Effect of a competing instrument on melodic contour identification by cochlear implant users

John J. Galvin III, Qian-Jie Fu, and Sandra I. Oba

*Department of Communication and Auditory Neuroscience, House Ear Institute, Los Angeles, California 90057
jgalvin@hei.org, qfu@hei.org, soba@hei.org*

Abstract: Cochlear implant (CI) users have limited access to fundamental frequency (F0) and timbre cues, which are needed to segregate competing voices and/or musical instruments. In the present study, CI users' melodic contour identification was measured for three target instruments in the presence of a masker instrument; the F0 of the masker was varied relative to the target instruments. Mean CI performance significantly declined in the presence of the masker, while mean normal-hearing performance was largely unaffected. However, the most musically experienced CI users were able to make use of timbre and F0 differences between instruments.

© 2009 Acoustical Society of America

PACS numbers: 43.66.Ts, 43.66.Jh, 43.66.Hg, 43.75.Cd [DD]

Date Received: September 8, 2008 **Date Accepted:** December 10, 2008

1. Introduction

In music, timbre cues are important for segregating multiple instruments presented in a polyphonic context. When multiple instruments are played simultaneously or non-simultaneously, timbre cues can be used to track the melodic components within a piece of music. Due to the limited spectral resolution currently available in cochlear implants (CIs), CI users often have great difficulty in understanding, perceiving, and appreciating music, especially large ensemble performances (e.g., Looi *et al.*, 2008). Melodic pitch perception is difficult for CI users, even with a single instrument (e.g., Guerts and Wouters, 2004; Kong *et al.*, 2004). Melodic contour identification (MCI) has been used to quantify listeners' melodic pitch perception. Galvin *et al.* (2007) found that CI users' MCI performance was quite varied, depending on musical experience before and after implantation, and was generally poorer than that of normal-hearing (NH) listeners. Galvin *et al.* (2008) also showed that timbre cues can significantly affect CI users' MCI performance, suggesting that melodies might sound quite different (in terms of pitch sequences), depending on which instrument is playing. In general, as musical listening tasks become more complex, CI users experience greater difficulty. Musically experienced CI users may encounter less difficulty, and training has been shown to improve CI users' timbre perception (Gfeller *et al.*, 2002) and MCI performance (Galvin *et al.*, 2007).

However, even musically experienced CI users can have great difficulty with polyphonic music. Looi *et al.* (2008) showed that instrument identification was much better with solo instruments than with small or large ensembles. Due to the lack of spectro-temporal fine structure cues, CI listeners have difficulty using timbre cues to segregate instruments, resulting in distorted or confusing melody lines. Currently, it is not well known how CI listeners' melodic pitch perception may be influenced by a competing instrument. In the present study, MCI was measured for different instruments (organ, violin, and piano), with or without a competing instrument (piano). The masking contour was always flat (no change in pitch). Four masker notes (A3, A4, A5, and A6) were tested to observe whether listeners could use F0 and timbre differences between the masker and target instruments to identify the target contour.

2. Methods

Seven CI and seven NH subjects participated in the present experiment. All CI subjects participated in previous MCI studies (Galvin *et al.*, 2007, 2008), and thus were familiar with the MCI task and instrument stimuli. CI subject demographics are shown in Table 1. Note that subjects

Table 1. CI subject demographics.

Subject	Gender	Age	CI experience (years)	Device/strategy
S1	M	50	1	Freedom/ACE
S2	F	63	4	N24/ACE
S3	M	56	17	N22/SPEAK
S4	M	49	15	N22/SPEAK
S5	M	77	11	N22/SPEAK
S6	M	74	8	CII/Fidelity 120
S7	M	60	16	N22/SPEAK

S1 and S2 had greater music experience before and after implantation, compared with the other CI subjects. The mean age for NH subjects was 37.9 years (range: 20–52 years). Two of the NH subjects were active musicians, and five of the NH subjects had some musical instruction (e.g., piano, cello, organ, or drum lessons) during childhood. All NH subjects participated in the previous study by Galvin *et al.* (2008), and thus were experienced with the MCI task and instrument stimuli. All subjects were paid for their participation, and all provided informed consent before participating in the experiment.

