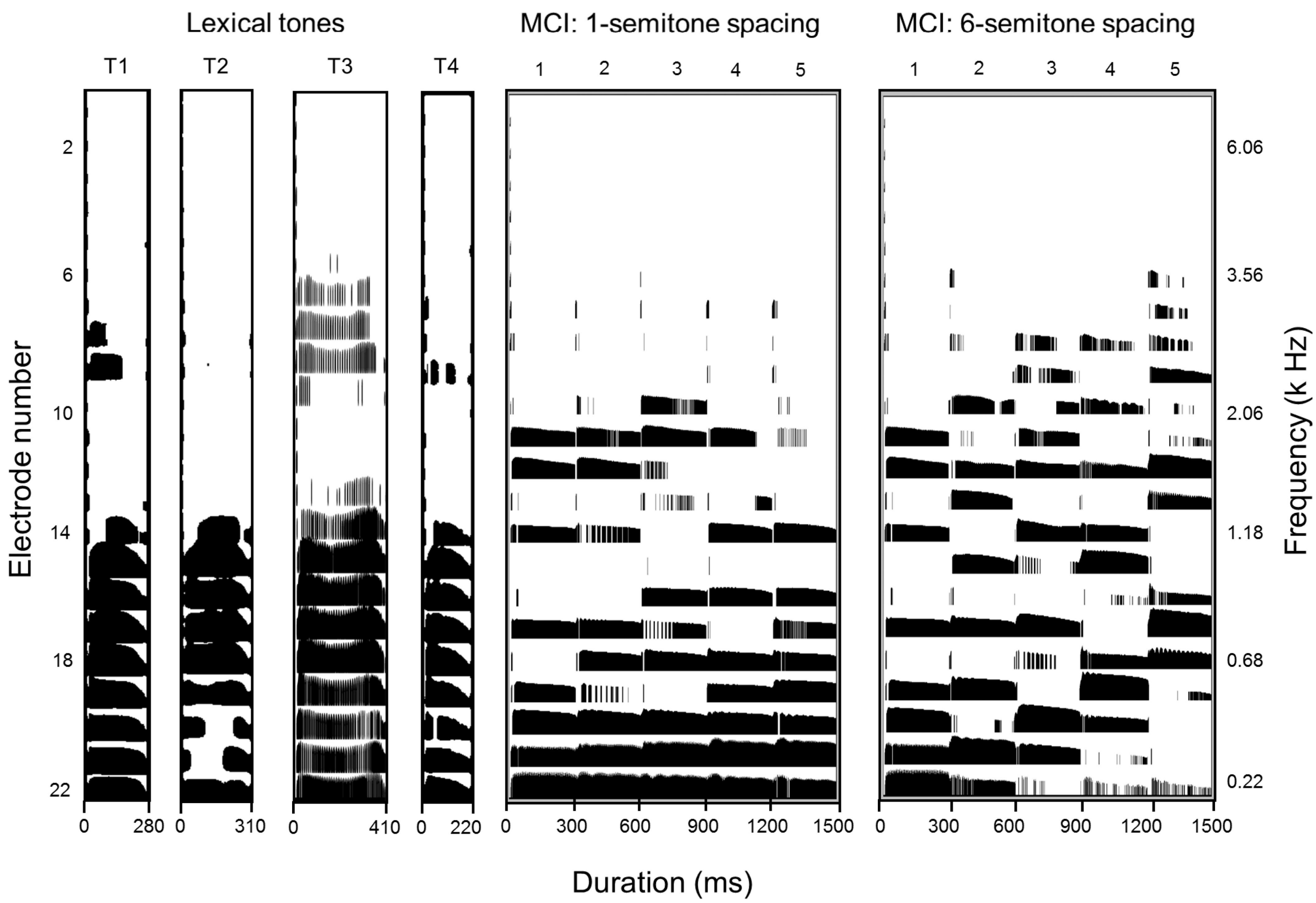

Figure 2.

