Music Perception with Temporal Cues in Acoustic and Electric Hearing

Ying-Yee Kong, Rachel Cruz, J. Ackland Jones, and Fan-Gang Zeng

Objective: **The first specific aim of the present study is to compare the ability of normal-hearing and cochlear implant listeners to use temporal cues in three music perception tasks: tempo discrimination, rhythmic pattern identification, and melody identification. The second aim is to identify the relative contribution of temporal and spectral cues to melody recognition in acoustic and electric hearing.**

Design: **Both normal-hearing and cochlear implant listeners participated in the experiments. Tempo discrimination was measured in a two-interval forced-choice procedure in which subjects were asked to choose the faster tempo at four standard tempo conditions (60, 80, 100, and 120 beats per minute). For rhythmic pattern identification, seven different rhythmic patterns were created and subjects were asked to read and choose the musical notation displayed on the screen that corresponded to the rhythmic pattern presented. Melody identification was evaluated with two sets of 12 familiar melodies. One set contained both rhythm and melody information (rhythm condition), whereas the other set contained only melody information (no-rhythm condition). Melody stimuli were also processed to extract the slowly varying temporal envelope from 1, 2, 4, 8, 16, 32, and 64 frequency bands, to create cochlear implant simulations. Subjects listened to a melody and had to respond by choosing one of the 12 names corresponding to the melodies displayed on a computer screen.**

Results: **In tempo discrimination, the cochlear implant listeners performed similarly to the normal-hearing listeners with rate discrimination difference limens obtained at 4–6 beats per minute. In rhythmic pattern identification, the cochlear implant listeners performed 5–25 percentage points poorer than the normal-hearing listeners. The normal-hearing listeners achieved perfect scores in melody identification with and without the rhythmic cues. However, the cochlear implant listeners performed significantly poorer than the normal-hearing listeners in both rhythm and no-rhythm conditions. The simulation results from normal-hearing listeners showed a relatively high level of performance for all numbers of frequency bands in the rhythm condition but required as many as 32 bands in the no-rhythm condition.**

Conclusions: **Cochlear-implant listeners performed normally in tempo discrimination, but significantly poorer than normal-hearing listeners in rhythmic pattern identification and melody recognition. While both temporal (rhythmic) and spectral (pitch) cues contribute to melody recognition, cochlear-implant listeners mostly relied on the rhythmic cues for melody recognition. Without the rhythmic cues, high spectral resolution with as many as 32 bands was needed for melody recognition for normal-hearing listeners. This result indicates that the present cochlear implants provide sufficient spectral cues to support speech recognition in quiet, but they are not adequate to support music perception. Increasing the number of functional channels and improved encoding of the fine structure information are necessary to improve music perception for cochlear implant listeners.**

(Ear & Hearing 2004;25;173–185)

Cochlear implants are designed to restore hearing in profoundly deafened individuals through electric stimulation. While good speech understanding has been achieved with the modern multi-electrode cochlear implants (70–80% sentence recognition in quiet; for the Clarion device, see Osberger, Fisher, & Kalberer, 2000; for the Med-El device, see Garnham, O'Driscoll, Ramsden, & Saeed, 2002; for the Nucleus devices, see Skinner et al., 1994), music appreciation has been reported to be exceptionally challenging. Some implant listeners reported they could enjoy music and were able to recognize melodies, but most described music as sounding unpleasant and noisy. Previous work demonstrated that cochlear implant listeners could utilize temporal information to perceive musical pitch in the form of electric pulse rate (Pijl, 1997; Pijl & Schwarz, 1995). However, most of the implant listeners appeared to be unable to receive pitch information when the musical stimuli were presented through their speech processor (Gfeller & Lansing, 1991; Gfeller, Woodworth, Robin, Witt, & Knutson, 1997; Leal et al., 2003; Pijl, 1997; Schultz & Kerber, 1994).

There have been a number of studies documenting the temporal resolution of cochlear implant listeners with gap detection, temporal modulation

Department of Otolaryngology (R.C., J.A.J., F.-G.Z.), Department of Biomedical Engineering (F.-G.Z.), and Department of Cognitive Sciences (Y.-Y.K., F.-G.Z.), University of California, Irvine.

DOI: 10.1097/01.AUD.0000120365.97792.2F

0196/0202/04/2502-0173/0 • Ear & Hearing • Copyright © 2004 by Lippincott Williams & Wilkins • Printed in the U.S.A.

detection, and rate discrimination tasks (Moore & Glasberg, 1988; Muller, 1981; Shannon, 1989, 1992). In general, cochlear implant listeners have normal or better than normal temporal resolution following amplitude changes in stimulus waveforms measured by gap detection and amplitude modulation detection tasks. In addition, cochlear implant listeners can use low-frequency temporal information to extract pitch information. Perception of temporal pitch was demonstrated in cochlear implant listeners for stimulation rates up to about 300 Hz (Eddington, Dobelle, Brackmann, Mladevosky, & Parkin, 1978; Shannon, 1983; Zeng, 2002). Pijl and Schwarz (1995) and Pijl (1997) studied music perception in a group of Nucleus 22 cochlear implant listeners and found that they could identify melodies with 80–100% accuracy from a closed set in the absence of rhythmic cues when the melodies were presented by systematic variation of low-frequency electric pulses to a single electrode. Relying solely on the pulse rate, their subjects could also label the pitch intervals as “flat,” “sharp,” or “in tune” with high levels of precision relative to their memory for a specific interval size from the familiar melodies. Their performance in labeling the musical intervals was comparable to that of normal-hearing listeners. While normal-hearing listeners could access both temporal and spectral information in the stimuli, cochlear implant listeners with direct electrical stimulation to a single electrode were presented with only the temporal information. The similar performance between normal-hearing and cochlear-implant listeners suggested that temporal information, in the form of electric pulse rate, is sufficient for musical pitch perception at low frequencies.

Except for Simultaneous Analog Stimulation (SAS), all current processing strategies for cochlear implants remove the temporal fine-structure information in the stimulus waveforms and preserve only the temporal envelopes extracted from 6 to 22 frequency bands. The envelopes are extracted from each frequency band by full-wave rectification and low-pass filtering at a low frequency (<500 Hz). The envelope outputs are finally compressed and then used to modulate biphasic pulses. Trains of modulated biphasic pulses are then delivered to the electrodes at a constant rate between 180 and 300 Hz for the Nucleus-22 device and up to 833 Hz for the Clarion-I device (for a review, see Loizou, 1998, 1999). While the temporal and amplitude information is preserved, this type of signal processing algorithm provides only the coarse spectral information of the signal. It has been shown that high levels of speech perception can be achieved with only limited spectral information for both cochlear implant listeners (Garnham et al., 2002; Osberger et al., 2000; Skinner et al., 1994) and normal-hearing listeners listening to a simulation of the cochlear implant (Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995). However, such processing strategies may not be adequate for music appreciation in cochlear implant listeners.

Regardless of their inability to receive spectral information for pitch perception, some implant listeners did express that they enjoyed listening to music and reported having the ability to follow the rhythm of the melody. Schulz and Kerber (1994) investigated the music perception abilities in a group of Med-El implant listeners and concluded that cochlear implant listeners could recognize songs with rhythmic structure better than songs without rhythmic structure. The notion that music perception in electric hearing is dominated by the rhythmic patterns rather than pitch/tonal information was further demonstrated in a series of studies conducted by Gfeller et al. (1991, 1997, 2002). Gfeller et al. (1997) used the adapted *Primary Measures of Music Audiation* test battery to measure Nucleus 22 cochlear implant listeners' ability to discriminate simple rhythmic patterns and pitch patterns via their speech processors using the F0F1F2 and SPEAK strategies. The music task included identification of 20 pairs of short, sequential pitch patterns, with each pattern consisting of two to five notes ranging in pitch from C3 (130.81 Hz) to F4 (349.23 Hz). For the rhythmic pattern task, 14 pairs of short rhythm patterns with differing note durations were used with all the notes presented at the same frequency (C4–261.63 Hz). They found that cochlear implant listeners performed significantly better with simple rhythmic patterns as opposed to sequential pitch patterns. They also performed significantly poorer on pitch pattern discrimination tasks than normal-hearing listeners. On the other hand, no significant difference in performance was found between normal-hearing and cochlear implant listeners on rhythmic tasks.

