Published in final edited form as:

Ear Hear. 2010 December ; 31(6): 796–805. doi:10.1097/AUD.0b013e3181e8b7bd.

Psychoacoustic abilities associated with music perception in cochlear implant users

Jong Ho Won, MS^{1,2}, Ward R. Drennan, PhD¹, Robert S. Kang, MD, MPH¹, and Jay T. Rubinstein, MD, PhD^{1,2}

¹Virginia Merrill Bloedel Hearing Research Center, Department of Otolaryngology-Head and Neck Surgery, University of Washington, Seattle, WA 98195, USA

²Department of Bioengineering, University of Washington, Seattle, WA 98195, USA

Abstract

Objectives—This study was designed to determine what acoustic elements are associated with musical perception ability in cochlear implant (CI) users; and to understand how acoustic elements, which are important to good speech perception, contribute to music perception in CI users. It was hypothesized that the variability in the performance of music and speech perception may be related to differences in the sensitivity to specific acoustic features such as spectral changes or temporal modulations, or both.

Design—A battery of hearing tasks was administered to forty-two CI listeners. The Clinical Assessment of Music Perception (CAMP) was used, which evaluates complex-tone pitch-direction discrimination, melody recognition, and timbre recognition. To investigate spectral and temporal processing, spectral-ripple discrimination and Schroeder-phase discrimination abilities were evaluated. Speech perception ability in quiet and noise was also evaluated. Relationships between CAMP subtest scores, spectral-ripple discrimination thresholds, Schroeder-phase discrimination scores, and speech recognition scores were assessed.

Results—Spectral-ripple discrimination was shown to correlate with all three aspects of music perception studied. Schroeder-phase discrimination was generally not predictive of music perception outcomes. Music perception ability was significantly correlated with speech perception ability. Nearly half of the variance in melody and timbre recognition was predicted jointly by spectral-ripple and pitch-direction discrimination thresholds. Similar results were observed on speech recognition as well.

Conclusions—The current study suggests that spectral-ripple discrimination is significantly associated with music perception in CI users. A previous report showed that spectral-ripple discrimination is significantly correlated with speech recognition in quiet and in noise (Won *et al.*, 2007). The present study also showed that speech recognition and music perception are also related to one another. Spectral-ripple discrimination ability seems to reflect a wide range of hearing abilities in CI users. The results suggest that materially improving spectral resolution could provide significant benefits in music and speech perception outcomes in CI users.

Correspondence to: Jong Ho Won, jhwon@uw.edu, Telephone: (206) 616-2041; Fax: (206) 616-1828.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

cochlear implants; psychophysical abilities; music perception; speech perception

I. INTRODUCTION

Speech perception performance by cochlear implant (CI) users has steadily improved (e.g. Wilson and Dorman, 2008). One reason for this success has been the improvement in understanding of the perceptual processes involved in electric hearing and their relation to speech perception outcome. Numerous psychophysical tests have been conducted; for example, (1) temporal modulation detection (Cazals *et al.*, 1994; Fu, 2002; Luo *et al.*, 2008); (2) electrode discrimination (Nelson *et al.*, 1995; Zwolan *et al.*, 1997; Donaldson and Nelson, 2000; Henry *et al.*, 2000); (3) pitch perception (Collins *et al.*, 1997; Gfeller *et al.*, 2007); (4) spectral-ripple discrimination (Henry and Turner, 2003; Henry *et al.*, 2005; Won *et al.*, 2007; Litvak *et al.*, 2007; Saoji *et al.*, 2009); (5) gap detection (Wei *et al.*, 2007); (6) auditory stream segregation (Hong and Turner, 2006); and (7) Schroeder-phase discrimination (Drennan *et al.*, 2008). These studies show various factors that predict the variability in speech outcomes in CI users.

For example, previous work (Won *et al.*, 2007) has demonstrated that spectral-ripple discrimination ability significantly correlated with speech reception thresholds (SRTs) in noise in CI users ($r = -0.55$, $p = 0.002$ in two-talker babble; $r = -0.62$, $p = 0.0004$ in steady-state noise). Other investigators also reported that spectral shape measures correlate with speech recognition (Henry and Turner, 2003; Henry *et al.*, 2005; Litvak *et al.*, 2007; Saoji *et al.*, 2009), suggesting that spectral selectivity with a sound processor is a fundamental underlying factor that partly accounts for the wide range of variability in speech perception in CI users. More recently, Drennan *et al.* (2008) showed that in CI users, Schroeder-phase discrimination was significantly correlated with monosyllabic word recognition in quiet ($r = 0.52$, $p = 0.03$) and speech reception thresholds in steady-state noise ($r = -0.48$, $p = 0.03$). Further, Schroeder-phase discrimination ability did not show any relationship with spectral-ripple discrimination ability, thus it was concluded that Schroeder-phase discrimination serves as a measure of temporal resolution in CI users, assessing CI users' ability to use across-channel timing differences in the temporal envelopes as well as within-channel temporal fine structure (Won *et al.* 2009). If the sound processor could provide improved representation of temporal and spectral details of sound, it is expected that speech understanding abilities would improve (e.g. Drennan *et al.* 2010).

Previous studies have shown that the performance of rhythm recognition in CI users is generally good (Gfeller *et al.*, 1997; Kong *et al.*, 2004; Gfeller *et al.*, 2005), but other aspects of music perception such as pitch (e.g., Gfeller *et al.*, 2007), melody (e.g. Gfeller *et al.*, 2002a), and timbre (e.g. Gfeller *et al.*, 2002b) are still generally weak in a majority of CI users. Considerable across-subject variability in music perception ability has also been found (e.g. Kong *et al.*, 2004; Galvin *et al.*, 2007; Nimmons *et al.*, 2008; Kang *et al.*, 2009); however, it is not clear why some CI users show especially good music perception relative to the population of implant users. Although the relationship between various psychophysical measures and speech perception in CI users has been extensively studied, the relationship between psychophysical measures and music perception has not. In order to develop CI technology for better music perception outcomes, it is important to understand which acoustic characteristics are most important for music perception.

The goal of the current study is to determine what psychoacoustic abilities are associated with musical perception ability. It was hypothesized that the variability in music perception

performance is related to differences in the psychophysical abilities relevant to speech and music perception in CI users. To test the hypotheses, a hearing test battery was implemented which included (1) a Clinical Assessment of Music Perception (CAMP) test (Nimmons *et al.*, 2008; Kang *et al.*, 2009), (2) a measure of resolving broad-band spectral differences (Henry *et al.*, 2005, Won *et al.*, 2007), (3) a measure of ability to track temporal modulations that sweep rapidly across channels (Schroeder-phase discrimination, Drennan *et al.*, 2008), (4) speech recognition in quiet using consonant–nucleus–consonant (CNC) words (Peterson and Lehiste, 1962), and (5) speech reception threshold (SRT) in noise using spondees in babble noise and steady-state, speech-shaped noise (Turner *et al.*, 2004; Won *et al.*, 2007). All of the psychophysical tests have been shown to be reliable measures, showing good test-retest reliability with minimal learning. For example, Kang *et al.* (2009) showed in 35 CI users that the test-retest intraclass correlation coefficient for pitch, melody, and timbre tests of the CAMP test were 0.85, 0.92, and 0.69, respectively. Won *et al.* (2007) reported strong test-retest reliability for the spectral-ripple test ($r = 0.89$). Drennan *et al.* (2008) demonstrated that test-retest reliability for the Schroeder-phase discrimination was good, showing no improvements throughout the six test blocks as well as no significant difference between the scores measured on two different days.