Electrodograms for 4 lexical tones (left four panels) and 2 melodic contours (right two panels). The x-axis shows time, the left y-axis shows electrode number (from most apical electrode 22 to most basal electrode 1) and the right y-axis shows the acoustic frequency assigned to the electrodes. For the melodic contours, the numbers 1–5 at the top of the panel indicate each note in the contour. The electrodograms were generated using the default clinical stimulation parameters for the Cochlear Freedom CI device.

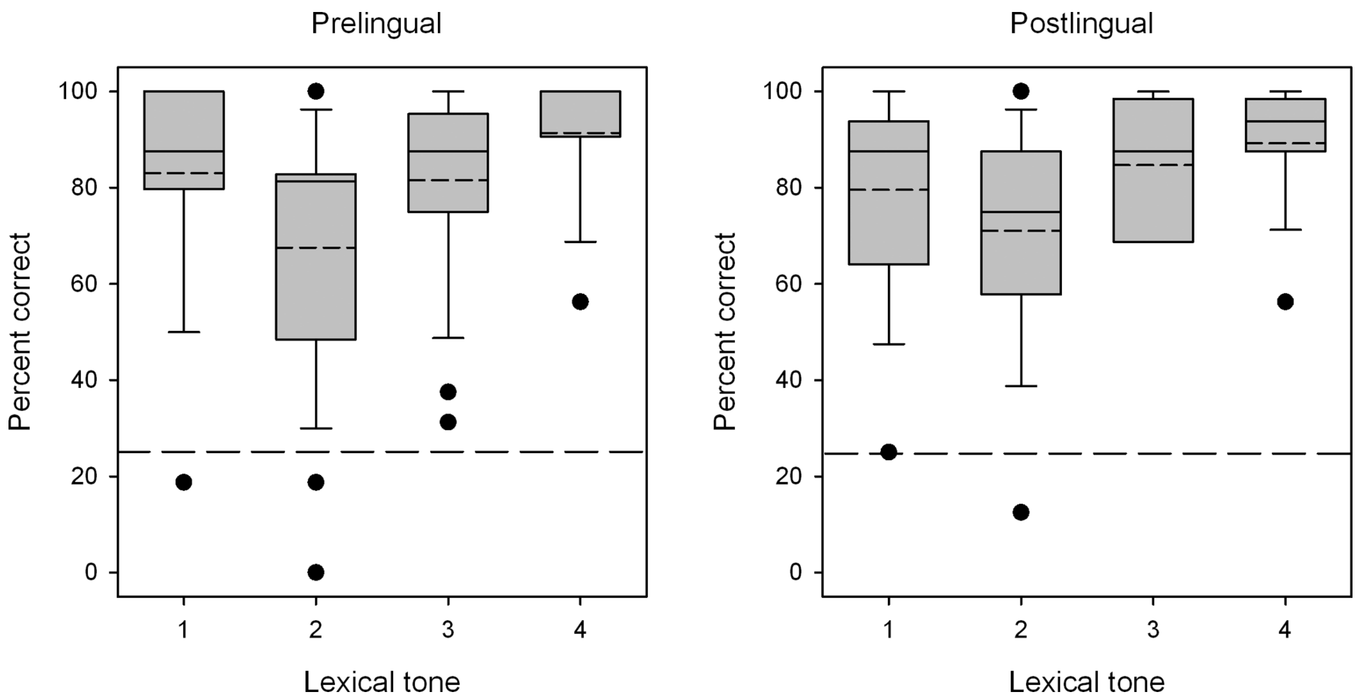

Figure 3.

Boxplots for lexical tone recognition performance for prelingually deafened (left panel) and postlingually deafened CI participants (right panel). The short dashed lines show median value, the solid lines show mean value, the error bars show the 10th and 90th percentiles, and the circles show outliers. The long dashed line indicates chance level performance (25.0% correct).