In contrast to most studies in cochlear implants that focused on temporal processing of simple stimuli, such as detection of gaps and amplitude modulations, Collins, Wakefield, and Feinman (1994) and Gfeller et al. (1997) examined cochlear implant listeners' ability to process complex temporal patterns. Collins et al. presented stimuli with different temporal patterns by direct electric stimulation to a single electrode. Stimuli consisted of sequences of twelve 35-ms pulse bursts at 1000 pulse/sec, separated by variable gaps. Subjects were asked if two sequences of pulse bursts were the same or different. Contrary to the normal temporal processing with simple temporal patterns, they observed large inter-

TABLE 1. Information on nine cochlear implant listeners

Subject	Age	Mus.	Age at Onset	Dur. Deaf	Etiology	Yr Exp	Device	Strategy	Vowel (%)	Sentence (%)
C1	40	0	10	14	Meningitis	16	Ineraid*	CIS	DNT	DNT
C2	57	0	44	5	Sudden	11	Nuc 22	SPEAK	60	82
C3	44	0	35	0	Trauma	9	Nuc 22	SPEAK	79	92
C4	69	<2	34	26	Unknown	2	Clarion I	MPS	DNT	DNT
C5	70	0	29	7	Menieres	24	Ineraid†	CIS	DNT	DNT
C6	70	>10	40	2	Unknown	3	Clarion I	CIS	54	16
C7	69	0	30	5	Otosclerosis	13	Nuc 22	SPEAK	69	78
C8	65	>35	55	1	Unknown	2	Clarion I	MPS	67	86
C9	64	0	42	6	Menieres	4	Clarion I	SAS	75	DNT

Mus: years of formal musical training; Age at Onset: age at onset of hearing loss; Dur. Deaf: duration of profound hearing loss prior to implantation; Yr Exp: years of experience with the implant; Strategy: speech processing strategy used during the experiments; Vowel: score (% correct) on vowel recognition in /hVd/ context; Sentence: score (% correct) on sentence recognition with IEEE sentences; CIS: Continuous Interleaved Sampler; MPS: Multiple Pulsatile Sampler; DNT: did not test.

** Subject C1 was implanted with the Ineraid device and used the Med-El processor.*

† Subject C5 was implanted with the Ineraid device and used the Geneva processor.

subject differences in cochlear implant listeners' ability to discriminate between complex temporal patterns. Among the seven subjects tested, three performed similarly to the normal-hearing range reported in Sorkin (1990), two performed poorer than the normal-hearing listeners, and the remaining two performed close to the chance level. Similarly, Gfeller et al. found that cochlear implant listeners performed significantly poorer than normal-hearing listeners in the six-pulse task, where the four different rhythmic patterns contained a train of six pulses separated by either long or short interpulse intervals. The reason for the cochlear implant listeners' poor performance in these complex temporal tasks is not clear, but Gfeller et al. suggested that it may be related to the implant listeners' cognitive capacities, such as deficits in central temporal processing, rather than peripheral processing. Recent studies on information processing capacities in cochlear implant children by Pisoni and his colleagues (Pisoni, 2000; Pisoni & Cleary, 2003) suggested working memory may play a role in speech recognition performance in cochlear implant listeners.

Previous studies have shown that both rhythmic and pitch cues contributed to melody recognition (Andrews & Dowling, 1991; Dowling & Bartlett, 1981). The inability of cochlear implant listeners to follow rhythmic patterns may impair their ability to recognize melodies and enjoy music. Here we propose three experiments to evaluate music perception with temporal cues in general, and the relationship between rhythmic and melody recognition in particular, in normal-hearing and cochlear implant listeners. The first two experiments were designed to evaluate normal-hearing and cochlear implant listeners' ability to discriminate and identify the rhythmic features that are found across all musical genres (i.e., tempo discrimination and rhythmic pattern identification tasks). The third experiment was designed to systematically measure the relative contributions of temporal and spectral cues to melody recognition in normal-hearing and cochlear implant listeners. We hypothesized that cochlear implant listeners would perform similarly to normal-hearing listeners in the low-frequency tempo discrimination task, but poorer than the normal-hearing listeners in the rhythmic pattern discrimination task based on the previous results reported in the complex temporal tasks by Gfeller et al. (1997). We further hypothesized that temporal envelope cues alone could not support melody recognition without rhythmic cues.

EXPERIMENT I—TEMPO DISCRIMINATION

Methods

Subjects • Four normal-hearing (N1–N4) and five cochlear implant listeners (C1-C5) participated in the tempo discrimination task. Pure tone audiometry was performed in normal-hearing listeners before the experiment to ensure that their hearing thresholds were within the normal range ($\leq$ 20 dB HL) from 125 to 8000 Hz. Three out of the four normal-hearing listeners (N1, N3, N4) were trained musicians, whereas only one cochlear implant listener (C4) had received musical training prior to the experiment.

Table 1 shows additional information regarding hearing history, implant device, and speech performance for the cochlear implant listeners in all three experiments in the present study. Speech recognition scores for some subjects (C1, C4, C5, and C9) were not tested due to time constraints. All cochlear-implant users were postlingually deafened adults with at least 2 yr of implant use at the time of testing. It should be noted that there was a large age difference between the normal and implant listeners in the three experiments presented in this paper. The average age of normal-hearing listeners was 20 yr (range, 16–35 yr) and the implant listeners was 61 yr (range, 40–70 yr). The difference in age may be a factor for differences in performance in our tasks. Studies have shown that limited memory capacity and slower processing speed may have detrimental effects on complex stimulus processing tasks (Gordon-Salant & Fitzgibbons, 1997; Wingfield, 1996).

Stimuli • Musical tempos were generated using an Alesis SR-16 drum machine. Stimuli contained a pattern of beats, a single bar in duration, performed in a 4/4 time signature that indicates four beats per measure with one quarter note equaling a beat. Four standard tempos were created at 60, 80, 100, and 120 beats per minute (bpm). Stimuli were then converted into .wav files using sound recording and editing software (Cool Edit Pro, version 1.2).

Procedure • Tempo discrimination was measured with four standard tempos at 60, 80, 100, and 120 bpm. For each standard tempo, 21 or more tempos were used to pair with the standard tempo. For instance, tempos ranging from 60 to 80 bpm were paired with the standard tempo of 60 bpm. Subjects listened to these tempo pairs and were asked to identify the faster tempo in a two-interval, forced-choice paradigm. Each pair was presented once in a random order within a block. To obtain a measure of the subject's response bias, one of the tempo pairs contained the same standard tempo. A total of 20 blocks were tested to obtain percent correct scores for each tempo pair. All subjects listened to the stimuli in the sound field in a sound attenuated, IAC double-walled booth at a comfortable listening level (55–65 dB SPL). Cochlear-implant listeners listened through their speech processors at their preferred settings.

A three-parameter sigmoidal function was used to fit the data,

$$P(C) = \left(\frac{S - 50}{1 + e^{\frac{-2(t-\theta)}{\beta}}}\right) + 50 \qquad (1)$$

where S is the saturation score, θ and β are related to the threshold and its slope, respectively. Tempo discrimination threshold (T) was calculated as the bpm that produced 75% correct score:

$$T = \theta - \left(\frac{\beta}{2}\right) * \ln\left(\frac{S - 75}{25}\right) \qquad (2)$$

Results

Figure 1 shows the mean tempo discrimination performance from both normal-hearing and cochlear implant listeners (the error bars are equal to one standard error of the mean). Each panel represents a different standard tempo condition at 60, 80, 100, and 120 bpm, respectively. A sigmoidal function was fitted to the individual data and the mean data for each tempo condition, for both normal-hearing and cochlear implant listeners. Table 2 displays the mean values and the standard error of the function parameters (S, θ, β) and the mean tempo discrimination thresholds (T).