In the present study, all tests were performed with the subjects' own sound processors; therefore, discrimination sensitivity for real-world acoustic sounds was evaluated as well as speech and music outcomes and relationships among measures. Examining a relationship between the components of music perception and psychophysical abilities will enhance past research (e.g. Gfeller and Lansing, 1991) by providing a better understanding of how specific acoustic elements contribute to music perception, and might provide insights into improving music perception with CIs.

II. METHODS

1. Subjects

Forty-two post-lingually deafened adult CI users participated in this study. They were 25 – 78 years old (mean = 50 years, 20 females and 22 males), and all were native speakers of American English. Individual subject information is listed in Table 1. The use of human subjects in this study was reviewed and approved by the University of Washington Institutional Review Board. A few subjects did not participate in all of the experiments due to time and scheduling constraints. For the correlation or regression analyses, all the available data have been included without ignoring any data point.

2. Test administration

All subjects listened to the stimuli using their own sound processor set to a comfortable listening level. Bilateral users were tested with both implants functioning. CI sound processor settings were not changed from the clinical settings throughout all test batteries. All tests were conducted in a double-walled, sound-treated booth (IAC). Custom MATLAB (The Mathworks, Inc.) programs were used to present stimuli on a Macintosh G5 computer with a Crown D45 amplifier. A single loudspeaker (B&W DM303), positioned 1-m in front of the subjects, presented stimuli in a sound-field. Five psychophysical test batteries were administered in this study. The order of test administration was randomized across the subjects. The duration of the entire test battery for one subject was approximately two and a half hours.

A. CAMP (Clinical Assessment of Music Perception) test—The CAMP was used for music perception testing as previously described by Nimmons *et al.* (2008) and Kang *et al.* (2009). All stimuli in this task were presented at 65 dBA. All of the forty-two CI subjects were

tested for music perception. The CAMP has three subtests; complex-tone pitch-direction discrimination, melody recognition, and timbre recognition.

Synthesized complex tones were created for the complex-tone pitch-direction discrimination and melody recognition tests. The synthetic complex tones with uniform temporal envelopes had identical spectral envelopes derived from a recorded piano note at middle C (262 Hz). A custom peak-detection algorithm was used to determine the fast Fourier transform components of the fundamental frequency (F0) of the note at middle C and its harmonic frequencies. For each tone, the magnitude and phase relationships derived from the fast Fourier transform peaks were applied to sets of sinusoidal waves corresponding to the F0 and its harmonics. The sinusoidal waves were summed to generate each tone. A uniform temporal envelope with exponential onset (50 ms) and linear decay (710 ms) was applied to the summed sinusoidal waves to create the final synthetic complex tones of 760 ms duration.

For the complex-tone pitch-direction discrimination test, a two-alternative forced-choice (AFC) with one-up and one-down adaptive procedure (Levitt, 1971) was used to measure the thresholds in semitones. Synthesized complex tones with duration of 760 ms were presented at 65 dBA. The presentation level was roved within trials (± 4 dB range in 1-dB steps) to minimize level cues. On each presentation, a tone at the reference F0 (i.e. standard stimulus) and a higher pitched tone (i.e. comparison stimulus with larger F0) were played in random order. The subjects were instructed to identify which tone sounded higher in pitch by clicking on a button using a computer mouse. Prior to actual pitch-direction discrimination testing, four single trials were done for familiarity with visual feedback. The actual testing was administered with three reference F0s (262, 330, and 392 Hz). Feedback was not provided during testing. The initial interval presented was always 12 semitones (1 octave), which is almost always discriminable by CI listeners. For the next two presentations, if the listeners made correct responses, the interval size of 6 and 3 semitones were used. For the rest of tracking, the step size of 1 semitone was used. If an incorrect response was made during the first three presentations, the interval size was immediately set to 1 semitone and the rest of the testing was completed with the step size of 1 semitone. The threshold was estimated as the mean interval size for three adaptive tracks, each determined from the mean of the final 6 of 8 reversals. A reversal at zero was automatically added by the test algorithm when the listeners answered correctly at 1 semitone. Due to theoretical concerns related to the 1-up 1-down staircase method with a 2-AFC task, the data were analyzed using the Spearman-Kärber method (Ulrich and Miller, 2004) to determine a threshold estimate corresponding to 75% correct performance. Instead of fitting a curve to the psychometric function and extracting a single point, the Spearman-Kärber method incorporates all of the data and estimates a midpoint for 75% correct in 2-AFC tasks. Pitch-direction discrimination thresholds estimated with the Spearman-Kärber method are used for all results and data analyses in this study.

In the melody recognition test, 12 isochronous and familiar melody clips were used. They include Frère Jacques, Happy Birthday, Here Comes the Bride, Jingle Bells, London Bridge, Mary Had a Little Lamb, Old MacDonald, Rock-a-Bye Baby, Row Row Row Your Boat, Silent Night Holy Night, Three Blind Mice, and Twinkle Twinkle Little Star. Synthesized complex tones with duration of 500 ms were used and they were ramped with 50 ms of exponential onset and 450 ms of linear decay. The tones were repeated in an eight note pattern at a tempo of 60 beats per minute to eliminate rhythm cues. Stimulus level was 65 dBA. The amplitude of each note was randomly roved by 4-dB. The melody test began with a brief training session in which subjects listened to each melody two times. In training, they were aware which melody they were listening to. During the actual testing, each melody was played 3 times in random order across the set of 36 presentations and the subjects were asked to identify melodies by clicking on the title corresponding to the melodies presented. A total score was calculated after

36 melody presentations as the percent of melodies correctly identified. Feedback was not provided.

In the timbre (musical instrument) recognition test, sound clips of live recordings for 8 western musical instruments playing an identical 5-note sequence were used. This melodic sequence encompassed the octave above middle C playing bidirectional interval changes at a uniform tempo of 82 beats per minute. The instruments used include piano, violin, cello, acoustic guitar, trumpet, flute, clarinet, and saxophone. The timbre test began with a training session in which subjects listened to each instrument two times. During the actual testing, each instrument sound clip was played 3 times in random order at 65 dBA and the subjects were instructed to click on the labeled icon of the instrument corresponding to the timbre presented. A total percent correct score was calculated after 24 presentations as the percent of instruments correctly identified. Feedback was not provided.