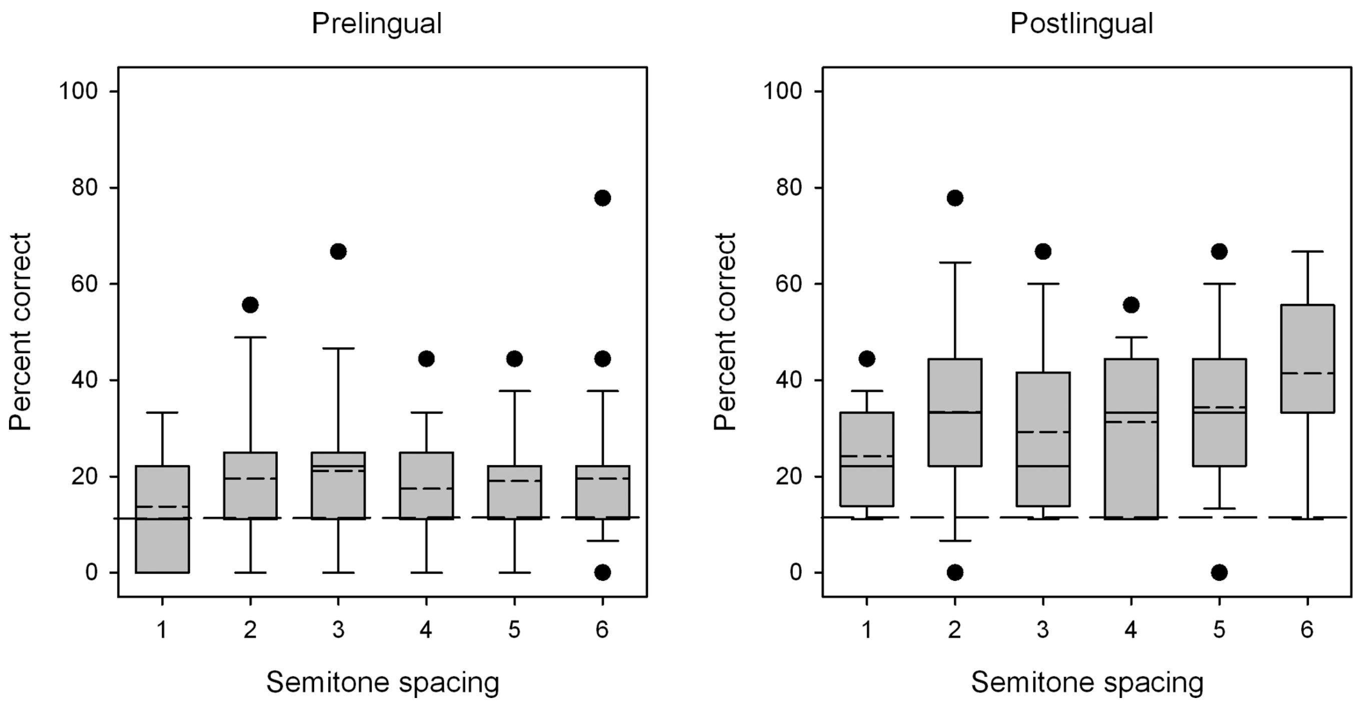

Figure 4.
Boxplots for MCI performance for prelingually deafened (left panel) and postlingually deafened CI participants (right panel). The short dashed lines show median value, the solid lines show mean value, the error bars show the 10th and 90th percentiles, and the circles show outliers. The long dashed line indicates chance level performance (11.1% correct).

