Perceptual fusion of polyphonic pitch in cochlear implant users

Patrick J. Donnelly
Department of Computer Science, Johns Hopkins University, Baltimore, Maryland 21218
donnell@cs.jhu.edu

Benjamin Z. Guo and Charles J. Limb
Department of Otolaryngology—Head and Neck Surgery, Johns Hopkins Hospital, Baltimore, Maryland 21287
bguo@fas.harvard.edu, climb@jhmi.edu

Abstract: In music, multiple pitches often occur simultaneously, an essential feature of harmony. In the present study, the authors assessed the ability of cochlear implant (CI) users to perceive polyphonic pitch. Acoustically presented stimuli consisted of one, two, or three superposed tones with different fundamental frequencies (f_0). The normal hearing control group obtained significantly higher mean scores than the CI group. CI users performed near chance levels in recognizing two- and three-pitch stimuli, and demonstrated perceptual fusion of multiple pitches as single-pitch units. These results suggest that limitations in polyphonic pitch perception may significantly impair music perception in CI users.

© 2009 Acoustical Society of America
PACS numbers: 43.66.Ts, 43.66.Hg, 43.75.Cd [QJF]
Date Received: July 29, 2009 **Date Accepted:** August 27, 2009

1. Introduction

The ability of cochlear implant (CI) users to perceive music remains severely limited, primarily due to limited pitch resolution (Gfeller *et al.*, 2002; Pijl, 1997). While studies have shown that CI users are able to adequately perceive temporal cues that convey rhythmic information, the perception of pitch and timbre remains quite poor with current CI hardware and processing strategies (Leal *et al.*, 2003; McDermott, 2004). Previous studies have shown that CI users are severely impaired compared to normal hearing (NH) subjects in tests of pitch perception using acoustically presented stimuli, with CI users rarely exhibiting pitch discrimination thresholds of less than several semitones (Gfeller *et al.*, 2002; Looi *et al.*, 2004). Most published studies on pitch perception have focused on pitch discrimination, in which subjects are required to detect whether two sounds differ in pitch, and pitch ranking, in which subjects are asked to listen to two sounds presented in sequence and judge which one has the higher pitch. While these approaches are certainly valid, various elements (i.e., melody, harmony, rhythm, and timbre) usually occur simultaneously in music. In the context of the impaired pitch resolution described in CI users, it is germane that nearly all forms of music utilize at least some degree of polyphony (where multiple pitches occur simultaneously), an essential feature of harmony. Comparatively little research, however, has been done on the perception of polyphonic pitch (or harmony) in CI subjects. One recent study (Galvin *et al.*, 2009) examined melodic contour segregation in CI subjects using acoustically presented stimuli and found that CI users have difficulty segregating competing melodic contours even in the presence of timbral cues.

The objectives of the present study were to evaluate the ability of post-lingually deafened adult CI users to perceive the number of pitches in acoustically presented stimuli and to compare their performance with that of NH adults. Subjects listened to acoustically presented stimuli consisting of one, two, or three simultaneous tones with different fundamental frequencies (f_0) within a single octave. Both pure tones and piano tones were used to assess the effect of harmonics on polyphonic pitch perception. We hypothesized that CI users, as a result of diminished pitch resolution, would show decreased ability to differentiate between single versus mul-

 © 2009 Acoustical Society of America

Downloaded 09 Oct 2011 to 153.90.5.75. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.jsp

Table 1. CI subject demographics.

Subject	Sex	Age	CI experience (years)	Device	Processor
CI1	M	69	5	ABC HiRes 90K	Harmony
CI2	F	58	6	CC Nucleus 24	Esprit 3G
CI3	M	27	2	ABC HiRes 90K	Harmony
CI4	F	68	1	CC Nucleus Contour	Freedom
CI5	M	58	5	ABC HiRes 90K	Harmony
CI6	M	46	2	CC Nucleus Contour	Freedom
CI7	F	33	4	CC Nucleus 24	Freedom
CI8	F	54	2	CC Nucleus Contour	Freedom
CI9	M	47	11	ABC Clarion	Platinum BTE
CI10	F	56	2	CC Nucleus Contour	Freedom
CI11	F	67	4	CC Nucleus Contour	Freedom
CI12	F	54	1	ABC HiRes 90K	Harmony

ABC: Advanced Bionics Corporation; CC: Cochlear Corporation; BTE: behind-the-ear.

tiple tones in comparison to NH controls. We further hypothesized that the ability of CI users to detect polyphony would increase as a function of interval distance between pitches, due to the presumptive relationship between increased frequency separation and improved perception of polyphony.

