**Proceedings of 20th International Congress on Acoustics, ICA 2010**

23-27 August 2010, Sydney, Australia

# Musical note perception and musical sound vocal production in cochlear implant children

**Shuai Wang, Juan Huang**

Key Laboratory of Machine Perception (Ministry of Education)

Department of Intelligence Science, Peking University, Beijing, 100871, People's Republic of China

**PACS:** 43.75.CD

## ABSTRACT

Cochlear implant (CI) successfully restores speech communication in profoundly deaf people. Speech perception by CI users can reach as high as ~100% in quiet condition. However, CI users perform poorly in pitch related tasks such as music perception. Studies have shown that the lack of explicit coding of low frequency information (< 500 Hz) by CI processors contributes to the impaired pitch perception. It remains unknown how the speech production relates to the impaired pitch perception in CI users. In this study, we tested the perception and vocal production of low-frequency musical notes in 14 CI children. In the perception task, subjects were asked to discriminate whether a pair of musical notes were the same or different. In the vocal production task, subjects were asked to vocalize the musical tones after listening to the same notes vocalized by a female speaker. Objective and subjective evaluations of each subject's vocal production were carried out. Correct response of note discrimination, pitch contour perception, and extracted F0 frequency range of sound samples were analyzed. OUR results show that CI Children can discriminate musical notes when there were two or more semitones between the notes. However, all CI users showed a significant loss in the production of the musical notes. F0 of subjects' vocal production of music notes was significantly lower than that of the original sounds they heard. Overall, the subjective evaluation of subjects' vocal production of musical notes showed a poor pitch contour recognition. Findings from the present study showed that the vocal production of low-frequency notes by CI users does not reflect the accuracy of their perception, suggesting that the distorted auditory feedback provided by CI disrupted their vocal production.

## INTRODUCTION

A cochlear implant is a device that receives sounds from the environment and stimulates the auditory nerve so that the deaf person wearing it could perceive these sounds. Cochlear implant (CI) can successfully restore speech perception, but its performance on music perception is limited. The Continuous Interleaved Sampling (CIS) processing strategy is generally used in clinical CI products (Wilson, 1991). Speech sounds are received by the microphone of CI and subfilted into severl frequency bands. The envelope of each frequency band is extracted and used to modulate high frequency electric pulse train,.Signals are then sent to corresponding electrodes which align to the tonotopic map of the cocholea. Current CI processing strategies do not explicitly encode frequency variations carried in sounds such that melody information is not encodedand delivered to the central auditory system.

Studies have shown that CI users' speech perception in quiet can reach as high as 100% giving some rehabilitation training (McDermott, 2004). Correspondingly, CI users' has been shown to produce relatively high intelligibility speech as long as CI were implanted at early age (Habib et al., 2010). However, melody recognition and pitch discrimination in CI users are poor (McDermott, 2004; Gefeller, 1991; Wilson&Dorman, 2008). Postlingual deaf CI users can only recognize simple familiar melodies with correct response at about 30-63% (Kong et a.,l 2004) and the frequency difference limen in CI users has been shown to be much larger at 1-11.7 semitones comparing to normal hearing people (Nimmons, 2008). Researchers have showned that the pitch and pitch contour produced by CI users are poor, especially showing the high compression of fundamental frequency (F0) in musical production by CI users (Nakata, 2006; Xu, 2009). However, how does such poor musical vocal production correlated to perception performance in CI users remains maily un-explored.

In the present study, we examined the musical notes discrimination and corresponding vocal production of these notes in congenitally deaf children with a uni-lateral CI. The aim of the study is to investigate the relationship between pitch and pitch contour perception with vocal production, and to explore the possibility of automatic modulation mechanisms of vocal production in CI users.

## METHOD AND PROCEDURE

### Subject

Fourteen prelingually deaf Mandarin-speaking children with a uni-lateral CI participated in the experiments. Four of them are female and 10 are male, aged between 2-8.8 years old.