B. Speech reception threshold (SRT) in background noise—The subjects were asked to identify one randomly chosen spondee word out of a closed-set of 12 equally difficult spondees (Harris, 1991) in the presence of background noise (Turner *et al.*, 2004; Won *et al.*, 2007). The spondees, two-syllable words with equal emphasis on each syllable (e.g. “birthday”, “padlock”, “sidewalk”), were recorded by a female talker (F0 range: 212–250 Hz, Turner *et al.*, 2004). Two background noises were used: steady-state, speech-shaped noise and two-talker babble noise. The two-talker babble consisted of a male voice saying “Name two uses for ice” and a female voice saying “Bill might discuss the foam.” The female talker for the babble was different from the female talker for the spondees. Duration of the background noise was 2.0 s and the onset of the spondees was 500 ms after the onset of the background noise. The same wave file of background noise was used on every trial in order to eliminate variance that might arise from variability in the background stimulus. The 12-AFC task used one-up, one-down adaptive tracking, converging on 50% correct (Levitt, 1971). The level of the target speech was 65 dBA. The noise level was varied with a step size of 2-dB. Feedback was not provided. For all subjects, the adaptive track started with +10 dB signal-to-noise ratio (SNR) condition. The threshold for a single adaptive track was estimated by averaging the SNR for the final 10 of 14 reversals. Forty-eight of the 51 subjects were tested. The primary dependent variable was the mean SRT of six adaptive tracks.

C. Spectral-ripple stimuli discrimination test—The spectral-ripple discrimination test in this study is the same as that previously described by Won *et al.* (2007), similar to Henry *et al.* (2005). Two-hundred pure-tone frequency components were summed to generate the rippled noise stimuli. The 200 tones were spaced equally on a logarithmic frequency scale. The amplitudes of the components were determined by a full-wave rectified sinusoidal envelope on a logarithmic amplitude scale. The ripple peaks were spaced equally on a logarithmic frequency scale. The stimuli had a bandwidth of 100–5,000 Hz and a peak-to-valley ratio of 30 dB. The mean presentation level of the stimuli was 61 dBA and randomly roved ± 4 -dB in 1-dB steps. The starting phases of the components were randomized for each presentation. The ripple stimuli were generated with 14 different densities, measured in ripples per octave. The ripple densities differed by ratios of 1.414 (0.125, 0.176, 0.250, 0.354, 0.500, 0.707, 1.000, 1.414, 2.000, 2.828, 4.000, 5.657, 8.000, and 11.314 ripples/octave). Standard (reference stimulus) and inverted (ripple phase reversed test stimulus) ripple stimuli were generated. For standard ripples, the phase of the full-wave rectified sinusoidal spectral envelope was set to zero radians, and for inverted ripples, it was set to $\pi/2$. The stimuli had 500 ms total duration and were ramped with 150 ms rise/fall times. Stimuli were filtered with a long-term, speech-shaped filter (Byrne *et al.*, 1994). A 3-AFC, two-up and one-down adaptive procedure was used to determine the spectral-ripple resolution threshold converging on 70.7% correct (Levitt, 1971). Each adaptive track started with 0.176 ripples/octave and moved in equal ratio steps of 1.414. Feedback was not provided. The threshold for a single adaptive track was estimated by

averaging the ripple spacing (the number of ripples/octave) for the final 8 of 13 reversals. The primary dependent variable for this test was the mean threshold estimated from six adaptive tracks. All 42 subjects were tested.

D. Schroeder-phase discrimination test—The Schroeder-phase discrimination test in the current study is the same as that previously described by Drennan *et al.* (2008), and similar to Dooling *et al.* (2002). Positive and negative Schroeder-phase stimulus pairs were created for four different F0s of 50, 100, 200 and 400 Hz. For each F0, equal-amplitude harmonics from the F0 up to 5 kHz were summed. Phase values for each harmonic were determined by the following equation:

$$\theta_n = \pm \pi n(n+1)/N$$

where θ_n is the phase of the n th harmonic, n is the n th harmonic, N is the total number of harmonics in the complex. The positive or negative sign was used for the positive or negative Schroeder-phase stimuli, respectively. The Schroeder-phase stimuli were presented at 65 dBA without roving the level. A 4-interval, 2-AFC procedure was used. One stimulus (i.e. positive Schroeder-phase, test stimulus) occurred in either the second or third interval and was different from three others (i.e. negative Schroeder-phase, reference stimulus). The subject's task was to discriminate the test stimulus from the reference stimuli. Visual feedback was provided. To determine a total percent correct for each F0, the method of constant stimuli was used. In a single test block, each F0 was presented 24 times in random order and a total percent correct for each F0 was calculated as the percent of stimuli correctly identified. Thirty-six of the 42 subjects participated in this test. The dependent variables for this test were the mean percent correct of six test blocks for each F0 and an overall average percent correct score by averaging the scores from each of four F0s. There were 144 presentations for each F0.

E. CNC word recognition test—Fifty CNC monosyllabic words (Peterson and Lehiste, 1962) were presented in quiet at 62 dBA. A CNC word list was randomly chosen out of ten lists for each subject. The subjects were instructed to repeat the word that they heard. A total percent correct score was calculated after 50 presentations as the percent of words correctly repeated. Thirty-seven of the 42 subjects were tested.

3. Data analysis

Correlations of the CAMP subtests with the other tests in the battery described above were assessed using a Pearson and Spearman correlation coefficients to ensure that parametric and non-parametric correlations were both significant in case any of the variables were not normally distributed. Partial correlation analyses were conducted to determine the extent to which the CAMP subtests scores correlated with speech tasks independent of spectral resolution. Additionally, using spectral-ripple thresholds, Schroeder-phase scores, and pitch-direction discrimination scores as predictors, multiple linear regression analyses with two predictors were conducted to explain the variability of melody recognition, timbre recognition and SRTs in noise. To be considered valid, the adjusted R^2 for the combination of factors had to be greater than the R^2 of the independent variables, the p -value for each coefficient had to be less than 0.05, the p -value for the F-ratio had to be less than 0.05, the 95% confidence interval of each of the regression coefficients could not include 0, and the regression standardized residual had to be normally distributed (Lomax, 2005).