2. Methods

Twelve monaurally implanted CI users aged 27–69 years (mean=53.1±12.6 years) and 12 NH controls participated in the study. CI subjects used a variety of devices and processing strategies (Table 1). Each CI user had at least 1 year of experience using their implant system. The 12 NH adults ranged in age from 21 to 45 years (mean=28.4±9.2 years). All subjects completed a musical experience questionnaire to ascertain the extent of their musical training. No CI or NH subjects had musical training beyond an amateur level. All experiments were performed at the Sound and Music Perception Laboratory of Johns Hopkins Hospital, under an IRB approved research protocol. Informed consent was obtained for all participants.

All piano stimuli were recorded using Ivory Grand Piano Virtual Instrument (Synthogy) on the Apple LOGIC PRO 7.0 platform. Pure tones were generated using AUDACITY 1.2.5 (Dominic Mazzoni, open source). All stimuli were exactly 2.5 s in duration and were normalized by root-mean square power with equal-loudness contour adjustment using ADOBE AUDITION 3.0. Pure tone stimuli were given linear rise/decay ramps of 200 ms to reduce onset clicks. All stimuli consisted of pitches from within a central octave ranging in f_0 from 261 (C4) to 523 Hz (C5). Single-pitch stimuli consisted of either pure tones or piano tones from C4 to B4 (12 unique pitches and 24 total stimuli). Two-pitch stimuli consisted of either pure tone or piano tone representations of all 12 possible intervals within the range C4–C5 (1–12 semitone interval distance and 24 total stimuli). Three-pitch stimuli consisted of either pure tone or piano tone representations of six unique symmetric chords (equal interval spacing between lower/middle and middle/higher pitches) within the range C4–C5. No stimuli contained both pure and piano tones. For each three-pitch stimulus, interval spacing ranged from one to six semitones (for both lower/middle and middle/upper pitches). The pitches were equally spaced in order to maintain a consistent musical dispersion of pitches within each of the three-pitch stimuli and minimize the effect of varying intervals on perception of polyphony. Because only six unique arrangements of three equally spaced pitches are possible within a single octave, two sets of three-pitch stimuli were created (24 total stimuli for both piano and pure tones). Pitches used in the sets of two- and three-pitch stimuli were mathematically distributed symmetrically across the octave

Downloaded 09 Oct 2011 to 153.90.5.75. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.jsp

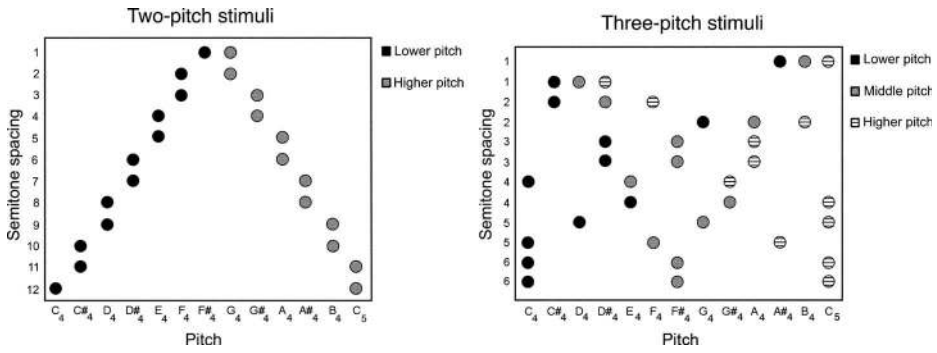


Fig. 1. Pitches and semitone spacing of two-pitch stimuli (left) and three-pitch stimuli (right).

so as to minimize over- or under-representation of any given pitch, as shown in Fig. 1. In total, 72 stimuli were presented to each subject.