The sigmoidal function provided excellent fits to the data, accounting for 71–96% of the variance in the individual data and 92–96% of the variance in the mean data. The mean thresholds for normal-hearing listeners were 62.93, 83.40, 103.50, and 124.18 bpm, and for cochlear implant listeners were 63.76, 84.40, 106.03, and 125.33 bpm for tempo conditions of 60, 80, 100, and 120 bpm, respectively. Although the tempo discrimination thresholds in cochlear implant listeners were slightly higher than that in normal-hearing listeners, their differences were not statistically significant in any tempo conditions (for standard condition of 60 bpm: $t_{(7)} = -1.6$, $p > 0.05$; 80 bpm: $t_{(7)} = -1.42$, $p > 0.05$; 100 bpm: $t_{(7)} = -2.27$, $p > 0.05$; 120 bpm: $t_{(7)} = -2.00$, $p > 0.05$).

EXPERIMENT II—RHYTHMIC PATTERN IDENTIFICATION

Methods

Subjects • The same four normal-hearing listeners (N1-N4) that participated in Experiment I also participated in the rhythmic pattern identification task. Three cochlear implant listeners (C2, C3, and C6) were tested in this task. Subjects C2 and C3 also participated in Experiment I. Among the three implant listeners, only C6 had received musical training prior to our experiment. Table 1 provides additional information about the cochlear implant listeners.

Stimuli • All musical patterns were played on and recorded by an Alesis SR-16 drum machine. The recorded patterns were quantized to remove any timing errors such as delay or rushing, and then were converted into .wav files using Cool Edit Pro, version 1.2.

Seven one-bar rhythmic patterns, created using permutations of quarter, eighth, and sixteenth notes, were used in this experiment. Figure 2 shows the original standard rhythmic pattern used in the tempo discrimination task (containing only quarter notes) and the additional six rhythmic patterns. All patterns were in a 4/4 time signature. In these patterns, beats one, three, and four always contained the same quarter note, but beat two was varied to contain one of the seven possible patterns. These patterns were played at four standard tempos: 60, 90, 120, and 150 bpm.

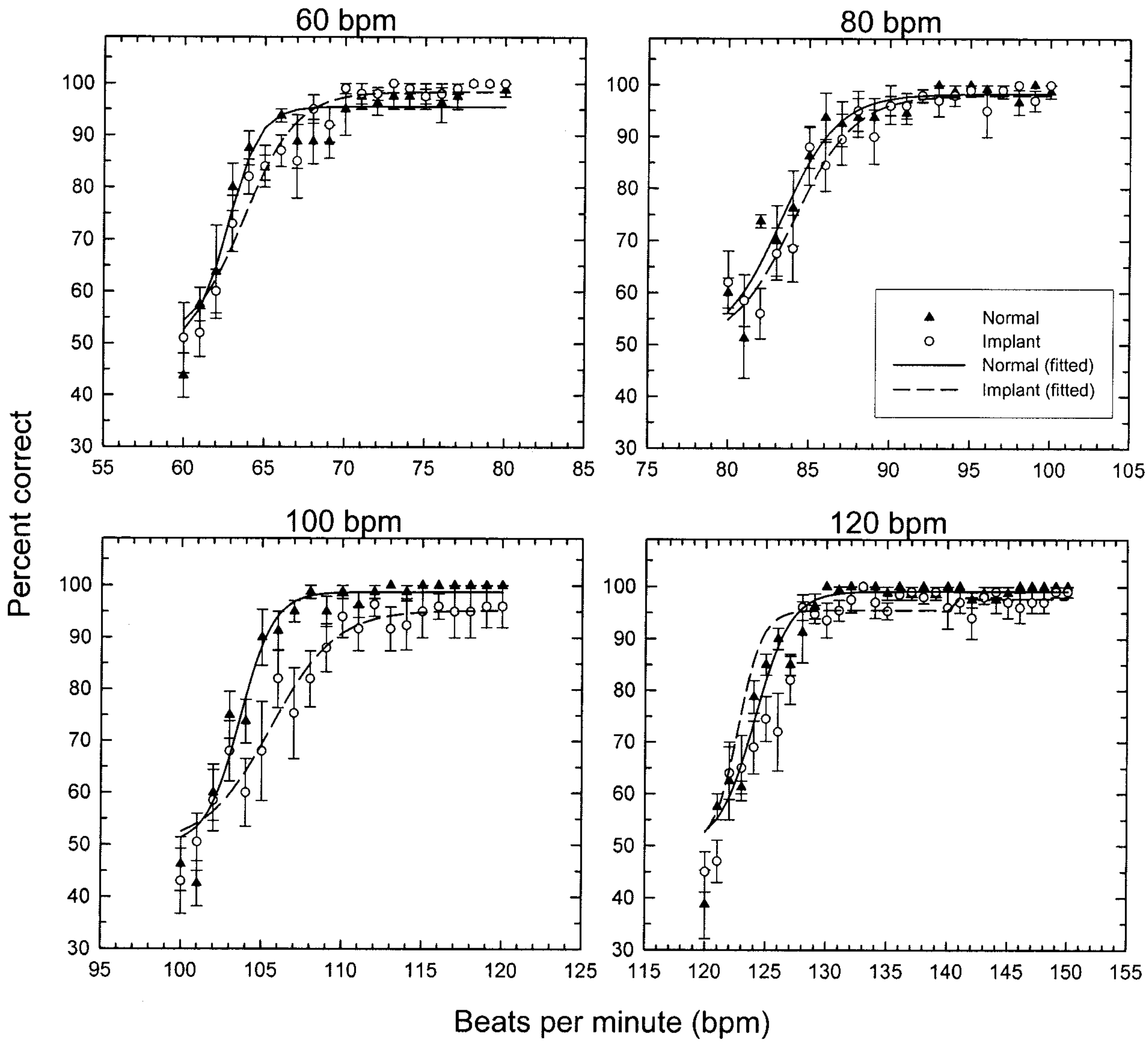

Figure 1. Mean tempo discrimination scores from normal-hearing (filled triangles) and cochlear implant (open circles) listeners. The four panels represent the four standard tempo conditions of 60, 80, 100, and 120 bpm. A sigmoidal function was fitted to the data, in which the solid and dashed lines represent the fitted functions for the normal-hearing and cochlear implant data, respectively. The vertical error bars indicate the standard error of the mean.

Procedure • Rhythmic pattern identification was measured at four standard tempos: 60, 90, 120, and 150 bpm. Each subject listened to two bars of drumbeats. The first bar was always the standard rhythmic pattern (the middle trace in Fig. 2). The second bar contained one of the seven patterns displayed in Figure 2. Subjects were asked to choose the musical notation corresponding to the rhythmic pattern they heard. We trained all of our subjects to read basic musical notation until their performance reached

TABLE 2. Mean values and standard error of the psychometric function parameters (*S*, *θ*, *β*) and the tempo discrimination thresholds (*T*) in normal-hearing and cochlear implant listeners

	S (%)		*θ* (bpm)		*β* (bpm)		*T* (bpm)	
Condition	NH	CI	NH	CI	NH	CI	NH	CI
60 bpm	95.50 ± 1.18	98.36 ± 1.02	62.73 ± 0.25	63.74 ± 0.24	1.97 ± 0.44	3.26 ± 0.44	62.93 ± 0.29	63.76 ± 0.40
80 bpm	98.31 ± 1.08	98.03 ± 1.16	83.29 ± 0.27	84.07 ± 0.28	3.51 ± 0.50	3.74 ± 0.52	83.40 ± 0.51	84.40 ± 0.48
100 bpm	98.75 ± 1.10	95.24 ± 1.50	103.7 ± 0.21	105.6 ± 0.37	2.07 ± 0.38	3.98 ± 0.65	103.50 ± 0.30	106.03 ± 0.95
120 bpm	99.11 ± 0.80	97.69 ± 0.80	124.1 ± 0.23	125.1 ± 0.26	2.93 ± 0.42	3.59 ± 0.46	124.18 ± 0.33	125.33 ± 0.44

Bpm: beats per minute; NH: normal hearing; CI: cochlear implant.