Target stimuli consisted of nine five-note melodic contours (rising, flat, falling, flat-rising, falling-rising, rising-flat, falling-flat, rising-falling, and flat-falling), similar to those used in the previous MCI studies (Galvin *et al.*, 2007, 2008). All notes in the contours were generated according to $f = 2^{x/12} f_{\text{ref}}$, where f is the frequency of the target note, x is the number of semitones relative to the root note, and f_{ref} is the frequency of the root note (A3, the lowest note of the contour). The spacing between successive notes in each contour was varied to be one, two, three, four, or five semitones. Across all target contours and interval conditions, the F0 range was 220–698 Hz. Each note was 250 ms in duration, and the interval between notes was 50 ms. The contours were played by sampled instruments with synthesis (Roland Sound Canvas with Microsoft Wavetable synthesis). The three target instruments were organ, violin, and piano. The three instruments were selected to represent a range of spectral and temporal complexities. For example, the organ has a relatively simple spectrum with little attack, the violin has a more complex spectrum with a slow attack, and the piano has the most complex spectrum with a sharp attack. In the previous study by Galvin *et al.* (2008), CI users' MCI performance was best with the organ and poorest with the piano. Baseline performance for the three target instruments was measured with no masker. The masker stimuli consisted of the flat contour played by the piano, i.e., each note was the same. To test the effect of F0 overlap between the target and masker contours, the F0 of the masker was varied to be either A3 (220 Hz), A4 (440 Hz), A5 (880 Hz), or A6 (1720 Hz). The masker and target contours were normalized to have the same long term rms amplitude (65 dB).

Polyphonic music often involves multiple instruments that interleave and overlap in time, and listeners may use temporal offsets between instruments to track different melodic components. When played simultaneously by multiple sound sources, it is more difficult to track the melodic components. In the present study, the masker and target were presented simultaneously; the onset, duration, and offset for each note of the target and masker were the same.

Figure 1 shows electrodiagrams for the target A3 rising contour (five semitones between notes) for the organ and piano, with or without the piano masker; audio examples of these stimuli are given in Mm. 1. The electrodiagrams were generated using custom software that implemented the default parameter settings of the Advanced Combination Encoder (ACE) (Vandali *et al.*, 2000) strategy typically implemented in Nucleus-24 and Freedom implant devices, i.e., frequency allocation Table 6, 900 Hz/channel, 8 maxima, etc. With no masker, the rising contour can be easily observed, as the F0s and harmonics span most of the electrode array; note that the stimulation pattern is more spectrally dense and temporally complex for the

Fig. 1. (Color online) Left panels: electrodegrams for the A3 rising contour (five semitones between notes) played by the organ with no masker (left), the A3 piano masker, and the A5 piano masker. The piano masker was the flat contour. The left axis shows the electrode number. Right panels: similar electrodegrams for the A3 piano target rising contour.

piano than for the organ. With the A3 piano masker, the stimulation pattern is denser, as the eight spectral maxima are shared between the masker and target instruments; again, the stimulation pattern is denser for the piano than for the organ. With the A5 piano masker, the stimulation pattern is less dense, and the rising contour can be observed across the apical channels; again, the stimulation pattern remains denser and more complex for the piano than for the organ. Note that these stimulation patterns could be different with different speech parameter settings (e.g., coding strategy, acoustic frequency allocation, etc.).

[Mm. 1. Audio example of the target A3 rising contour (5 semitones between notes) for the organ and piano, with or without the piano masker. This is a file of type wav (546 Kb).]