III. RESULTS

Average scores for all tests are reported in Table 2A. The results for the present study are consistent with previously reported data (the CAMP subtests: Nimmons *et al.*, 2008, Kang *et*

al., 2009; spectral-ripple discrimination: Won *et al.*, 2007; Schroeder-phase discrimination: Drennan *et al.*, 2008). For tests which were repeated 6 times to determine average thresholds, the variability of the test was evaluated by comparing the first 3 blocks to the second 3 blocks (Table 2B). As shown in Table 2B, most tests showed strong correlation between the thresholds obtained from the first and second 3 blocks. A paired t-test was also conducted to determine if learning had occurred throughout the course of 6 repetitions. Spectral-ripple discrimination and Schroeder-phase discrimination did not show a difference between the first and second 3 blocks, whereas speech recognition in noise showed improvement (1.4 dB in babble, 0.6 dB in steady-noise). These observations are also consistent with previous reports (Won *et al.*, 2007; Drennan *et al.*, 2008). The Shapiro-Wil test (alpha level = 0.05) showed that the timbre recognition scores, Schroeder-phase discrimination, average scores across the four F0s, SRTs in two-talker babble, and CNC word recognition scores were not significantly different from a normal distribution. Pearson and Spearman correlations are reported to indicate that parametric and nonparametric correlations are both significant, respectively. If both correlations are significant, it ensures the fact that the distributions are not purely normal is irrelevant, because nonparametric correlation does not assume normal distribution of variables (Hotelling and Pabst 1936). The number of subjects varies in the correlation analysis depending upon the number of subjects that did both of the tests being correlated.

1. Correlation between spectral-ripple discrimination and all music measures

Significant correlations were found between the spectral-ripple thresholds and performances on all of the CAMP subtests. Figure 1 shows the scattergram of spectral-ripple thresholds and mean pitch-direction discrimination thresholds for 41 subjects (Pearson $r = -0.49$, $p = 0.001$; Spearman $\rho = -0.58$, $p < 0.0001$). Table 3 shows the correlations of the spectral-ripple thresholds with the pitch-direction discrimination thresholds for each of three F0s. There was a trend that the correlation became greater with higher F0s for the pitch-direction discrimination, but the difference was not statistically significant. The left panel of Figure 2 shows that the melody recognition scores significantly correlated with spectral-ripple thresholds (Pearson $r = 0.65$, $p < 0.0001$; Spearman $\rho = 0.62$, $p < 0.0001$, $N = 42$). The right panel of Figure 2 shows the scattergram of spectral-ripple threshold and timbre recognition scores (Pearson $r = 0.57$, $p < 0.0001$; Spearman $\rho = 0.60$, $p < 0.0001$, $N = 42$). The results suggest that good spectral resolution is important for music perception in CI users.

2. Correlation of Schroeder-phase discrimination

A significant correlation was found between Schroeder-phase discrimination scores at 200 Hz and timbre recognition scores (Pearson $r = 0.37$, $p = 0.03$; Spearman $\rho = 0.38$, $p = 0.02$, $N = 36$). Other than that, Schroeder-phase discrimination scores were not correlated with any CAMP subtests as shown in Table 4. Note that only 36 of the 42 subjects participated in the Schroeder-phase discrimination test. Consistent with a previous report (Drennan *et al.*, 2008), Schroeder-phase discrimination significantly correlated with CNC word recognition (Pearson $r = 0.52$, $p < 0.01$; Spearman $\rho = 0.59$, $p = 0.001$; $N = 28$) and with speech perception in steady-state noise (Pearson $r = -0.48$, $p = 0.007$; Spearman $\rho = -0.57$, $p = 0.001$, $N = 31$).

3. Prediction for melody and timbre recognition using the combination of the spectral-ripple and pitch-direction discrimination

Multiple linear regression analyses were conducted to explain variability of melody and timbre recognition scores and SRTs in noise. Table 5 shows the multiple linear regression data which met the criteria described in the Methods section. Figure 3 shows that the combination of spectral-ripple thresholds and mean pitch-direction discrimination thresholds predicted nearly half of the variance in melody and timbre scores (49% for melody and 47% for timbre). These multiple linear regression analyses showed that when coupling spectral ripple discrimination

with mean pitch-direction discrimination ability (averaged over 3 F0s), pitch-direction discrimination ability accounted for *an additional* 2% to 15% of the variance of the music and speech performance over variance accounted for with spectral-ripple discrimination alone (Table 5). One might speculate that some aspect of CI users' pitch perception ability that is not captured by spectral-ripple discrimination ability which would have contributed to the additional predictive power.

4. Relationship between music perception and speech perception

Pitch-direction discrimination was significantly correlated with CNC word recognition and speech perception in noise. Figure 4 shows that mean pitch-direction discrimination thresholds were significantly correlated with CNC recognition scores (left panel, Pearson $r = -0.56$, $p < 0.001$; Spearman $\rho = -0.55$, $p < 0.001$) and SRTs in two-talker babble (middle panel, Pearson $r = 0.54$, $p < 0.001$; Spearman $\rho = 0.54$, $p < 0.001$) and in steady-state noise (right panel, Pearson $r = 0.50$, $p < 0.01$; Spearman $\rho = 0.49$, $p < 0.01$). Correlations between the pitch-direction discrimination and CNC scores and SRTs in noise are summarized in Table 6A.

In order to determine the extent to which the pitch-direction discrimination thresholds correlated with the speech measures independent of spectral resolution, a partial correlation analysis was done controlling for the effects of spectral-ripple resolution (Table 6B). For the pitch-direction discrimination at 262 Hz, there was little change in the correlation when the effects of spectral ripple discrimination ability were factored out. This suggests that (1) spectral-ripple discrimination has little or no effect on the relationship between the pitch-direction discrimination at 262 Hz and speech perception; and (2) spectral resolution and complex-tone pitch-direction discrimination ability at low fundamental frequencies (< 300 Hz) tend to contribute independently to predicting speech perception in noise. However, the correlation of pitch-direction discrimination thresholds at 392 Hz with SRTs went to nearly zero ($r = 0.07$ in babble; $r = 0.10$ in steady-state noise) when the spectral-ripple effects was factored out. The relationship between pitch direction discrimination at 392 Hz and CNC scores also became weak ($r = -0.28$) and nonsignificant ($p = 0.12$). One might speculate that temporal factors are contributing to pitch discrimination performance for fundamental frequencies less than 300 Hz, but above 300 Hz, spectral factors are playing the primary role.

The melody and timbre recognition scores were also significantly correlated with CNC word recognition and speech perception in noise. Figure 5 shows that CNC word recognition scores were significantly correlated with melody recognition scores (left panel, Pearson $r = 0.43$, $p = 0.007$; Spearman $\rho = 0.48$, $p = 0.003$) and timbre recognition scores (right panel, Pearson $r = 0.62$, $p < 0.001$; Spearman $\rho = 0.67$, $p < 0.001$). The upper panel of Figure 6 shows the scattergram of melody recognition vs. SRTs in two-talker babble (left panel, Pearson $r = -0.49$, $p = 0.002$; Spearman $\rho = -0.54$, $p < 0.001$) for 38 subjects and in steady-state noise (right panel, Pearson $r = -0.42$, $p = 0.008$; Spearman $\rho = -0.54$, $p < 0.001$) for 40 subjects. The lower panel of Figure 6 shows with the scattergram of timbre recognition vs. SRTs in two-talker babble (left panel, Pearson $r = -0.61$, $p < 0.0001$; Spearman $\rho = -0.63$, $p < 0.0001$) for 38 subjects and in steady-state noise (right panel, Pearson $r = -0.63$, $p < 0.0001$; Spearman $\rho = -0.68$, $p < 0.0001$) for 40 subjects.