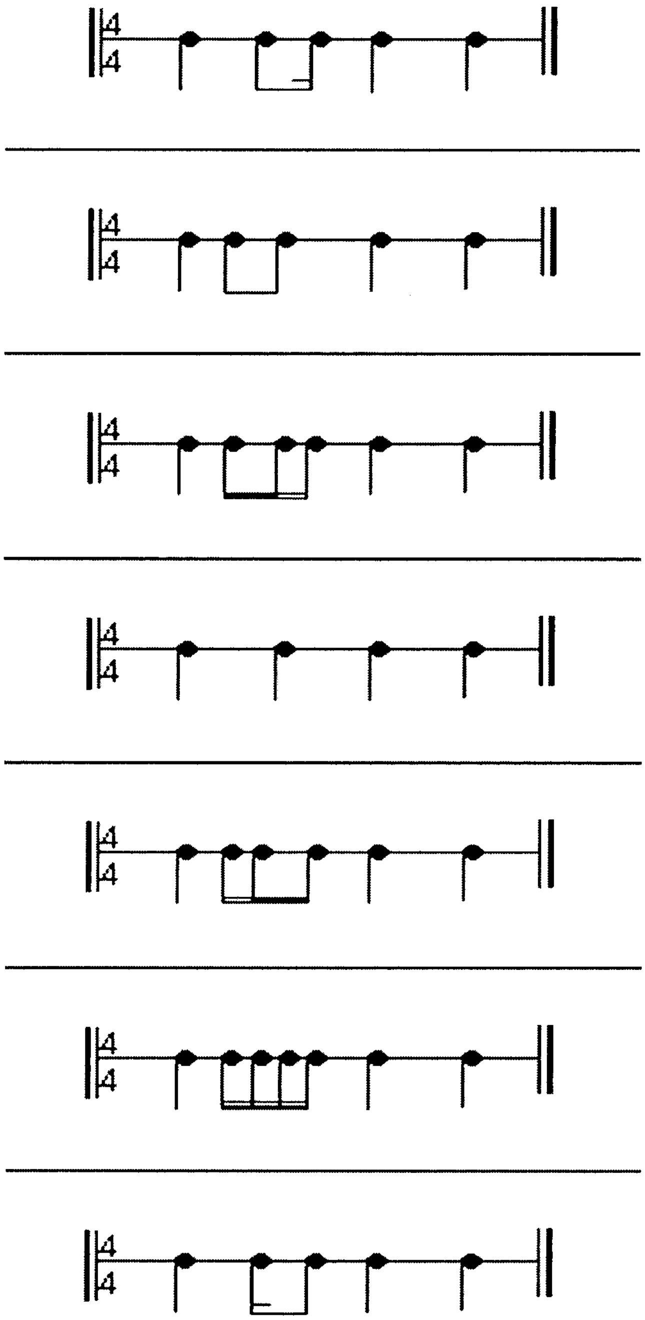

Figure 2. Seven different patterns created in the pattern identification task. The first bar of the pattern always contained four quarter-notes (four sounds of equal duration). The second bar was chosen from the options above. The first, third, and fourth beats of the second bars were always the same (quarter notes). The second beat varied with permutations using eighth and sixteenth notes.

85% correct in the training condition. Most of the subjects were tested over the course of 5 mo or less, depending on their familiarity with musical concepts. Feedback was given both by interactive lessons with the investigator (a professional musician) and the software during the training sessions. During the actual test, no feedback was given.

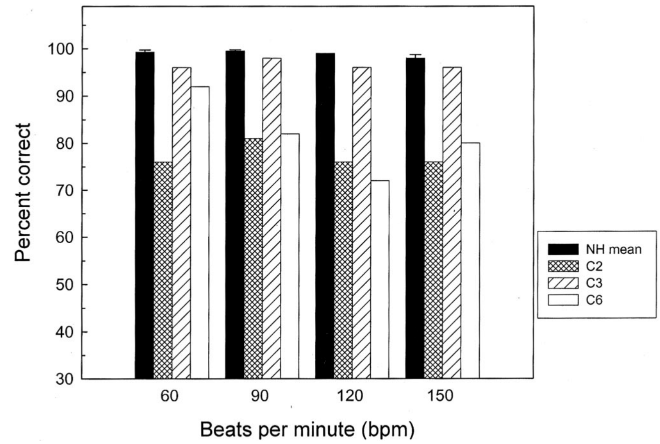

Figure 3. Mean rhythmic pattern identification scores from normal-hearing listeners and individual data from three cochlear implant listeners (C2, C3, and C6) at different tempo conditions (60, 90, 120, and 150 bpm).

Results

Figure 3 shows the mean rhythmic pattern identification scores from four normal-hearing listeners (solid bar) and the individual data from three cochlear implant listeners. The normal-hearing subjects achieved nearly perfect performance (mean scores = 99.3, 99.5, 99, and 98% for 60, 90, 120, and 150 bpm, respectively; standard error <0.3%). On the other hand, only one of the three cochlear implant subjects (C3) achieved a performance similar to the normal-hearing listeners. The other two subjects (C2 and C6) achieved scores of 10–25 percentage points lower than the normal score.

A one-way repeated measure analysis of variance, using the Greenhouse-Geisser adjustment, revealed no significant difference in performance among the four tempo conditions in both normal-hearing [$F_{(3,9)} = 3.41$, $p > 0.05$] and cochlear implant [$F_{(3,6)} = 1.15$, $p > 0.05$] listeners. This indicates that listeners' ability to identify rhythmic patterns was not affected by the tempo.

EXPERIMENT III—MELODY IDENTIFICATION

Methods

Subjects • Six additional normal-hearing (N5–N10) and six cochlear implant listeners (C2, C3, C6, C7, C8, and C9) participated in the melody identification task. Two normal-hearing listeners (N5 and N10) had more than 12 yr of formal piano training, and two cochlear implant listeners (C6 and C8) had received prior musical training. Table 1 provides additional information about the cochlear implant listeners.

Stimuli • Two sets of 12 familiar songs, played by single notes (monophonic melodies), were generated using a software synthesizer (ReBirth RB-338, version 2.0.1). One set contained both rhythmic and melodic information (rhythm condition), whereas the other set contained only melodic information (no-rhythm condition). In the no-rhythm condition, all melodies were played using notes of the same duration (quarter notes with 350 msec in duration) with a silent period of 150 msec between notes. Therefore, pitch was the only available cue for melody identification in the no-rhythm condition. Each song consisted of 12–14 notes of its initial phrase. All songs were played within a frequency range from 207 (G#3) to 523 Hz (C5). The largest range was 16 semitones between the highest and the lowest notes of the melody, while the smallest range was 7 semitones. Detailed information about each song is listed in the Appendix.

The choice of the 12 songs used in this experiment was based on a melody recognition study performed on a group of normal-hearing listeners. A list of 30 melodies with and without rhythmic cues was presented to 15 normal-hearing listeners ranging from 20 to 40 yr of age. They were asked to write the title of the songs they heard in an open-set recognition task. The songs that received 90% or greater recognition were then used in this experiment. All subjects participating in this experiment, except subject C9, reported that they had heard and remembered the 12 melodies. Subject C9 was not raised in the United States and had limited exposure to the melodies. He reported only five songs on the list were familiar to him.