Stimuli were randomly presented in a soundproof booth through a single calibrated loudspeaker (Sony SS-MB150H) at a presentation level of 80 dB sound pressure level through an OB822 clinical audiometer (Madsen Electronics); the speaker was positioned directly in front of the listener. For CI users, the contralateral ear (which was profoundly impaired in all individuals) was occluded with an earplug to diminish the effects of any minimal residual hearing, and no hearing aids were used. No subjects reported being able to hear any stimuli through their non-implanted ear. Stimuli were presented in a three alternative single-interval forced-choice procedure in which the subjects were instructed to choose whether the given stimuli consisted of one, two, or three pitches. Subjects were familiarized with the stimuli and procedure prior to formal testing. No feedback was given regarding the correctness of responses. The number of correct responses for each subject was averaged across the separate tone and pitch-number conditions to obtain an overall mean score.

3. Results

For all conditions, the CI group scored significantly lower than the NH group. The overall mean scores for each subject group were $43.1 \pm 12.3\%$ for CI users and $66.9 \pm 9.4\%$ for NH subjects. An unpaired t-test revealed a significant difference in overall mean scores between the NH and CI groups ($p < 0.001$). No statistically significant difference was found between average scores for pure tones and piano tones across subject groups. As none of the subjects had significant music training, musical experience was not analyzed as a covariate in this study. The mean scores of both subject groups for all stimuli, pure tones, and piano tones are shown in Fig. 2.

The CI group was significantly impaired in perceiving two- and three-pitch stimuli and scored close to 33% (i.e., near chance levels) when identifying two- and three-pitch stimuli (CI:

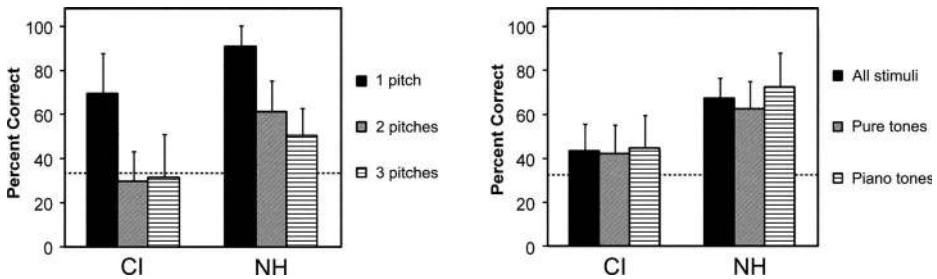


Fig. 2. Mean performance accuracy across CI subjects and NH subjects for one-, two-, and three-pitch stimuli (left) and for all stimuli, pure tones, and piano tones (right). The error bars show one standard deviation of the mean, and the dashed line shows chance performance level (33.3% correct).

Downloaded 09 Oct 2011 to 153.90.5.75. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.jsp

Table 2. Confusion matrix for NH subjects.

		Identified			
		One pitch	Two pitches	Three pitches	No response
Presented	One pitch	260	24	2	2
		90.3%	8.3%	0.7%	0.7%
	Two pitches	63	172	53	0
		21.9%	59.7%	18.4%	0%
	Three pitches	23	122	143	0
		8.0%	42.4%	49.6%	0%

one-pitch – 69.1±18.6%, two-pitch – 29.1±14.1%, and three-pitch – 30.9±20.0%). In comparison, the NH group was much more successful at distinguishing single from multiple pitches, but demonstrated difficulties at distinguishing between two- and three-pitch stimuli (NH: one-pitch – 90.6±9.9%, two-pitch – 60.4±14.8%, and three-pitch – 49.6±13.0%). With the exception of 1 CI subject who achieved a higher mean score (72.2±11.5%) than 7 out of the 12 NH subjects, performance ranges for the 2 groups had little overlap. Unpaired *t*-tests revealed significant differences in scores between the NH and CI groups for all pitch conditions (one-pitch – $p=0.0026$, two-pitch – $p<0.001$, and three-pitch – $p=0.0136$). While NH subjects often identified three-pitch stimuli as having two pitches (suggesting that three-pitch condition was the most difficult for NH subjects), CI subjects often identified both three-pitch and two-pitch stimuli as a single pitch. NH subjects were less likely to identify two- and three-pitch stimuli as one pitch compared to CI users ($p=0.004$ for two-pitch stimuli and $p<0.001$ for three-pitch stimuli, unpaired *t*-test). Confusion matrices for NH and CI subjects are presented in Tables 2 and 3, respectively. For 4 out of 1728 total stimulus presentations (0.023%), the response period eventually timed out without a subject response. While this altered chance levels to a very small extent, we did not feel that this was a relevant variable to include in the analysis given the small magnitude of this effect.