The partial correlation analyses described above were also performed for the relationships among melody recognition, timbre recognition, CNC word recognition, and speech perception in noise. When the spectral-ripple thresholds were factored out, the correlations between the melody scores and speech scores (SRTs in noise and CNC percent correct) became zero ($r = -0.02$ for SRTs in babble; $r = 0.04$ for SRTs in steady-state noise; $r = 0.04$ for CNC scores), suggesting that spectral resolution ability accounts for the relationship between these speech measures and melody recognition.

IV. DISCUSSION

1. Spectral-ripple resolution and music perception

Using chimaeric sounds, Smith *et al.* (2002) suggested that high spectral resolution would be required for melody recognition. Kong *et al.* (2004) found that as many as 32 vocoder frequency bands would be needed for rhythmless melody recognition in normal-hearing listeners. Laneau *et al.* (2006) showed a negative effect of spectral smearing on complex harmonic tone pitch perception of noise-band vocoded signals in normal-hearing listeners. All of these previous studies using normal-hearing listeners and acoustic simulation suggest that better spectral resolution would contribute to better music perception in CI users. The present study found significant relationships between spectral-ripple discrimination and pitch-direction discrimination, melody recognition, and timbre recognition, showing that CI users who had better spectral-ripple discrimination ability had better music perception. Thus, the hypothesis that better spectral resolution contributes to better music perception is supported by the present study.

In order to discriminate standard and inverted spectral-ripple stimuli, CI listeners must be able to detect and discriminate relative amplitude changes across channels (Henry and Turner, 2003). Figure 7 shows the CI sound processor output for Fidelity120 corresponding to the 16 electrodes for the spectral-ripple densities of 0.5, 1.0, 2.0, and 4.0 ripples/octave. Average outputs over the duration of the spectral-ripple stimuli (0.5 sec) for each electrode are plotted. For the ripple densities of 0.5 to 2 ripples/octave, multiple peaks and valleys are well represented across the electrodes. In addition, there is a gradual decrease in the distance between the peaks and valley as the ripple density increases, resulting in reduced spectral contrast between the standard and inverted ripple stimuli, especially at high ripple density (4 ripples/octave).

A theoretical mechanism for the spectral-ripple discrimination follows the place theory of pitch perception and such a mechanism is also likely to contribute to the ability to discriminate changes in the F0 of harmonic complexes, timbre and pitch contours. Timbre is based on the physical characteristics of sound including the spectral energy distribution and temporal envelope, especially its attack and decay characteristics (e.g. Plomp, 1970; Grey, 1977). The significant correlation found between the spectral-ripple thresholds and timbre scores demonstrate that CI subjects depend in part (32% of the variance) on the spectral properties of sound to identify the timbre. This finding in CI users parallels previous findings in normal-hearing listeners that the spectral envelope of the sound is the most salient parameter in timbre discrimination (e.g. McAdams *et al.*, 1999; Gunawan and Sen, 2008).

2. Pitch-direction discrimination thresholds: Comparison with previous results

Pitch-direction discrimination thresholds in this study were higher than those from previous studies (Nimmons *et al.*, 2008; Kang *et al.*, 2009). Thresholds in the present study were estimated with the Spearman-Kärber method (Ulrich and Miller, 2004), incorporating all of the raw data of the responses (right or wrong) for each and every presentation in the pitch-direction discrimination task. The Spearman-Kärber method estimates thresholds for 75% correct, whereas the mean-of-reversals methods estimates near 50% correct. Thresholds reported in the results above were therefore greater (mean = 4.6 semitones) than thresholds estimated using mean-of-reversals approach (2.6 semitones in the present study; 3.0 semitones in Kang *et al.*, 2009). The thresholds estimated with each method were, however, highly correlated ($r = 0.97$), and consequently, the method of threshold calculation had no material effect on the correlation results. Kang *et al.* (2009) also showed the intraclass correlation of 0.85 for test-retest reliability of the pitch-direction discrimination using the 1-up 1-down staircase method with a 2-AFC task, demonstrating that approach produced a reliable result.

3. Schroeder-phase discrimination ability and music measures

Except the weak correlation between Schroeder-phase discrimination scores at 200 Hz and timbre recognition scores (Pearson $r = 0.37$, $p = 0.03$; Spearman $\rho = 0.38$, $p = 0.02$, $N = 36$), no correlations were found between the music measures and Schroeder-phase discrimination. CIs sound processors using pulsatile stimulation strategies transform the acoustic Schroeder-phase stimuli into sweeps of envelope packets that either rise or fall in frequency and repeat over time. Drennan *et al.* (2008) suggested that CI subjects use between-channel timing differences in the temporal envelopes to discriminate positive and negative Schroeder-phase stimuli. In order to use a between-channel cue, good sensitivity to detect and discriminate fast changing spectral differences is required. However, there is no such dynamic spectral change in the pitch-direction discrimination and timbre recognition tests. In the melody recognition test, the tones had 500 ms duration and were repeated in an eight note pattern. This dynamic change in the melody recognition test is much slower than in the Schroeder-phase discrimination test. In contrast to music, correlations between Schroeder-phase discrimination and speech measures were reported (50-Hz Schroeder vs. CNC: $r = 0.52$; 200-Hz Schroeder vs. SRT in steady-state noise: $r = -0.48$; Drennan *et al.*, 2008) and similar correlations were also found in the current study. Drennan *et al.* suggested that the ability to discriminate dynamic spectral change might underlie the ability of CI listeners to hear the trajectory of a typical consonant-vowel transitions and to segregate the speech of a female talker (which has a F0 of approximately 200-Hz) from steady-state noise. Hearing fast spectral changes such as those in speech were not necessary in CAMP tasks. Good Schroeder-phase discrimination might be related to the ability to perceive faster musical passages such as rapidly ascending or descending scales or arpeggios.