The original melody stimuli were also processed to contain either temporal envelope cues or coarse spectral cues via a cochlear implant simulation program (Shannon et al., 1995). The original broadband (80–8800 Hz) stimuli were divided into 1, 2, 4, 8, 16, 32, and 64 frequency bands. The cutoff frequencies of each band were determined by approximating equal cochlear distance for each band according to the Greenwood map (1990). For instance, cutoff frequencies for the 8-band simulation were 80, 220, 440, 780, 1300, 2100, 3500, 5500, and 8800 Hz. The temporal envelope was then extracted from each analysis band by full-wave rectification and low-pass filtering at 500 Hz. This envelope signal was then used to modulate a white noise, which was then bandpass filtered with the same analysis filters as were used on the original signal. Stimuli were re-synthesized by summing these envelope-modulated narrow-band signals.

Procedure • Both normal-hearing and cochlear implant listeners were first tested with the original, unprocessed stimuli. In addition, normal-hearing listeners were tested with all the rhythm and no-rhythm processed stimuli (simulation). Cochlear-implant listeners, however, were only tested using 1-band with rhythm processed stimuli. The condition of 1-band without rhythm was not performed in the cochlear implant listeners because their performance was already at or near the chance level when presented with original stimuli.

Normal-hearing listeners were tested in a double-walled, soundproof booth. Stimuli were presented monaurally through headphones at the listeners' most comfortable levels (65–70 dB SPL), which was about 5 dB higher than the tempo discrimination and rhythmic pattern identification experiments. A higher presentational level was needed to achieve subjects' most comfortable level with the melody stimuli. Cochlear-implant listeners listened to the stimuli through their speech processors using a direct connection, connecting from the output of the sound card of a computer to the auxiliary electric input in their speech processors. They were advised to adjust the volume and the sensitivity dial to the place where the stimuli were perceived at their most comfortable level. The titles of the 12 melodies were displayed on a computer screen and the subjects were asked to identify the melody from this closed set. Both normal-hearing and cochlear implant listeners were given 5–10 minutes to become familiar with the songs by listening to the original melodies with and without rhythmic cues before actual data collection. Visual feedback was given during this practice session. All melodies were presented three times in random order for each experimental condition. Repetition of the stimulus was not allowed and visual feedback was given immediately after the subject's response.

Results

Results show differences in performance between normal-hearing and cochlear implant listeners in the melody identification task with and without rhythmic cues. The mean chance level performance for this task was about 8%. While the normal-hearing listeners similarly achieved near perfect performance for the rhythm (98.3%) and no-rhythm (97.5%) conditions, the cochlear implant listeners performed less accurately than the normal-hearing listeners in the rhythm condition [63.2%, $t_{(10)} = 6.97$, $p < 0.001$] and essentially at chance [11.7%, $t_{(10)} = 23.84$, $p < 0.001$] in the no-rhythm conditions.

Figure 4 shows the individual data from the six cochlear implant subjects. In the rhythm condition, scores ranged from 45 to 78% with a 5% standard error. In the no-rhythm condition, all subjects performed near chance (~8%) or within the 95% chance range (~17%) calculated using the binominal distribution equation, except for subject C9, who achieved

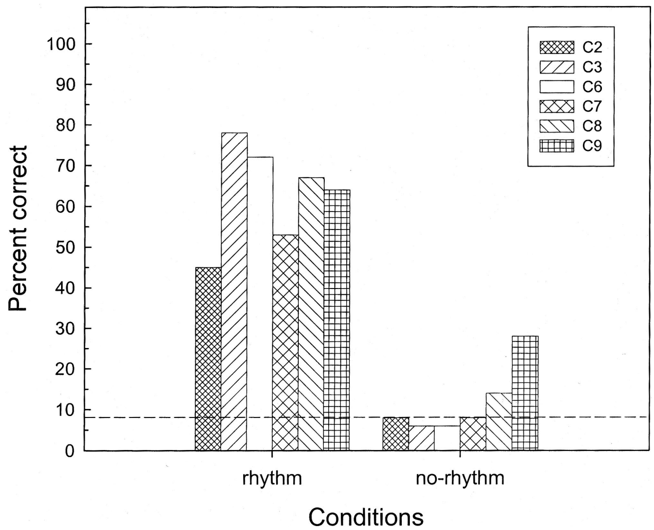

Figure 4. Melody identification scores from individual cochlear implant listeners with the original melodies. The horizontal dashed line indicates the mean chance performance. The vertical bars represent different subjects in each condition.

28% correct. Note that the probability of achieving 28% by chance would be less than 0.5%.

Figure 5 shows the results of the 1-band conditions in normal-hearing and cochlear implant listeners, along with the results from the original conditions for comparison. On average, the normal-hearing listeners achieved an 80.0% correct score in the 1-band with rhythm condition, which was significantly better than the 52.5% correct score achieved by the cochlear implant listeners in the same condition [$t_{(10)} = 3.310$, $p < 0.05$]. For comparison, the normal-hearing listeners' performance in the 1-band with rhythm condition was significantly poorer (17.8 percentage points) than in the original with rhythm condition [$F_{(1,5)} = 13.35$, $p < 0.05$], but the cochlear implant listeners' performance between the two conditions was not significantly different [$F_{(1,5)} = 2.86$, $p > 0.05$]. In addition, there was no difference between the cochlear implant listeners' scores (11.7%) in the original no-rhythm condition and the normal-hearing listeners' scores (12.5%) in the 1-band no-rhythm condition [$t_{(10)} = 0.183$, $p > 0.05$].

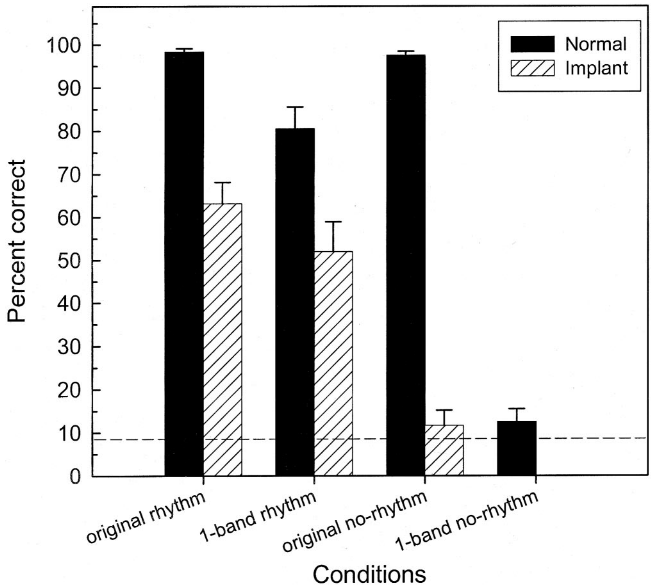

Figure 5. Mean scores from normal-hearing and cochlear-implant listeners in four melody identification conditions. The horizontal dashed line indicates the mean chance performance and the vertical error bars represent the standard error of the mean.

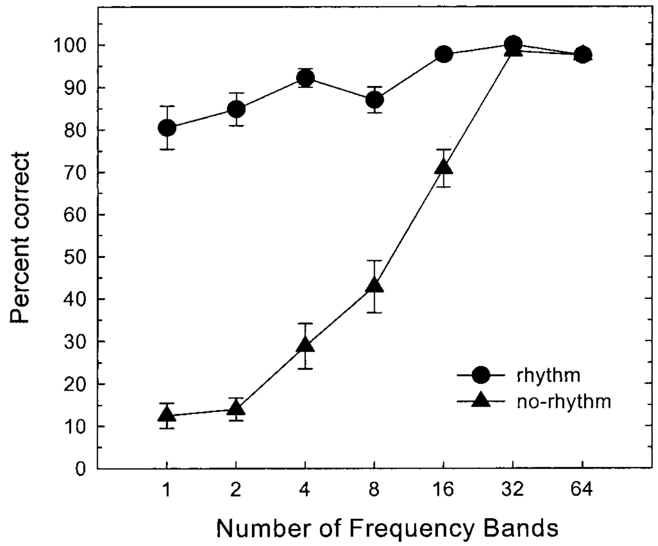

Figure 6. Mean melody identification scores as a function of number of frequency bands obtained from normal-hearing listeners with amplitude modulated stimuli in both rhythmic conditions. The solid circles represent the rhythm condition, whereas the solid triangles are for the no-rhythm condition. The vertical error bars indicate the standard error of the mean.