Figure 3 shows the two-pitch performance accuracy of both subject groups as a function of interval distance. CI subjects performed near chance levels for all three-pitch conditions and most two-pitch conditions. For three-pitch conditions, there was no apparent relationship between interval spacing and ability to detect polyphony. For two-pitch conditions, increased interval spacing did not lead to better performance for detection of polyphony. In fact, an inverse relationship was suggested for identification of the one semitone interval spacing in two-pitch conditions (minor second interval), for which CI users were nearly as accurate as NH subjects.

Table 3. Confusion matrix for CI subjects.

		Identified			
		One pitch	Two pitches	Three pitches	No response
Presented	One pitch	199	63	26	0
		69.1%	21.9%	9.0%	0%
	Two pitches	136	84	67	1
		47.2%	29.2%	23.3%	0.3%
	Three pitches	112	86	89	1
		38.9%	29.9%	30.9%	0.3%

Downloaded 09 Oct 2011 to 153.90.5.75. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.jsp

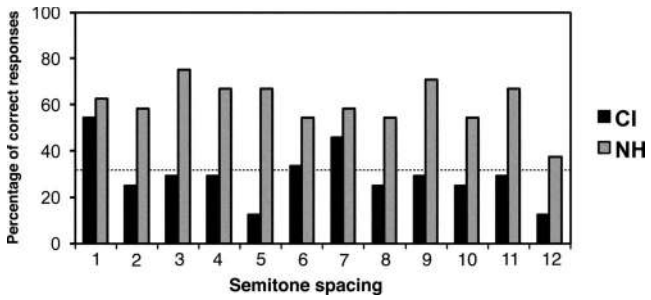


Fig. 3. Percentage of correct responses (number of correct responses/number of stimuli presented) across CI subjects and NH subjects as a function of semitone spacing in two-pitch stimuli. The dashed line shows chance performance level (33.3% correct).

4. Discussion

While previous studies have examined pitch resolution in CI subjects using pitch ranking and pitch discrimination tasks, the present study instead examined the ability of CI users to perceive acoustic polyphonic pitch using a novel pitch separation task in which subjects were asked to distinguish between one-, two-, and three-pitch stimuli. These categories of musical stimuli were chosen for their ubiquitous usage in Western music for melodies, intervals, and chords. The results from this study show that CI users obtain significantly lower average scores than NH subjects when asked to distinguish between single and multiple acoustically presented tones. CI users identified two- and three-pitch stimuli near chance levels, and demonstrated frequent perceptual fusion of multiple-pitch stimuli as single-pitch units. Both groups demonstrated a bias toward identifying all stimuli as one pitch (NH: one pitch—40%, two pitches—37%, and three pitches—23%; CI: one pitch—52%, two pitches—27%, and three pitches—21%). However, NH subjects were less likely to identify two- and three-pitch stimuli as one pitch. While a listener's ability to identify the number of components in a polyphonic stimulus does not necessarily correspond to the ability to perceive differences between polyphonic stimuli in a musical context (e.g., whether a musical triad is major or minor), perceptual fusion of polyphonic pitch likely impairs CI users in accurately perceiving many aspects of music, such as harmony, consonance, dissonance, and tonality.

One explanation for the lower average scores of CI subjects is the limited pitch resolution afforded by current CI devices. While pitch resolution varies widely across CI users, studies have shown that CI users rarely exhibit pitch discrimination thresholds of less than several semitones when two tones are presented sequentially (Gfeller *et al.*, 2002; Pretorius and Hanekom, 2008). Additionally, studies in NH listeners have shown that much greater frequency differences are required for the resolution of two tones sounding simultaneously than for the discrimination of two tones presented sequentially. For NH listeners, the limit for separation of two superposed pure tones within 100–500 Hz is roughly one semitone, almost ten times larger than the just noticeable difference for single pure tones within the same frequency range (Roederer, 1995). CI users would be expected to demonstrate even poorer separation of superposed tones than NH subjects due to poor pitch resolution.