4. Implications for development of cochlear implants

In order to develop a CI system to better represent musical sound signals while still maintaining or improving speech signal representation, it is important to understand which acoustic cues underlie both music and speech perception. Although speech and music have somewhat different perceptual demands, previous studies have shown that the performance of speech correlates with music perception in CI users. Gfeller *et al.* (2002a) investigated the relationship between melody recognition and several types of speech tests and found correlations ranged from 0.51 to 0.61 in 42 CI users. Galvin *et al.* (2007) showed a correlation between the melodic contour identification test and vowel recognition ($r = 0.73$, $p = 0.01$) in 11 CI subjects. For 7 CI subjects, Gfeller *et al.* (2007) found a correlation between melody recognition and speech recognition in noise ($r = -0.55$ in broadband noise; $r = -0.50$ in babble noise), although it was not statistically significant ($p > 0.05$). In the current study, melody and timbre recognition scores significantly correlated with CNC word recognition in quiet listening condition ($r = 0.43$ for melody; $r = 0.62$ for timbre), SRTs in two-talker babble ($r = -0.49$ for melody; $r = -0.61$ for timbre), and SRTs in steady-state noise ($r = -0.42$ for melody; $r = -0.63$ for timbre). With the effect of spectral-ripple thresholds removed, the correlation between melody recognition and speech recognition (CNC scores and SRTs in two types of noise) became nearly zero. This suggests that the ability to resolve spectral changes contributes to both music and speech perception abilities.

Significant correlations from behavioral data do not necessarily reflect a causative mechanism. It is plausible that both spectral ripple discrimination ability and music and speech understanding abilities are limited by the condition of the peripheral auditory system. The peripheral auditory systems of individual CI users have a variety of different conditions which might account for the correlations observed; however, “higher-level” auditory processing capabilities could also play a role.

Analysis of performance using the psychoacoustic tests might be valuable for analysis of the effectiveness of engineering manipulations to sound processing. The present data suggest that attempts to improve spectral resolution by manipulating sound processor parameters or developing new hardware and software could lead to improvement in multiple clinical outcomes such as music and speech perception.

Acknowledgments

We appreciate the dedicated efforts of our subjects. Elyse Jameyson assisted with data collection. This study was supported by NIH grants R01-DC007525, P30-DC04661, P50-DC00242, F31-DC009755, Cochlear Corporation, and Advanced Bionics Corporation. Neither company played any role in data analysis or the composition of this paper.

REFERENCES

- Cazals Y, Pelizzone M, Saudan O, et al. Low-pass filtering in amplitude modulation detection associated with vowel and consonant identification in subjects with cochlear implants. *J Acoust Soc Am* 1994;96:2048–2054. [PubMed: 7963020]
- Byrne D, Dillon H, Tran K, Arlinger S, et al. An international comparison of long-term average speech spectra. *J. Acoust. Soc. Am* 1994;96:2108–2120.
- Collins LM, Zwolan TA, Wakefield GH. Comparison of electrode discrimination, pitch ranking, and pitch scaling data in postlingually deafened adult cochlear implant subjects. *J Acoust Soc Am* 1997;101:440–455. [PubMed: 9000735]
- Donaldson GS, Nelson DA. Place-pitch sensitivity and its relation to consonant recognition by cochlear implant listeners using the MPEAK and SPEAK speech processing strategies. *J Acoust Soc Am* 2000;107:1645–1658. [PubMed: 10738818]
- Doolling RJ, Leek MR, Gleich O, et al. Auditory temporal resolution in birds: Discrimination of harmonic complexes. *J Acoust Soc Am* 2002;112:748–759. [PubMed: 12186054]
- Drennan WR, Longnion JK, Ruffin C, et al. Discrimination of schroeder-phase harmonic complexes by normal-hearing and cochlear-implant listeners. *J Assoc Res Otolaryngol* 2008;9:138–149. [PubMed: 18066624]
- Drennan WR, Rubinstein JT. Music perception in cochlear implant users and its relationship with psychophysical capabilities. *J Rehabil Res Dev* 2008;45:779–790. [PubMed: 18816426]
- Drennan WR, Won JH, Jameyson E, Nie K, Rubinstein JT. Sensitivity of psychophysical measures to signal processor modifications in cochlear implant users. *Hear Res.* 2010 In press.
- Fu QJ. Temporal processing and speech recognition in cochlear implant users. *Neuroreport* 2002;13:1635–1639. [PubMed: 12352617]
- Galvin JJ 3rd, Fu QJ, Nogaki G. Melodic contour identification by cochlear implant listeners. *Ear Hear* 2007;28:302–319. [PubMed: 17485980]
- Gfeller K, Lansing CR. Melodic, rhythmic, and timbral perception of adult cochlear implant users. *J Speech Hear Res* 1991;34:916–920. [PubMed: 1956198]
- Gfeller K, Woodworth G, Robin DA, et al. Perception of rhythmic and sequential pitch patterns by normally hearing adults and adult cochlear implant users. *Ear Hear* 1997;18:1–15. [PubMed: 9058033]
- Gfeller K, Turner C, Mehr M, et al. Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. *Cochlear Implants International* 2002a;3:31–55.
- Gfeller K, Witt S, Woodworth G, et al. Effects of frequency, instrumental family, and cochlear implant type on timbre recognition and appraisal. *Ann Otol Rhinol Laryngol* 2002b;111:349–356. [PubMed: 11991588]
- Gfeller K, Olszewski C, Rychener M, et al. Recognition of "real-world" musical excerpts by cochlear implant recipients and normal-hearing adults. *Ear Hear* 2005;26:237–250. [PubMed: 15937406]
- Gfeller K, Turner C, Oleson J, 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]
- Grey JM. Multidimensional perceptual scaling of musical timbres. *J Acoust Soc Am* 1977;61:1270–1277. [PubMed: 560400]