The duration of using CI is within 0.5 to 3.5 years. All subjects took the musical note perception tests, while only 7 of the subjects finished musical vocal production recording.

### Testing material

Musical notes are generated by Vocaloid singing vocal-sythesizing software (Yamaha, 2003) using the phoneme “ma” as the basic sound which makes the sound easier to produce by CI users comparing to pure tone note. Nine musical notes are used in the tests, C4 (262 Hz), D4 (294 Hz), E4 (330 Hz), F4 (349 Hz), G4 (392 Hz), A4 (440 Hz), B4 (494 Hz), C5 (446 Hz), and D5 (568 Hz).

For the musical note perception test, 7 notes (C4, D4, E4, F4, G4, A4, B4) which coveres one octave frequency range are presented to 49 pairs which consist of all the combinations, of which 42 pairs are different and 7 pairs are the same. In the musical vocal production recording, three musical notes are presented in groups of three notes which produce 5 pitch contours, “flat” (7 samples), “rising” (7 samples), “falling” (8 samples), “rising and then falling” (10 samples), and “falling and rising” (12 samples). The duration of each note is 500 ms and the interval between notes is 500 ms.

### Procedure

Sound samples were presented through Creative E-MU 0404 USB sound card (Creative) from a SwansH2 loudspeaker (HiVi Inc. California) in a quiet room. The sound level is 75 dB. Subjects were sitting at 1 meter in front of the speaker which was put at the same azimuthal level to their ears. In music note discrimination test, the 49 pair of notes was presented to subjects in a random order. The subject was asked to tell if a pair of notes sounded the same or different. An experimenter recorded the subject’s response from an experimental interface. In music note production recording, the sound samples consisted of three notes were presented to a subject in a random order. The subject was asked to repeat what they heard following each sound sample. A high sensitive Pro49Q microphone (Audio-technica) was used to record both the sound samples and subjects’ vocal production. The recording continued if the process went smoothly, or selse an experimenter would stop the recording. Both musical note presentation and vocal production were controlled by an experimenter via a Thinkpad W500 computer (Lenovo Co. Ltd.) and an experiment interface.

### Data analysis

#### Objective analysis

For musical note perception data, two analysis methods were used. The percentage of correct discrimination of musical notes was calculated based on the number of semitone difference between the pair of notes. The “Hit” rate (when the pair of notes are different and subject’s reponse is “different”) and “False alarm” rate (when the pair of notes are the same and subject’s reponse is “different”) and d’ were calculated.

For musical note vacal production data, both music sounds samples and subject’s vocalizations were segmented manually. Signals were denoised using the Denoise tool of Adobe Audition 3.0 (Adobe systems Inc.) signal processing software. The denoise level was set at 100%, attenuation level was set at 60 dB, precision factor was set at 7, smooth sum was set at 1, spectrum attenuation rate was set at 65%, the number of denoise FFT point was set at 4096. All signals were processed using the same denoise parameters. F0, duration of notes, and duration of intervals between notes of the denoised signals were extracted using STRAIGHT signal processing software for further data analysis.

#### Subjective analysis

Five normal hearing Mandarin-speaking subjects were recruited to subjectively evaluate the vocal production sound samples we recorded. Two of the subjects are 40-50 years old and three are 21-23 years old colleage students. All subjects have general intelligence and no experience of communication with deaf people . Sound samples were presented to the subjects in a random order. The subjects were asked to evaluate if the pitch contour of the three notes in each sound sample was “Flat”, “Up”, “Down”, “Up-Down”, and “Down-Up”. Subjective correct response ratio was defined as “the number of sound samples that the subjective evaluation agrees with the original sound samples” devided by the number of sound samples of each type of pitch contour.