Figure 6 shows the averaged normal-hearing listeners' scores as a function of the number of bands in the rhythm (circles) and no-rhythm (triangles) conditions. Note the relatively flat performance function in the rhythm condition and the steeply increasing performance function in the no-rhythm condition. The performance reached the 90% correct level with only four bands in the rhythm condition, but not until 32 bands in the no-rhythm condition. A two-way repeated measures analysis of variance indicated significant band [$F_{(6,30)} = 95.64$, $p < 0.001$] and rhythm [$F_{(1,5)} = 909.18$, $p < 0.001$] main effects, as well as a significant interaction between the two variables [$F_{(6,30)} = 63.34$, $p < 0.001$]. A *t*-test revealed that the interaction was due to a significant difference in performance between the rhythm and no-rhythm conditions for the number of bands less than or equal to 16, but no significant differences for the 32- or 64-band conditions.

Finally, an analysis of error patterns was conducted on cochlear implant performance with original melodies and normal-hearing performance with 8-band simulation with and without rhythmic cues. Without rhythmic cues neither cochlear implant nor normal-hearing listeners showed systematic error patterns. When collapsed across subjects, all melodies appeared to be equally difficult for cochlear implant listeners, but not for normal-hearing listeners with the 8-band simulation. Melodies with large intervals (e.g., wider than a sixth) and melodies with starting notes substantially lower than other melodies were identified with greater accuracy (i.e., "Lullaby, and Good Night," "Take Me Out to the Ball Game," and "Auld Lang Syne"). With rhythmic cues, both normal-hearing and cochlear-implant listeners shared similar error patterns. Melodies with similar rhythmic patterns were confused more frequently (e.g., "London Bridge is Falling Down" versus "Mary Had a Little Lamb," and "Old MacDonald Had a Farm" versus "Twinkle, Twinkle, Little Star") than melodies with very different rhythmic structure from other melodies in the list (e.g., "Lullaby, and Good Night," "Take Me Out to the Ball Game," and "Auld Lang Syne").

DISCUSSION

Temporal Processing in Cochlear Implants

Previous studies reported relatively normal temporal processing ability with simple stimuli in cochlear implant listeners. Shannon (1989, 1992) showed normal gap detection and modulation detection thresholds in cochlear implant listeners, with a 2–4 msec gap and 5–10% modulation detection thresholds measured at the most comfortable level. Nelson and Donaldson (2002) measured recovery from forward masking in a group of 21 Nucleus and Clarion users and found that most of their subjects (18/21) had normal forward masking recovery functions with an average time constant of 54 msec.

In the present tempo discrimination experiment, the cochlear implant listeners also yielded performance comparable to the normal-hearing listeners. The tempo discrimination test can be viewed as a rate discrimination test at extremely low rates, i.e., in the 1–2 Hz range corresponding to 60–120 bpm conditions. Previous studies have reported that cochlear implant listeners had rate discrimination limens that were 10–100 orders of magnitude higher than normal-hearing listeners at rates of 50 Hz or above (e.g., Zeng, 2002). Unlike rate discrimination at the higher rates, the present results showed that cochlear implant listeners had normal rate discrimination at very low rates.

Cochlear-implant listeners appeared to have difficulty performing complex temporal processing tasks. In the present rhythmic pattern identification task, two out of the three cochlear implant listeners performed 10–25 percentage points poorer than normal-hearing listeners. Furthermore, in the 1-band with rhythm condition where only the temporal cues were present, the cochlear implant listeners' melody recognition was about 30 percentage points less than the normal-hearing listeners (Fig. 5). These data were in agreement with previous findings reporting both less accurate temporal pattern discrimination performance and a large inter-subject variability in cochlear implant listeners (Collins et al., 1994; Gfeller et al., 1997). The discrepancy between normal tempo discrimination and abnormal rhythmic pattern identification performance in cochlear-implant listeners can be attributed to differences in cognitive function. The simple temporal detection and discrimination tasks require sensory processing, whereas the complex temporal pattern identification tasks demand cognitive processes. Degraded and distorted sensory inputs in electric hearing may demand a high memory load to process and encode the incoming signals in cochlear implant listeners. Gfeller et al. (1997) suggested that abnormal performance for cochlear implant listeners may be attributed to "deficits in central temporal processing." Pisoni and his colleagues (Pisoni, 2000; Pisoni & Cleary, 2003) also noted that measures of information processing capacities such as digit span recalls are correlated with measures in speech performance. Another possibility, is that in the present study the cochlear implant listeners were, on average, 41 yr older than the normal-hearing listeners. The performance difference might be due, at least partially, to the age difference between the two groups (Fitzgibbons & Gordon-Salant, 2001; Gordon-Salant & Fitzgibbons, 1999). Gordon-Salant & Fitzgibbons (1999) reported that older listeners (age ranged 65–76 yr) performed significantly poorer than younger (age ranged 18–40 yr) listeners on temporally mediated measures, particularly in more complex stimulus conditions. The age range of their younger and older listeners was similar to our subjects. Gordon-Salant & Fitzgibbons (1997) further suggested that the differences in temporal processing and speech understanding between the two age groups were influenced by their working memory capacity.

Rhythmic Pattern Identification and Speech Perception

We obtained IEEE sentence recognition scores in quiet in all but one (the non-native speaker, C9) cochlear implant listeners who participated in the melody identification task. No reliable relationship was found between the sentence recognition performance and the melody identification performance with only rhythmic cues (1-band with rhythm condition). Subject C6's performance in melody identification was comparable to other subjects, but her sentence recognition score was extremely low (16%). The other four subjects achieved between 80–90% correct sentence scores despite the large variability in their melody recognition scores, which ranged from 40–80% correct. This result is inconsistent with Collins et al. (1994), who suggested that temporal pattern tasks can be an "accurate and efficient predictor" of speech recognition ability in cochlear-implant subjects. This inconsistency might be due to the relatively small sample size ($N = 5$), as well as the ceiling effects in speech recognition in four out of five subjects in the present study.

Music Perception and Music Experience

We found no relationship between melody identification and music experience in the present study. In the original with rhythm condition, both the lowest score (C2) and the highest score (C3) were obtained from subjects without any music experience (Fig. 4). Similarly, the only subject (C9) who performed significantly above the chance level in the no-rhythm condition (Fig. 4) had no music experience, and limited exposure to the melodies used in this study since he was not a native English speaker. The melodies that were used in this study are heard by children and adults in everyday life, and are therefore familiar to persons who have had little or no musical training. Thus, the extent of musical training was not an important factor in the interpretation of these data. The results of this study are also consistent with prior research (Gfeller et al., 2000) indicating that formal music training is not a particularly strong predictor of perceptual accuracy for melody recognition by implant listeners.

Music Perception with Temporal Cues

We have provided clear evidence showing that cochlear implant listeners rely mostly on the rhythmic cues for melody recognition (Figs. 4 and 5). On average, the cochlear implant listeners could recognize about 60% of the melodies with the rhythmic cues, but dropped to the chance level without the rhythmic cues. More convincingly, cochlear implant listeners produced relatively unchanged performance (50%) in the 1-band condition where only the rhythmic cues were preserved.

In addition, we observed a positive relationship between rhythmic pattern identification and melody identification performance with rhythmic cues in the three implant listeners (C2, C3, and C6) who participated in both tasks. Among the three implant listeners, subject C2 performed the worst in the rhythmic pattern identification task and also produced the lowest score in the melody identification task with original stimuli and rhythmic cues. Subject C3, on the other hand, achieved the highest scores in both tasks compared with the other two subjects. Despite the small sample size, this observation further enhances the suggestion that the cochlear implant listeners almost exclusively used the rhythmic cues to perform the melody identification task. A third piece of evidence for this suggestion comes from the absence of consistent error patterns in the no-rhythm condition but the presence of such patterns in the rhythm condition.