- Gunawan D, Sen D. Spectral envelope sensitivity of musical instrument sounds. *J Acoust Soc Am* 2008;123:500–506. [PubMed: 18177177]
- Harris, RW. Speech audiometry materials compact disk. Provo, UT: Brigham Young University; 1991.
- Henry BA, McKay CM, McDermott HJ, et al. The relationship between speech perception and electrode discrimination in cochlear implantees. *J Acoust Soc Am* 2000;108:1269–1280. [PubMed: 11008827]
- Henry BA, Turner CW. The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. *J Acoust Soc Am* 2003;113:2861–2873. [PubMed: 12765402]
- Henry BA, Turner CW, Behrens A. Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners. *J Acoust Soc Am* 2005;118:1111–1121. [PubMed: 16158665]
- Hong RS, Turner CW. Pure-tone auditory stream segregation and speech perception in noise in cochlear implant recipients. *J Acoust Soc Am* 2006;120:360–374. [PubMed: 16875232]
- Hotelling H, Pabst MR. Rank correlation and tests of significance involving no assumption of normality. *Ann. Math. Statist* 1936;7:29–43.
- Kang RS, Nimmons GL, Drennan WR, et al. Development and validation of the University of Washington clinical assessment of music perception (CAMP) 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]
- Laneau J, Moonen M, Wouters J. Factors affecting the use of noise-band vocoders as acoustic models for pitch perception in cochlear implants. *J Acoust Soc Am* 2006;119:491–506. [PubMed: 16454303]
- Levitt H. Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* 1971;49 Suppl 2:467–477. [PubMed: 5541744]
- Litvak LM, Spahr AJ, Saoji AA, et al. Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. *J Acoust Soc Am* 2007;122:982–991. [PubMed: 17672646]
- Lomax, RG. Statistical Concepts: A Second Course. Mahwah, NJ: Lawrence Erlbaum Associates; 2005.
- Luo, Xin; 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]
- McAdams S, Beauchamp JW, Meneguzzi S. Discrimination of musical instrument sounds resynthesized with simplified spectrotemporal parameters. *J Acoust Soc Am* 1999;105:882–897. [PubMed: 9972573]
- Nelson DA, Van Tasell DJ, Schroder AC, et al. Electrode ranking of "place pitch" and speech recognition in electrical hearing. *J Acoust Soc Am* 1995;98:1987–1999. [PubMed: 7593921]
- Nimmons GL, Kang RS, Drennan WR, et al. Clinical assessment of music perception in cochlear implant listeners. *Otol Neurotol* 2008;29:149–155. [PubMed: 18309572]
- Peterson GE, Lehiste I. Revised CNC lists for auditory tests. *J Speech Hear Disord* 1962;27:62–70. [PubMed: 14485785]
- Plomp, R. Timbre as a multidimensional attribute of complex tones in frequency analysis and periodicity detection in hearing. In: Plomp, R.; Smoorenburg, GF., editors. *Frequency Analysis and Periodicity Detection in Hearing*. Leiden: Sijthoff; 1970.
- Saoji AA, Litvak L, Spahr AJ, Eddins DA. Spectral modulation detection and vowel and consonant identifications in cochlear implant listeners. *J Acoust Soc Am* 2009;126:955–958. [PubMed: 19739707]
- Smith ZM, Delgutte B, Oxenham AJ. Chimaeric sounds reveal dichotomies in auditory perception. *Nature* 2002;416:87–90. [PubMed: 11882898]
- Turner CW, Gantz BJ, Vidal C, et al. Speech recognition in noise for cochlear implant listeners: Benefits of residual acoustic hearing. *J Acoust Soc Am* 2004;115:1729–1735. [PubMed: 15101651]
- Ulrich R, Miller J. Threshold estimation in two-alternative forced-choice (2AFC) tasks: The Spearman-Kärber method. *Percep & Psychophys* 2004;66:517–533.
- 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, Dorman MF. Cochlear implants: A remarkable past and a brilliant future. *Hear Res* 2008;242:3–21. [PubMed: 18616994]
- Won JH, Drennan WR, Rubinstein JT. Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. *J Assoc Res Otolaryngol* 2007;8:384–392. [PubMed: 17587137]
- Won JH, Drennan WR, Jameyson E, Rubinstein JT. Single-channel Schroeder-phase discrimination as a measure of within-channel temporal fine-structure sensitivity. *Assoc. Res. Otolaryngol.* 2009 Abstract #884.
- Zwolan TA, Collins LM, Wakefield GH. Electrode discrimination and speech recognition in postlingually deafened adult cochlear implant subjects. *J Acoust Soc Am* 1997;102:3673–3685. [PubMed: 9407659]

Figure 1. Relationship between spectral-ripple discrimination and pitch-direction discrimination. Linear regression is represented by the solid line.

Figure 2. Relationship between spectral-ripple thresholds and melody scores (left panel) and timbre scores (right panel). Linear regressions are represented by the solid lines.

Figure 3.

Left panel shows the multiple regressions for melody recognition using spectral-ripple thresholds and mean pitch-direction thresholds ($r = 0.72$, adjusted $R^2 = 0.49$, $p < 0.0001$; whereas, independently, $R^2 = 0.42$, $p < 0.0001$ and $R^2 = 0.34$, $p < 0.0001$ for spectral-ripple and mean pitch-direction discrimination thresholds, respectively). Right panel shows the multiple regressions for timbre recognition using spectral-ripple thresholds and mean pitch-direction thresholds ($r = 0.70$, adjusted $R^2 = 0.47$, $p < 0.0001$; whereas, independently, $R^2 = 0.32$, $p < 0.0001$ and $R^2 = 0.41$, $p < 0.0001$ for spectral-ripple and mean pitch-direction discrimination thresholds, respectively). Linear regressions between the true and predicted scores are represented by the solid lines.

Figure 4. Relationship between mean pitch thresholds and CNC scores (left panel), SRTs in two-talker babble (middle panel), and SRTs in steady-state noise (right panel). Linear regressions are represented by the solid lines.

Figure 5. Relationship between CNC scores and melody recognition scores (left panel) and timbre recognition scores (right panel). Linear regressions are represented by the solid lines.

Figure 6.
Relationship between SRTs and melody scores (left panel) and timbre scores (right panel).
Linear regressions are represented by the solid lines.

Figure 7. Sound processor output for spectral-ripple densities of 0.5, 1.0, 2.0, and 4.0 ripples/octave with standard ripple (filled circles) and inverted ripple (open circles). Electrode 16 represents for the highest channel.

Table 1

Subject characteristics

Subject	Age (yrs)	Duration of hearing loss (yrs) ^a	Duration of implant use (yrs)	Etiology	Implant type	Speech Processor Strategy	CNC word score (% correct)
S01	61	0.3	2	Unknown	Nucleus 24	ACE	92
S02	69	3	8	Unknown	Nucleus 24	ACE	42
S03	61	5	11	Genetic	Nucleus 22	SPEAK	72
S04	62	1	3	Unknown	Nucleus 24	ACE	88
S09	41	0.5	0.5	Head trauma	HiRes90K	HiRes	40
S10	76	5	0.3	Otosclerosis	HiRes90K	HiRes	50
S12	49	0	2	Connexin 26	MedEl Combi40+	CIS	
S14	46	10	1	Unknown	Nucleus 22	ACE	
S15				Unknown	Nucleus 22	SPEAK	
S16	50	17	1	Unknown	HiRes90K	HiRes	46
S17	55	19	7	Unknown	Nucleus 24	ACE	
S18	74		6	Sudden hearing loss	Nucleus 24	ACE	52
S19	63	12	3	Genetic	Nucleus24	ACE	88
S20	53	10	2	Unknown	Nucleus24	ACE	80
S21	59	8	6	Otosclerosis	(L) Nucleus24; (R) Nucleus24	ACE	90
S25	43	2	1.7	Radiation	Nucleus 24	ACE	80
S30	66	5	3.5	Amikacin	(L)Nucleus Freedom; (R)Nucleus 22	ACE	80
S31	55	1	1	Hereditary	Nucleus Freedom	ACE	90
S33	57	15	16	Unknown	(L) Nucleus 24; (R) Nucleus 22	(L)CIS; (R)SPEAK	70
S34	55		1.5	Noise exposure	HiRes90K	HiRes	62
S35	72	3	2	Unknown	Nucleus 24	ACE	70
S36	56	40	6	Unknown	Nucleus Freedom	ACE	18