## RESULTS

### Musical note discrimination

In order to examine the difference of the perception of different frequency range in CI children, we devided the 7 musical notes into two frequency ranges, C4-F4 (relatively “Low”) and F4-B4 (relatively “High”). The frequency response range of 22 electrodes of the cochlear implants subjects weared is about 188-7938 Hz. The frequency range of the musical notes used in the experiment is 250-600 Hz which coveres the responding frequency range of the 4 basal CI electrodes. Dividing the testing notes into two “low” and “high” frequency groupes aimed to examine CI users’ musical note discrimination according to the perceived pitch of normal hearing subjects rather than the frequency response of electrodes in the CI. Figure 1 shows that CI children’s musical notes discrimination performance for “low” frequency notes varied more comparing to “high” frequency notes, as indicated by the distribution of d’. The distribution of d’ for “high” frequency notes discrimination has smaller range and high average than that for “low” frequency notes discrimination. There is a trend that the perception of “high” frequency notes is better than that of “low” frequency notes. However, T test did not yield significant difference between the d’ of “low” and “high” frequency conditions T=0.169，p>0.05 (Figure 1).

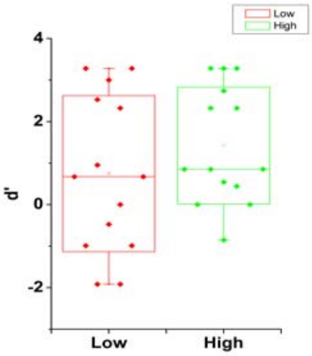

**Figure 1**. Musical note discrimination in children with CI. The X axis is frequency range and Y axis is d’. The error bars are standard deviation. The short horizontal lines are average d’

.

In order to examine if a subject’s musical note discrimination is directly correlated with the frequency difference between musical notes, we divided the pairs of signal samples into groups based on the number of semitones between notes. Data show that subject’s performance increase as a function of frequency difference between notes (Figure 2). Pearson Correlation analysis shows that subject’s musical note discrimination is positively correlated with the number of semitone difference between the notes, r=0.91, p<0.05. The result indicates that subject’s discrimination of the musical notes is maily based on the frequency difference between notes. Figure 2 shows that subject’s performance is all above chance level and monotonically increase when the number of semitone is higher than 2. The correct reponse to pairs of same notes (frequency difference is 0) is higher than 50%, which means subjects have the ability to detect the pairs of identical notes.

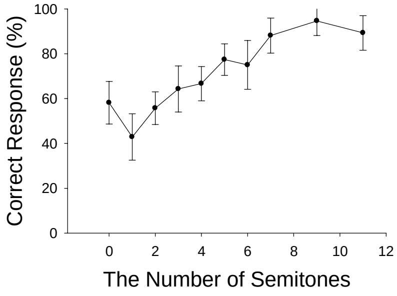

**Figure 2.** The correlation between musical note discrimination and the frequency difference. Error bars represent standard error for subject’s performance.

**Objective analysis of musical note vocal production**

F0 and duration of musical notes and subject’s vocal production of notes were extracted using signal processing software “STRAIGHT”. Figure 3 show the F0 of vocal production of musical notes in CI children. The average F0 of subject’s vocal production of the musical notes does not match the F0 original musical notes (Figure 3). The frequency difference between the average highest and lowest notes that subjects produced is only 14.2, which is only 8.9% of the original frequency difference between B4 and C4. Data show that CI users’ music note production has high frequency compression.

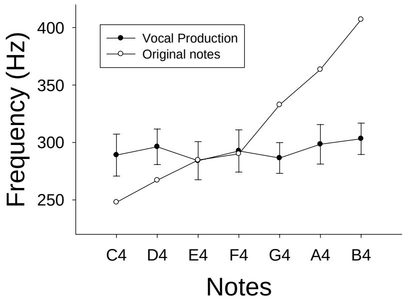

**Figure 3.** The comparison of F0 between vocal production in CI children and original musical notes. The filled circles are F0 of vocallized notes in CI users. Error bars represent Standard Error. Opened circles are the F0 of original musical notes which were played to subjects.