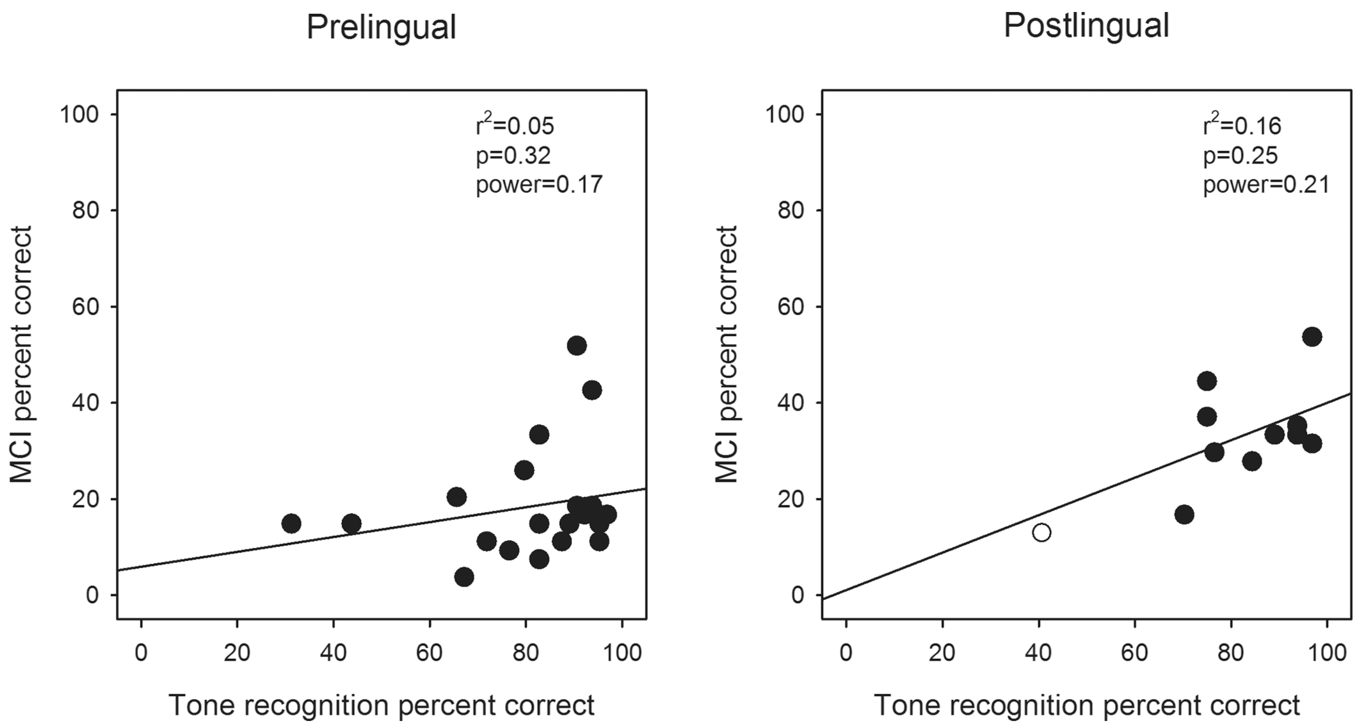

Figure 5.
MCI performance for individual subjects as a function of lexical tone recognition performance, for prelingually deafened (left panel) and postlingually deafened CI participants (right panel). For lexical tone data, performance was averaged across talkers and tones; for MCI data, performance was averaged across semitone spacings. The dashed lines show linear regressions fit to the data; r^2, *p*, and power values are shown near the regression lines. For postlingual participants, the regression analysis shown was performed excluding participant S28 (white circle), as tone recognition performance was much poorer than for the other participants.

Table 1

Demographic information for CI participants.