For both NH and CI subjects, stimuli were presented acoustically via a single loudspeaker (Tannoy Reveal) at 70 dBA. Testing was conducted in a sound-treated booth (IAC) with subjects directly facing the loudspeaker. CI subjects were tested using their clinically assigned speech processors. CI subjects were asked to use their everyday sensitivity and volume settings; once set, subjects were asked not to change these settings. NH subjects listened to unprocessed signals. The different target/masker conditions were tested independently and randomized across subjects; each condition was tested three times. During each test block, a target stimulus was randomly selected (without replacement) from among the 45 stimuli (9 contours $\times$ 5 intonations). Subjects were asked to click on one of the nine response choices shown onscreen, after which the next target stimulus was presented. Subjects were instructed to guess if they were unsure. Subjects were explicitly instructed that they would hear two simultaneous contours, one of which would always be the flat piano masker, and that the pitch of the masker might change from test to test. Subjects were allowed to repeat each stimulus up to three times; however, no preview or trial-by-trial feedback was provided.

3. Results

Figure 2 shows mean MCI performance for CI and NH subjects, as a function of target instrument. For CI subjects, baseline performance with no masker was 79.1% correct for the organ, 73.3% correct for the violin, and 60.5% correct for the piano. A two-way repeated-measures analysis of variance (RM ANOVA) showed that CI performance was significantly affected by

Fig. 2. Mean MCI performance across CI subjects (left) and NH subjects (right) for the different piano masker F0s, as a function of target instrument. The error bars show one standard deviation of the mean, and the dashed line shows chance performance level (11.1% correct).

target instrument [$F(2, 48) = 10.02, p = 0.003$] and masker note [$F(4, 48) = 8.03, p < 0.001$]. *Post hoc* Bonferroni t-tests showed that performance was significantly better without the piano masker ($p < 0.05$), and that there was no significant difference among masker note conditions. *Post hoc* Bonferroni t-tests also showed that masked MCI performance was not significantly affected by the target instrument, except for the A4 piano masker (violin > piano; $p > 0.05$). For NH subjects, mean performance with no masker was 93.3% correct for the organ, 93.0% correct for the violin, and 90.6% correct for the piano. A two-way RM ANOVA showed no significant effects for target instrument [$F(2, 48) = 3.83, p = 0.052$] or masker note [$F(4, 48) = 0.49, p = 0.743$].

While mean CI users' masked MCI performance was largely unaffected by differences in target instrument timbre or by differences in masker note, individual CI subjects exhibited markedly different patterns of results. Figure 3 shows individual CI subjects' MCI performance, as a function of target instrument. Subject S1's baseline performance was >97% correct for all three target instruments. While the piano masker generally reduced performance, S1 was sensitive to target instrument timbre (i.e., better performance when masker and target instruments were different) and masker note (i.e., better performance as the F0 separation between masker and target was increased). Subject S2 also exhibited some sensitivity to target instrument timbre and masker note. S2's performance with the organ was virtually unchanged by the presence of the piano masker. For the violin and piano targets, increasing the F0 separation between the masker and target improved performance. The pattern of results was different for subject S4. While S4 was somewhat sensitive to target instrument timbre, increasing the F0 separation between masker and target reduced performance. For subjects S3 and S5, the effects of instrument timbre and masker note were inconsistent. Subjects S6 and S7 exhibited little sensitivity to target instrument timbre or masker note.

4. Discussion

In general, CI subjects generally had great difficulty in perceiving the target melodic contours in the presence of a competing simultaneous instrument masker. Most CI subjects were unable to make use of timbre and/or pitch differences between the masker and target stimuli. This deficit in masked MCI performance is in agreement with previous studies showing poorer instrument identification in the context of small or large instrument ensembles (Looi *et al.*, 2008). The

Fig. 3. MCI performance for individual CI subjects for the different piano masker F0s, as a function of target instrument. The error bars show one standard deviation of the mean, and the dashed line shows chance performance level (11.1% correct).

mean CI results are also in agreement with [Stickney *et al.* \(2007\)](#), who showed that CI listeners were unable to make use of F0 differences between talkers to recognize speech in the presence of a competing talker. Note that the MCI task in the present experiment required listeners to attend exclusively to pitch cues, as opposed to a speech recognition task in which pitch cues are used to segregate talkers and access lexically meaningful information. Given the weak F0 coding in speech processors, CI users must make use of spectral envelope differences and/or periodicity cues to track changes in pitch. Because of the relatively coarse spectral resolution, competing speakers or instruments cannot be segregated, even with fairly large F0 differences.