**Subjective analysis of musical note vocal production**

Musical notes produced by CI users were prelayed to normal hearing subjects who were asked to subjectively catigorize the sound samples into one of the five types of pitch contour, “flat”, “up”, “down”, “up and then down”, and “down and then up”. The percentage of correct catigorization was calculated. Chance level is 20%. Results show that CI children’s musical note production is poor, with the correct subjective catigorization of the pitch contour at 30% as the highest performance (Figure 4).

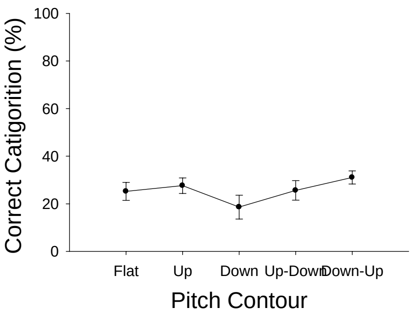

**Figure 4.** The subjective evaluation of the pitch contour of musical note production in children with cochlear implant. X axis is the types of pitch contour and Y axis is the percentage of correct catigorization of pitch contour of CI user’s vocal production by normal hearing subjects. Error bars represent Standard Error for subject’s performance.

Sound samples of CI users’ musical note production were segmented into pairs of adjacent notes. The note pairs were then presented to normal hearing subjects. Subjects were asked to evaluate the pairs of notes were the same or different. The pairs of notes were subgrouped according to the number of semitone difference. The number of correct responses was divided by the number of sound samples with the same number of semitone difference. Unlike musical note perception, subject’s vocal production performance is not correlated to the number of semitone differences (Figure 5).

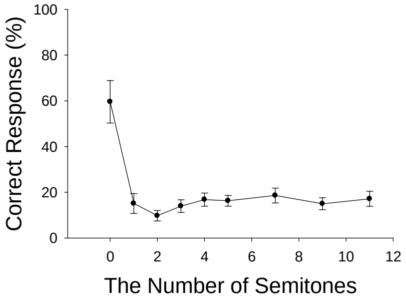

**Figure 5.** The correlation between musical note production and the frequency difference. Error bars represent Standard Error for subjects performance.

## Discussion

The present study shows that CI children can discriminate musical notes as long as the frequency difference between the notes is 2 or more semitones. The threshold of frequency discrimination for CI children is larger comparing to normal hearing children. Musical pitch perception performance increases as the function of frequency difference between notes. Previous studies have shown that frequency discrimination threshold for CI users to perceive the difference between piano notes is around 4 semitones (Fujita and Ito, 1999). In the present study, with frequency difference at 4 or more semitones, CI users can discriminate musical tones with over 60% correct response (Figure 2), which is in accordance with previous studies.

By analizing spontaneous singing in children with cochlear implant, Xu and colleagues have shown that the music vocal producation in CI children is poorer compare to normal hearing children. They found that the frequency difference between vocal produced notes and original musical notes is around 0.91-2.03 and 2.91-3.98 semitones respectively for normal hearing and cochlear implant children (Xu et al., 2009). Comparing to spontaneous singing, the procedure of the current experiment has a better control of the frequency pitch and pitch contours of the target musical notes of vocal production. Present study shows that CI users' musical note production is poor in conveying pitch information. For one octave musical notes from C4 to B4 which has frequency difference with 232 Hz, CI children could only produce an average pitch range at 14.2 Hz (Figure 3), suggesting that CI children may not be able to sing a melody in a correct way.

The present experiments allows us to make comparison between musical note perception and production performance. Although CI children can discriminate musical notes at above 60% correct percentage giving adequet frequency difference between notes. However, the musical note production performance does not agree with the perception data. Spearman correlation analysis result indicates that there was no significant correlation between the two types of performance, r=0.17, p>0.05 (Figure 2, Figure 5).