CI subjects were most accurate in identifying two-pitch stimuli for both piano and pure tones when the two pitches were one semitone apart. This result was unexpected, given that sounds with more similar f_0 are more likely to be heard as a single stream (Oxenham, 2008). In CI subjects, all f_0's were likely processed within a single analysis band. Corresponding harmonics were also likely processed in the same bands. It is possible that temporal cues in the envelope extraction from the constructive and destructive interferences between two tones at slightly different frequencies aided CI subjects in the correct identification of the number of pitches present. This interference (also known as a beat) is perceived as periodic variations in volume whose rate is the difference between the two frequencies (Δf). Numerous studies have shown

Downloaded 09 Oct 2011 to 153.90.5.75. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.jsp

that macroscopic temporal cues are readily perceived by CI users (Leal *et al.*, 2003), and it is possible that temporal envelope cues from beats aided CI subjects in separating superposed pitches. It is possible that beats provide optimal temporal cues to CI users when Δf is near one semitone, while temporal envelope cues with higher rates (when Δf is greater than one semitone) are less useful for helping CI users in separating polyphonic pitch.

The similarity in average scores between pure tones and piano tones across both subject groups was also an unexpected result. In the present study, it was anticipated that both subject groups would utilize the additional pitch information present in the overtones of the wider bandwidth piano tones to aid in the identification of polyphonic pitch. However, our results show that the presence of additional harmonic information does not necessarily increase the ability of both normal and CI subjects to perceive polyphony. Further research incorporating a larger set of complex tones and an examination of electrical stimulation patterns may be needed to better assess the effect of harmonics on pitch resolution in CI subjects.

5. Conclusion

CI users were found to obtain significantly lower scores than NH subjects when asked to distinguish between stimuli consisting of one, two, or three superposed tones, demonstrating perceptual fusion of multiple tones as single-pitch units. However, CI users were nearly as accurate as NH subjects at identifying the number of pitches present when two superposed tones were separated by one semitone. Overall, no statistically significant difference was found between average scores for pure tones and piano tones across both subject groups. This finding indicates that the presence of additional pitch information in complex tones may not aid either subject group in the resolution of polyphonic pitch. Perceptual fusion of polyphonic pitch likely contributes to the poor perception of harmony in CI users. As most of the music that CI users encounter is polyphonic, these findings indicate the need for further research on how polyphonic pitch is perceived by CI users. The development of processing strategies directed toward a more accurate representation of polyphonic pitch should greatly improve the ability of CI users to perceive complex musical stimuli.

Acknowledgment

The authors would like to thank all the subjects who participated in this study.

References and links

Galvin, J. J., Fu, Q., and Oba, S. I. (**2009**). "Effect of a competing instrument on melodic contour identification by cochlear implant users," J. Acoust. Soc. Am. **125**, EL98–EL103.
Gfeller, K., Turner, C., Mehr, M., Woodworth, G., Fearn, R., Knutson, J., Witt, S., and Stordahl, J. (**2002**). "Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults," Coch. Imp. Inter. **3**, 29–53.
Leal, M. C., Shin, Y. J., and Laborde, M.-I. (**2003**). "Music perception in adult cochlear implant recipients," Acta Oto-Laryngol. **123**, 826–835.
Looi, V., McDermott, H. J., McKay, C. M., and Hickson, L. (**2004**). "Pitch discrimination and melody recognition by cochlear implant users," Int. Congr. Ser. **1273**, 197–200.
McDermott, H. J. (**2004**). "Music perception with cochlear implants: A review," Trends Amplif. **8**, 49–82.
Oxenham, A. J. (**2008**). "Pitch perception and auditory stream segregation: Implications for hearing loss and cochlear implants," Trends Amplif. **12**, 316–31.
Pijl, S. (**1997**). "Labeling of musical interval size by cochlear implant patients and normal hearing subjects," Ear Hear. **18**, 364–72.
Pretorius, L. L., and Hanekom, J. J. (**2008**). "Free field frequency discrimination abilities of cochlear implant users," Hear. Res. **244**, 77–84.
Roederer, J. G. (**1995**). "Superposition of pure tones: First order beats and the critical band," in *The Physics and Psychophysics of Music* (Springer-Verlag, New York), pp. 28–36.

Downloaded 09 Oct 2011 to 153.90.5.75. Redistribution subject to ASA license or copyright; see http://asadl.org/journals/doc/ASALIB-home/info/terms.jsp