Music Perception with Spectral Cues

We demonstrated that fine spectral resolution with as many as 32 bands is necessary for melody identification without rhythmic cues. Similar results were reported previously by Burns, Lineaweaver, & Sanborn (Reference Note 1). Unlike speech recognition that only requires approximately four bands of coarse spectral cues to reach nearly perfect performance in quiet (Shannon et al., 1995), these simulation results suggest that the cochlear implant listeners need at least 16 or more functional spectral channels to recognize melodies without rhythmic cues.

Unfortunately, the present cochlear implants simply cannot provide 16 functional channels. Despite the 16 or more electrodes present in all modern multi-electrode cochlear implants, their functional numbers for speech recognition in quiet is at best between 4 and 8 (Dorman & Loizou, 1997; Fishman, Shannon, & Slattery, 1997; Friesen, Shannon, Baskent, & Wang, 2001; Garnham et al., 2002). If the same 4–8 functional channels are also available for melody recognition, cochlear implant listeners would perform 30–40% correct in the no-rhythm condition (Fig. 6). However, only the chance level performance was achieved by the six cochlear implant listeners in this study and most others in previous studies (Gfeller & Lansing, 1991; Gfeller et al., 2002; Gfeller et al., 1997; Pijl, 1997). In other words, multi-electrode cochlear implant listeners performed as if they were listening through a single-channel device in melody recognition. We should note that the six subjects in the present study had relatively high levels of vowel recognition ranging from 54% to 79% correct, similar to the normal-hearing listeners' 4-band vowel recognition performance in acoustic simulations (Shannon et al., 1995). These results suggest that while cochlear-implant listeners were able to learn to extract distorted spectral cues for speech recognition but not for melody recognition.

To accurately perceive melodies, both absolute and relative pitch cues are needed but are likely not available in cochlear implants. First, the frequency-to-electrode map is not aligned along the absolute pitch dimension in cochlear implants. The fundamental frequency of the musical notes in the present study ranged from 207 to 523 Hz, but the present cochlear implants are not inserted deep enough to access these low-frequency neurons. In addition, there is most likely a mismatch between the place of the electrodes and the characteristic frequency of the neurons at that location. Second, the mismatch of frequency to electrode was further intensified due to the problems of non-tonotopic trajectory of the electrode array (Skinner et al., 2002). Using computed tomography, Skinner et al. reported that a number of patients in their study had kinks, bends, and compressions in their implanted arrays, which may have adversely affected the encoding of relative pitch. Although such a mismatch may decrease speech recognition performance (Fu & Shannon, 1999; Shannon, Zeng, & Wygonski, 1998), it completely prevents the cochlear implant listeners from using place pitch in melody recognition. Finally, we note that the only cochlear implant subject (C9) who performed above the chance level in the no-rhythm condition was an SAS user. He achieved the above-chance-level performance based on the temporal pitch that was likely preserved and present in the analog waveform of the filter's output. Such temporal pitch was unlikely to be present at the envelope level in other implant listeners due to the low carrier rate (<1000 Hz) and the low envelope cutoff frequencies (<500 Hz).

SUMMARY

The present study systematically compared performance between cochlear implant and normal-hearing listeners in tempo discrimination, rhythmic pattern identification, and melody recognition. Consistent with previous studies, the present study found that the cochlear implant listeners performed normally in tempo discrimination but significantly poorer in rhythmic identification and melody recognition. Except for one subject who used the SAS strategy, cochlear implant listeners relied solely on rhythmic cues for melody recognition. The cochlear implant simulation conducted as a function of the number of frequency bands further demonstrated that both temporal (rhythmic) and spectral (pitch) cues contribute to melody recognition. Without rhythmic cues, high spectral resolution with as many as 32 bands is needed for melody recognition. While the present cochlear implants provide sufficient spectral cues to support speech recognition in quiet, they are not adequate for music perception. Increasing the number of functional channels and better encoding of the fine structure information (Smith, Delgutte, & Oxenham, 2002) are needed to improve music perception in cochlear implant listeners.

ACKNOWLEDGMENTS

We thank our cochlear implant and normal-hearing subjects for their time and dedication. The simulation program and computer programs for data collection were provided by Kaibao Nie. IEEE sentence scores were provided by Ginger Stickney, and vowel recognition scores were provided by Michael Vongphoe. This work was supported by National Institutes of Health Grant 2 R01 DC02267 (F.-G.Z.) and the research supplement award (Y.-Y.K.).

Address for correspondence: Ying-Yee Kong, M.A., and Fan-Gang Zeng, Ph.D., 364 Med Surge II, University of California, Irvine, Irvine, CA 92697. E-mail: ykong@uci.edu or fzeng@uci.edu.

Received April 4, 2003; accepted December 12, 2003

REFERENCES

Andrews, M. W., & Dowling, W. J. (1991). The development of perception of interleaved melodies and control of auditory attention. *Music Perception*, *8*, 349–368.

Collins, L. M., Wakefield, G. H., & Feinman, G. R. (1994). Temporal pattern discrimination and speech recognition under electrical stimulation. *Journal of the Acoustical Society of America*, *96*, 2731–2737.

Dorman, M., & Loizou, P. (1997). Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. *Journal of the Acoustical Society of America*, *102*, 2403–2411.

Dowling, W. J., & Bartlett, J. C. (1981). The importance of interval information in long-term memory for melodies. *Psychomusicology*, *1*, 30–49.

Eddington, D. K., Dobelle, W. H., Brackmann, D. E., Mladevosky, M. G., & Parkin, J. L. (1978). Auditory prosthesis research with multiple channel intracochlear stimulation in man. *The Annals of Otology, Rhinology & Laryngology*, *87(Suppl. 53)*, 1–39.

Fishman, K. E., Shannon, R. V., & Slattery, W. H. (1997). Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. *Journal of Speech, Language, and Hearing Research*, *40*, 1201–1215.

Fitzgibbons, P. J., & Gordon-Salant, S. (2001). Aging and temporal discrimination in auditory sequences. *Journal of the Acoustical Society of America*, *109*, 2955–2963.

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

Fu, Q-J, & Shannon, R. V. (1999). Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing. *Journal of the Acoustical Society of America*, *105*, 1889–1900.

Garnham, C., O'Driscoll, M., Ramsden, R., & Saeed, S. (2002). Speech understanding in noise with a Med-El COMBI 40+ cochlear implant using reduced channel sets. *Ear and Hearing*, *23*, 540–552.

Gfeller, K., & Lansing, C. R. (1991). Melodic, rhythmic, and timbral perception of adult cochlear implant users. *Journal of Speech and Hearing Research*, *34*, 916–920.

Gfeller, K., Christ, A., Knutson, J. F., Witt, S., Murray, K. T., & Tyler, R. S. (2000). Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients. *Journal of American Academy of Audiology*, *11*, 390–406.

Gfeller, K., Turner, C., Mehr, M., Woodworth, G., Fearn, R., Knutson, J. F., Witt, S., & Stordahl, J. (2002). Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. *Cochlear Implants International*, *3*, 29–53.

Gfeller, K., Woodworth, G., Robin, D. A., Witt, S., & Knutson, J. F. (1997). Perception of rhythmic and sequential pitch patterns by normally hearing adults and adult cochlear implant users. *Ear and Hearing*, *18*, 252–260.

Gordon-Salant, S., & Fitzgibbons, P. J. (1997). Selected cognitive factors and speech recognition performance among young and elderly listeners. *Journal of Speech Language and Hearing Research*, *40*, 423–431.

Gordon-Salant, S., & Fitzgibbons, P. J. (1999). Profile of auditory temporal processing in older listeners. *Journal of Speech Language and Hearing Research*, *42*, 300–311.

Greenwood, D. D. (1990). A cochlear frequency-position function for several species—29 years later. *Journal of the Acoustical Society of America*, *87*, 2592–2605.