Subject	Age (yrs)	Duration of hearing loss (yrs) ^a	Duration of implant use (yrs)	Etiology	Implant type	Speech Processor Strategy	CNC word score (% correct)
S37	81	3	2.5	Unknown	Nucleus 24	ACE	24
S38	51	9	4	Noise-induced and ototoxicity	Nucleus 24	ACE	100
S39	42	0	6	Genetic	Nucleus 24		
S40	72	5	6	Genetic	HiRes90K	HiRes	58
S41	52	7	5	Hereditary	HiRes90K	HiRes	86
S42	68	5	3	Genetic	HiRes90K	Fidelity 120	52
S43	34	3	6	Unknown	Clarion CII	Fidelity 120	24
S48	67	10	0.5	Unknown	HiRes90K	HiRes	72
S49	64	4	0.75	Hereditary	HiRes90K	Fidelity 120	86
S50	40	15	6	Genetic	Clarion CII	Fidelity 120	100
S51	56	7	6	Hereditary	Clarion CII	HiRes	92
S52	77	0	0.5	Noise exposure	HiRes90K	Fidelity 120	66
S53	63	3	7	Unknown	Clarion CII	Fidelity 120	96
S54	25	0.5	2.5	Unknown	HiRes90K	HiRes	94
S55	65	40	1	Genetic	(L) HiRes90K; (R) HiRes90K	(L)HiRes; (R)HiRes	66
S56	28	15	11	Unknown	Nucleus 22	SPEAK	80
S58	64	57	7	Noise exposure	Clarion CII	Fidelity 120	62
S59	47	12	2.5	Noise exposure	HiRes90K	Fidelity 120	85
S61	78	10	1	Genetic	HiRes90K	Fidelity 120	60
S62	32	3	1	Unknown	(L)HiRes90K; (R)HiRes90K	Fidelity 120	78

^aThe duration of their hearing loss before implantation.

Table 2

(A) Mean scores and 95% confidence intervals for individual tests		
Test	Average scores	± 95% confidence interval
Mean Pitch (semitones)	4.6	1.2
Pitch at 262 Hz (semitones)	5.1	1.6
Pitch at 330 Hz (semitones)	5.0	1.6
Pitch at 392 Hz (semitones)	3.7	1.7
Melody (%)	29.6	7.7
Timbre (%)	48.2	5.2
SRT in two-talker babble (dB SNR)	−8.1	1.9
SRT in steady-state noise (dB SNR)	−8.0	1.5
CNC recognition (%)	69.5	7.1
Spectral-ripple (ripples/octave)	2.1	0.4
Mean Schroeder-phase (%)	64.4	3.7
Schroeder-phase at 50 Hz (%)	76.4	5.4
Schroeder-phase at 100 Hz (%)	71.5	6.0
Schroeder-phase at 200 Hz (%)	57.4	3.0
Schroeder-phase at 400 Hz (%)	52.6	2.1

(B) Variability for individual tests. P-values for paired t-test and correlation between the 1st and 2nd 3 blocks are reported.		
Test	Paired t-test	Correlation
Spectral-ripple discrimination	$p = 0.80$	$r = 0.84, p < 0.0001$
SRT in two-talker babble	$p = 0.006$	$r = 0.90, p < 0.0001$
SRT in steady-state noise	$p = 0.04$	$r = 0.95, p < 0.0001$
Schroeder-phase at 50 Hz	$p = 0.46$	$r = 0.82, p < 0.0001$
Schroeder-phase at 100 Hz	$p = 0.10$	$r = 0.93, p < 0.0001$
Schroeder-phase at 200 Hz	$p = 0.65$	$r = 0.65, p < 0.0001$
Schroeder-phase at 400 Hz	$p = 0.97$	$r = 0.45, p < 0.05$

Table 3

Correlations between spectral-ripple thresholds and complex-tone pitch-direction discrimination thresholds for 41 subjects

	Pitch 262 Hz	Pitch 330 Hz	Pitch 392 Hz	Pitch Average
Pearson	-0.27	-0.45 **	-0.50 **	-0.49 **
Spearman	-0.40 *	-0.58 ***	-0.65 ***	-0.58 ***

* Significant at $p < 0.05$;

** $p < 0.01$;

*** $p < 0.001$

Table 4
Pearson correlations between Schroeder-phase discrimination scores and CAMP subtests scores

Schroeder-phase	F0	50 Hz	100 Hz	200 Hz	400 Hz	AVG
Pitch AVG		0.12	-0.06	-0.22	-0.18	-0.07
Melody		-0.03	0.25	0.17	0.09	0.17
Timbre		0.05	0.34	0.37*	0.06	0.27

* Significant at $p < 0.05$

Table 5
Multiple linear regression analysis results for each combination of the independent and dependent variables

Independent variables		Dependent Variable	R ² (var1 vs. dep)	R ² (var2 vs. dep)	Adjusted R ² (multiple regression)	F-ratio *
Var1	Var2					
Ripple	Pitch AVG	Melody	0.42	0.34	0.49	20.11
Ripple	Pitch AVG	Timbre	0.32	0.41	0.47	18.38
Ripple	Pitch AVG	SRT in babble	0.54	0.29	0.56	24.13
Ripple	Pitch AVG	SRT in SSN	0.33	0.25	0.37	12.37
Ripple	Sch 200-Hz	SRT in SSN	0.33	0.35	0.48	16.90
Ripple	Pitch AVG	CNC	0.36	0.31	0.43	14.37

* Significance for all F-ratios is less than 0.0001

Table 6

(A) Correlations of pitch thresholds with CNC scores and SRTs in two types of noise

Pearson	262-Hz	330-Hz	392-Hz	AVG
CNC scores	-0.48 **	-0.34 *	-0.56 ***	-0.56 ***
SRTs in babble	0.47 **	0.44 **	0.37 *	0.54 ***
SRTs in steady noise	0.48 **	0.36 **	0.31	0.50 **
<hr/>				
Spearman	262-Hz	330-Hz	392-Hz	AVG
CNC scores	-0.45 **	-0.49 **	-0.46 **	-0.55 ***
SRTs in babble	0.42 **	0.51 **	0.48 **	0.54 ***
SRTs in steady noise	0.42 **	0.40 *	0.35 *	0.49 **

(B) Partial correlations of pitch thresholds with CNC scores and SRTs in two types of noise controlling for predictive effect of spectral-ripple thresholds

	262-Hz	330-Hz	392-Hz	AVG
CNC scores	-0.32	-0.04	-0.28	-0.28
SRTs in babble	0.41 *	0.15	0.07	0.31
SRTs in steady noise	0.39 *	0.07	0.10	0.27

* Significant at $p < 0.05$;

** $p < 0.01$;

*** $p < 0.001$

* Significant at $p < 0.05$