Participant	Sex	Age test (yrs)	CI Exp (yrs)	Age imp (yrs)	Age deaf (yrs)	Dur deaf (yrs)	HA use	Pre/ post	Device	Unaided PTA (non-implanted ear)
S1	F	15.2	11.2	4.0	0.0	4.0	No	Pre	CI24M	105
S2	F	10.3	8.3	2.0	0.0	2.0	No	Pre	CI24M	105
S3	F	6.0	3.9	2.1	0.0	2.1	No	Pre	CI24R	90
S4	M	19.1	7.2	11.9	0.1	11.8	Yes	Pre	CI24M	78
S5	M	11.2	8.2	3.0	1.0	2.0	No	Pre	CI24M	105
S6	F	10.6	5.2	5.4	0.2	5.2	No	Pre	CI24R	83
S7	F	8.2	3.9	4.3	0.2	4.1	No	Pre	CI24R	83
S8	M	13.2	4.1	9.1	0.9	8.2	Yes	Pre	CI24R	77
S9	F	8.9	4.7	4.2	0.0	4.2	Yes	Pre	CI24M	80
S10	M	8.6	4.7	3.9	1.8	2.1	No	Pre	CI24M	105
S11	F	6.7	2.0	4.7	0.0	4.7	Yes	Pre	CI24RE	73
S12	M	13.2	10.9	2.3	0.0	2.3	No	Pre	CI24M	103
S13	M	9.0	5.0	4.0	0.0	4.0	No	Pre	CI24M	103
S14	F	10.7	8.8	1.9	0.1	1.8	No	Pre	CI24M	105
S15	F	12.8	11.1	1.7	0.0	1.7	No	Pre	CI24M	102
S16	M	15.9	3.8	12.1	0.4	11.7	No	Pre	CI24R	105
S17	M	11.0	9.2	1.8	0.0	1.8	No	Pre	CI24M	105
S18	M	10.2	8.0	2.2	0.00	2.2	No	Pre	CI24M	105
S19	F	9.9	5.7	4.2	1.1	3.1	No	Pre	CI24M	105
S20	F	8.1	5.2	2.9	0.0	2.9	No	Pre	CI24M	103
S21	M	8.3	5.1	3.2	0.0	3.2	Yes	Pre	CI24M	93
S22	M	11.9	4.5	7.4	5.7	1.7	Yes	Post	CI24M	78
S23	F	18.3	0.9	17.4	16.3	1.1	Yes	Post	CI24RE	77
S24	M	12.8	2.9	9.9	8.8	1.1	Yes	Post	CI24R	105
S25	M	15.7	4.7	11.0	5.8	5.2	Yes	Post	CI24M	78
S26	M	23.1	0.8	22.3	20.1	2.2	Yes	Post	CI24RE	75
S27	M	23.4	0.3	23.1	19.9	3.2	Yes	Post	CI24RE	82

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Participant	Sex	Age test (yrs)	CI Exp (yrs)	Age imp (yrs)	Age deaf (yrs)	Dur deaf (yrs)	HA use	Pre/ post	Device	Unaided PTA (non-implanted ear)
S28	M	25.8	0.9	24.9	21.8	3.1	No	Post	CI24RE	83
S29	M	9.3	2.1	7.2	6.2	1.0	Yes	Post	CI24R	72
S30	F	10.8	3.7	7.1	5.9	1.2	Yes	Post	CI24M	73
S31	F	23.2	2.9	20.3	17.2	3.1	Yes	Post	CI24R	82
S32	F	13.9	5.8	8.1	6.1	2.0	Yes	Post	CI24M	80

Age test = age at testing; CI exp = cochlear implant experience; Age implant = age at implantation; Age deaf = age at profound deafness; Dur deaf = duration of deafness; Pre/post = prelingually or postlingually deafened; PTA = pure-tone average threshold in dB HL calculated across 125, 250, and 500 Hz,

Table 2

Regression analyses between lexical tone recognition or MCI performance and demographic variables.

	Prelingual						Postlingual					
	Tone			MCI			Tone			MCI		
	r^2	p	power	r^2	p	power	r^2	p	power	r^2	p	power
Age deaf	0.04	0.37	0.14	0.18	0.06	0.49	0.32	0.07	0.45	0.59	0.006*	0.82
CI exp	0.05	0.33	0.16	0.34	0.006*	0.81	0.10	0.35	0.15	0.50	0.014*	0.71
Dur deaf	0.32	0.008*	0.77	0.01	0.61	0.07	0.08	0.39	0.13	0.09	0.37	0.14

Age deaf = age at profound deafness; CI exp = cochlear implant experience; Dur deaf = duration of deafness.

* $p < 0.017$ (Bonferroni corrected)