Interestingly, some CI subjects (S1 and S2) were able to consistently use timbre and F0 differences to identify the target contours. These subjects were the most musically experienced, which may have contributed to their better attention to timbre and pitch cues. Subject S1 exhibited a somewhat predictable pattern, i.e., masked MCI performance improved as the timbre and/or F0 differences were increased between the target and the masker. S2 exhibited a similar pattern (albeit poorer overall performance) with the violin and piano target instruments. For the remaining less musically experienced CI subjects, there were no clear effects for target instrument timbre and/or masker F0. For example, performance worsened for S4 as the masker F0 was increased, suggesting that the highest spectral component dominated the percept. For subject S7, the presence of the competing masker only marginally reduced the already low MCI scores. The different devices and processing strategies used by CI subjects may have also contributed to differences in individual performance, although there are too few subjects to fairly compare across CI devices. Despite differences among devices and strategies, contemporary CIs do not provide the spectro-temporal fine structure cue to support complex pitch perception and timbre discrimination. As such, musically experienced CI users may be better able to use the limited spectral and temporal envelope cues for music perception. Previous studies also show that targeted training improved CI users' music perception ([Galvin *et al.*, 2007](#); [Gfeller *et al.*, 2002](#)), even though CI subjects had years of experience with their clinical devices and strategies. Until CIs can provide fine structure cues, music experience and/or training may provide the greatest advantage for music perception with electric hearing.

Acknowledgments

The authors would like to thank all the subjects who participated in this study. The authors would also like to thank Dr. Diana Deutsch and two anonymous reviewers, as well as David Landsberger and Bob Shannon, for insightful comments on this work. This work was supported by NIH Grant No. DC004993.

References and links

- Galvin, J., Fu, Q.-J., and Nogaki, G. (2007). "Melodic contour identification by cochlear implant listeners," *Ear Hear.* **28**, 302–319.
- Galvin, J., Fu, Q.-J., and Oba, S. (2008). "Effect of instrument timbre on melodic contour identification by cochlear implant users," *J. Acoust. Soc. Am.* **124**, EL189–EL195.
- Geurts, L., and Wouters, J. (2004). "Better place coding of the fundamental frequency in cochlear implants," *J. Acoust. Soc. Am.* **115**, 844–852.
- Gfeller, K., Witt, S., Adamek, M., Mehr, M., Rogers, J., Stordahl, J., and Ringgenberg, S. (2002). "Effects of training on timbre recognition and appraisal by postlingually deafened cochlear implant recipients," *J. Am. Acad. Audiol.* **13**, 132–145.
- Kong, Y.-Y., Cruz, R., Jones, J., and Zeng, F.-G. (2004). "Music perception with temporal cues in acoustic and electric hearing," *Ear Hear.* **25**, 173–185.
- Looi, V., McDermott, H., McKay, C., and Hickson, L. (2008). "Music perception of cochlear implant users compared with that of hearing aid users," *Ear Hear.* **29**, 421–434.
- Stickney, G., Assmann, P., Chang, J., and Zeng, F.-G. (2007). "Effects of cochlear implant processing and fundamental frequency on the intelligibility of competing sentences," *J. Acoust. Soc. Am.* **122**, 1069–1078.
- Vandali, A. E., Whitford, L. A., Plant, K. L., and Clark, G. M. (2000). "Speech perception as a function of electrical stimulation rate: Using the Nucleus 24 cochlear implant system," *Ear Hear.* **21**, 608–624.