Our results indicate that CI children can perceive the frequency difference between musical notes, but can not vacally reproduce the sounds. Assume that CI users make attempts to convey the information they perceive in their vocal production, how do they convey the frequency difference? We observed that a few CI children tend to produce the higher pitch note with stronger sound level than the lower pitch note when the frequency difference between the two notes is large enough. Besides, one of the subjects produced the syllable "ma" as "wu" for all notes that higher than A4 (440 Hz). From musical note perception point of view, CI children might perceive the notes with different frequency as different sound level or syllable. The other possibility is that CI children tend to produce different sound level or syllable to represent the frequency difference they perceived from the musical notes. Further study need to manipulate both sound frequency, sound level, and syllables used in testing material to examine the hypothesis.

## ACKNOWLEDGEMENT

This research is supported by NSFC Grant 30670697 awarded to Dr. Juan Huang.

## REFERENCES

1. Xu L, Zhou N, Chen XW, Li YX, Schultz HM, Zhao XY, and Han DM. Vocal Singing by Prelingually-deafened children with cochlear implants. *Hearing Research. Vol. 255, pp. 129-134 (2009).*
2. Gfeller K, Lansing CR. Melodic, rhythmic, and timbral perception of adult cochlear implant users. *J. Speech Hear Res. Vol. 34, pp. 916–920 (1991).*
3. Nakata, et al. Pitch and Timing in the songs of deaf children with cochlear implants. *Music Percept.Vol. 24,pp. 147-154 (2006).*
4. Fujita S, Ito J. Ability of nucleus cochlear implantees to recognize music.*Ann Otol Rhinol Laryngol. Vol. 108, pp. 634-40 (1999).*
5. Niparko, et,al. Spoken Language Development in Children Following Cochlear Implantation, *J. Ame. Med. Associ (JAMA). Vol. 303, pp. 1498-1506(2010)*
6. McDermott HJ. Music perception with cochlear implants: a review. *Trends in Amplif. Vol. 8, pp. 49-82 (2004).*
7. Gfeller K, Lansing C. Musical perception of cochlear implant users as measured by the "Primary Measures of Music Audiation": An item analysis. *J Music Ther Vol. 29, pp. 18-39 (1992).*
8. P.W. Dawson, P.J. Blamey, S.J. Dettman, L.C. Rowland, E.J. Barker, E.A. Tobey, P.A. Busby, R.C. Cowan. A clinical report on speech production of cochlear implant users. *Ear and Hearing, vol. 16, pp. 551-561 (1995).*
9. M.G. Habib, S.B. Waltzman, B. Tajudeen, M.A. Svirsky. Speech production intelligibility of early implanted pediatric cochlear Implant users. *Int. J. Pediatr Otorhinolaryngol. May 14 (2010) [Epub ahead of print]*.
10. B.S. Wilson and M.F. Dorman. Cochlear implants: A remarkable past and a brilliant future. *Hearing Research, Vol. 242, pp 3–21(2008).*
11. K. Gfeller, C. Lansing. Musical perception of cochlear implant users as measured by the "Primary Measures of Music Audiation": An item analysis. *J Music Ther Vol. 29, pp. 18-39 (1992).*
12. Y.Y. Kong, R. Cruz, J.A. Jones, et al. Music perception with temporal cues in acoustic and electric hearing. *Ear Hear. Vol. 25, pp. 173-185 (2004).*
13. G.L. Nimmons, R.S. Kang, W.R. Drennan, J. Longnion, C. Ruffin, T. Worman, B. Yueh, J.T. Rubenstien. Clinical Assessment of Music Perception in Cochlear Implant Listeners, *Otol Neurotol. Vol. 29, pp. 149–155 (2008).*
14. E. Schulz and M. Kerber. Music perception with MED EL implants. *Advances in Cochlear Implants. Vienna: Manz, 326-332 (1994).*
15. B.S. Wilson, C.C. Finley, D.T. Lawson, R.D. Wolford, D.K. Eddington, and W.M. Rabinowitz. Better speech recogntion with cochlear implants. *Nature. Vol. 352, pp. 236-238 (1991).*