Leal, M. C., Shin, Y. J., Laborde, M. L., Calmels, M. N., Verges, S., Lugardon, S., Andrieu, S., Deguine, O., & Fraysse, B. (2003). Music perception in adult cochlear implant recipients. *Acta Otolaryngology*, *123*, 826–835.

Loizou, P. C. (1998). Mimicking the human ear. *IEEE Signal Processing Magazine*, *15*, 101–130.

Loizou, P. C. (1999). Signal-processing techniques for cochlear implants. *IEEE Engineering in Medicine and Biology Magazine*, *18*, 34–46.

Moore, B. C. J., & Glasberg, B. R. (1988). Gap detection with sinusoids and noise in normal, impaired, and electrically stimulated ears. *Journal of the Acoustical Society of America*, *83*, 1093–1101.

Muller, C. (1981). Comparison of percepts found with cochlear implant devices. *Annals of the New York Academy of Sciences*, *405*, 412–420.

Nelson, D. A., & Donaldson, G. S. (2002). Psychophysical recovery from pulse-train forward masking in electric hearing. *Journal of the Acoustical Society of America*, *112*, 2932–2947.

Osberger, M. J., Fisher, L., & Kalberer, A. (2000). Speech perception results in adults implanted with the CLARION Multi-Strategy cochlear implant. *Advances in Otorhinolaryngology*, *57*, 421–424.

Pijl, S. (1997). Labeling of musical interval size by cochlear implant patients and normally hearing subjects. *Ear and Hearing*, *18*, 364–372.

Pijl, S., & Schwarz, D. W. F. (1995). Melody recognition and musical interval perception by deaf subjects with electrical pulse trains through single cochlear implant electrodes. *Journal of the Acoustical Society of America*, *98*, 886–895.

Pisoni, D. B. (2000). Cognitive factors and cochlear implants: Some thoughts on perception, learning, and memory in speech perception. *Ear and Hearing*, *21*, 70–78.

Pisoni, D. B. & Cleary, M. (2003). Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation. *Ear and Hearing*, *24(1 Suppl)*, 106S–120S.

Schultz, E., & Kerber, M. (1994). Music perception with the MED-EL implants. In I. J. Hochmair-Desoyer & E. C. Hochmair (Eds.), *Advances in Cochlear Implants* (pp. 326–332). Vienna, Austria: Manz.

Shannon, R. V. (1983). Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics. *Hearing Research*, *11*, 157–189.

Shannon, R. V. (1989). Detection of gaps in sinusoids and pulse trains by patients with cochlear implants. *Journal of the Acoustical Society of America*, *85*, 2587–2592.

Shannon, R. V. (1992). Temporal modulation transfer functions in patients with cochlear implants. *Journal of the Acoustical Society of America*, *91*, 2156–2164.

Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, *270*, 303–304.

Shannon, R. V., Zeng, F.-G, & Wygonski, J. (1998). Speech recognition with altered spectral distribution of envelope cues. *Journal of the Acoustical Society of America*, *104*, 2467–2476.

Skinner, M. W., Clark, G. M., Whitford, L. A., Seligman, P. M., Staller, S. J., Shipp, D. B., Shallop, J. K., Everingham, C., Menapace, C. M., Arndt, P. L., Antogenelli, T., Brimacombe, J. A., Pijl, S., Daniels, P., George, C. R., McDermott, H. J., & Beiter, A. L. (1994). Evaluation of a new spectral peak coding strategy for the Nucleus 22 channel cochlear implant system. *American Journal of Otology*, *15(Suppl. 2)*, 15–27.

Skinner, M. W., Ketten, D. R., Holden, L. K., Harding, G. W., Smith, P. G., Gates, G. A., Neely, J. G., Kletzker, G. R., Brunsden, B., & Blocker, B. (2002). CT-Derived estimation of cochlear morphology and electrode array position in relation to word recognition in Nucleus-22 recipients. *Journal of the Association for Research in Otolaryngology*, *3*, 332–350.

Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, *416*, 87–90.

Sorkin, R. D. (1990). Perception of temporal patterns defined by tonal sequences. *Journal of the Acoustical Society of America*, *87*, 1695–1701.

Wingfield, A. (1996). Cognitive factors in auditory performance: Context, speech or processing, and constraints of memory. *Journal of American Academy of Audiology*, *7*, 175–182.

Zeng, F. G. (2002). Temporal pitch in electric hearing. *Hearing Research*, *174*, 101–106.

REFERENCE NOTES

1 Burns, E., Lineaweaver, S., & Sanborn, E. (February 2000). Perception of melodies via cochlear implants and implant simulations. Poster presented at 2000 ARO Mid-Winter Meeting, St. Petersburg Beach, Florida.

Appendix

Frequency Components of the 12 Familiar Melodies

Mel	Range	Notes (in Hz)														Largest interval	Int extent
1	220–369	292	292	292	220	245	245	220	369	369	329	329	292			6th	9
	A3–F#4	D4	D4	D4	A3	B3	B3	A3	F#4	F#4	E4	E4	D4				
2	261–440	261	261	392	392	440	440	392	348	348	329	329	292	292	261	5th	9
	C4–A4	C4	C4	G4	G4	A4	A4	G4	F4	F4	E4	E4	D4	D4	C4		
3	292–440	392	440	392	348	329	348	392	292	329	348	329	348	392		4th	7
	D4–A4	G4	A4	G4	F4	E4	F4	G4	D4	E4	F4	E4	F4	G4			
4	261–392	329	292	261	292	329	329	329	292	292	292	329	392	392		m3rd	7
	C4–G4	E4	D4	C4	D4	E4	E4	E4	D4	D4	D4	E4	G4	G4			
5	245–369	329	276	329	329	276	329	369	329	292	276	245	276	292		m3rd	7
	B3–F#4	E4	C#4	E4	E4	C#4	E4	F#4	E4	D4	C#4	B3	C#4	D4			
6	220–369	292	292	329	369	292	369	329	220	292	292	329	369	292	276	5th	9
	A3–F#4	D4	D4	E4	F#4	D4	F#4	E4	A3	D4	D4	E4	F#4	D4	C#4		
7	220–348	261	292	348	348	348	348	292	261	220	261	348	348	348	348	4th	8
	A3–F4	C4	D4	F4	F4	F4	F4	D4	C4	A3	C4	F4	F4	F4	F4		
8	261–392	261	261	292	261	348	329	261	261	292	261	392	348			5th	7
	C4–G4	C4	C4	D4	C4	F4	E4	C4	C4	D4	C4	G4	F4				
9	220–348	220	220	261	220	220	261	220	261	348	329	292	292	261		4th	7
	A3–F4	A3	A3	C4	A3	A3	C4	A3	C4	F4	E4	D4	D4	C4			
10	220–440	220	440	369	329	276	329	245	220	440	369	329	276	329		octave	12
	A3–A4	A3	A4	F#4	E4	C#4	E4	B3	A3	A4	F#4	E4	C#4	E4			
11	207–466	207	276	261	276	348	309	276	309	348	276	276	348	414	466	4th	14
	G#3–A#4	G#3	C#4	C4	C#4	F4	D#4	C#4	D#4	F4	C#4	C#4	F4	G#4	A#4		
12	207–523	309	261	207	261	309	414	523	466	414	261	292	309			m6th	16
	G#3–C5	D#4	C4	G#3	C4	D#4	G#4	C5	A#4	G#4	C4	D4	D#4				

1. Old MacDonald Had a Farm; 2. Twinkle, Twinkle, Littler Star; 3. London Bridge is Falling Down; 4. Mary Had a Little Lamb; 5. This Old Man; 6. Yankee Doodle; 7. She'll be Coming 'Round the Mountain; 8. Happy Birthday; 9. Lullaby, and Good Night; 10. Take Me Out to the Ball Game; 11. Auld Lang Syne; 12. Star Spangled Banner. Largest interval: largest interval in melody (m = minor); Int extent: range in semitones between the highest and lowest notes in the melodies.