Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

# The modulation of tactile stimulation on Mandarin tone production in children with cochlear implant

**Ningyuan 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:** 43.70.Mn

## ABSTRACT

The impoverished auditory information provided by cochlear implant (CI) results in impaired auditory perception and production in mnay CI users. Speech perception in CI users has been well studied and fully documented. In contrast, much less has been studied for speech production in CI users. Tactile stimulation has been shown to enhance CI performance. In the present study, we aim to discover to what extent that tactile-enhanced speech perception can modulate speech production in CI users. Twelve Mandarin-Chinese speaking children with unilateral cochlear implant were tested under two conditions: vocalizing Mandarin syllables after listening to the audio playback of the same syllables with or withoutfeeling tactile stimulation of the same syllables applied to fingertip. Objective and subjective evaluations were conducted on the recorded Mandarin syllables produced by the subjects. The clarification rate (CLR) and correction rate (CRR) of subject's tone production evaluated by normal hearing subjects were analized. The perception index (Pt) that indexes the role of tone perception to tone production was calculated. The results show that subjects' tone production is clearer and more recognizable with the combined tactile and CI stimulation than CI stimulation alone, and the role of tone perception on tone production is higher in CI+Tactile condition than CI only condition. Our results provide important clues for designing efficient speech rehabilitation training program for CI users.

## INTRODUCTION

The cochlear implant (CI) has been successful in restoring hearing to people with profound deafness. In recent years, there has been a boom in the number of CI user in China. Mandarin Chinese, a tonal language, uses voice pitch (also called fundamental frequency, F0) to signal a difference in meaning between words (Avery, 1992). For example, depending on the pitch variation, the same syllable, /huang/ in Mandarin Chinese could mean 'desert' (荒，Tone-1), 'yellow' (黄，Tone-2), 'lie' (谎，Tone-3), or 'swing' (晃，Tone-4), respectively. Unfortunately, the speech processing strategy that is commonly used in cochlear implants does not explicitly extract and transmit voice pitch information (Wilson et al., 1991, Zeng, 2008). Tonal language speaking CI users thus have more difficulties in speech perception than their counterparts who speak non-tonal languages.

Low-frequency acoustic stimulation (<500 Hz) can enhance cochlear-implant (CI) performance in noise and in pitch-related tasks, but it is not available to most CI users. Tactual presentation of F0 information have shown to enhance the perception of frequency contours and speech pattern contrasts in normal hearing subject (Breevwer & Plomp, 1986; Boothroyd, 1985; Grant et al., 1986; Waldslein & Boothroyd, 1995; Kishon-Rabin, 1996; Hanin et al., 1988Hnath-Chisolm & Medwetsky, 1988). Studies have shown that low-frequency tactile stimulation enhances Mandarin Tone recognition in both CI users and normal hearing subjects while listening to CI simulated sound (Huang et al., 2008). Presenting low-frequency information through tactile aid can compensate for the lack of explicit transmission of low-frequency information in cochlear implant, thus enhance cochlear implant users in Mandarin Tone perception.

Previous studies have shown that Mandarin Tone information is mainly carried by the variation of F0, the first and second formants (Liang, 1963). Other studies have indicated that temporal envelope and the variation of sound level along time domain for a syllable also contain mandarin tone information (Whalen and Xu, 1992; Fu and Zeng, 1998; Fu and Zeng, 2000; Xu et al., 2002).

To provide evidences of tactile stimulation on cochlear implant rehabilitation training, here we tested weather the effect of tactile stimulation on Mandarin tone perception could generate better tone production so that benefit speech communication in CI users. We recorded Manrarin tone production at two conditions, CI only and CI+Tactile conditions. For the former, CI users produce Mandarin tones following listening to the words while seeing some pictures. For the latter, CI users produce Mandarin tones following listening to the words and feeling tactile stimulation of the same words from their fingertip. F0, sound level, the first and second formants of the recorded signals were analized. We aim to examine how tactile stimulation may influence Mandarin tone production in children with CI.

## METHOD AND PROCEDURE

### Subject

Twelve prelingually deaf Mandarin-speaking children with a uni-lateral cochlear implant participated in the experiments.

Four of them are female and 10 are male, aged between 5-8.5 years old. The duration of using cochlear implant is between 2 to 50 months. Subjects were asked to vocalize Mandarin syllables after listening to the playback of the same syllables or vocalize Mandarin syllables after listening to the playback of the syllables while feeling tactile stimulation of the same syllables applied to their fingertip. All subjects signed consent forms provided by parents and teachers.

## Testing material

**The selection of Chinese words**

Words used in the experiment are 32 mono-syllable Chinese characters selected from a list of words that usually occurr in daily speech communication of toddlers. The frequency of the characters occuring in daily life use is similar. There are 8 Chinese characters for each tone.

**Pictures of the Chinese words**

Thirty two pictures corresponding to the words were also used in the experiments. Each picture contains three parts of information, image, Chinese character, and PinYin (Figure 1), which helps young subjects recognize the words in order to vocalizing the sylables.


**Figure 1.** Sample visual stimuli used in the experiments

**Testing Sound of the Chinese words**

Neospeech TTS synthesizer was used to generate the Chinese characters in female Mandarin speech sound (Lily, 2009) with the lowest speed. To exclude the duration of the tones that might be used in tone perception (Fu and Zeng, 2000), the duration of each word was normalized to the average duration of the 32 words, about 658.9 ms. The duration of each tones was 634.6, 647.1, 713.4, and 689.8 ms, respectively, after normalization.

**Tactile stimulation**

Tactile stimuli were generated by extracting F0 from the sound of the Chinese words by STRAIGHT software with sampling rate at 22.1Hz and filter band is at 40-800 Hz. The frequency range of the tactile stimuli is at 123.37Hz-346.77Hz，average frequency is 220.15 Hz with Standard Deviation at 46.01 Hz (Figure 2).

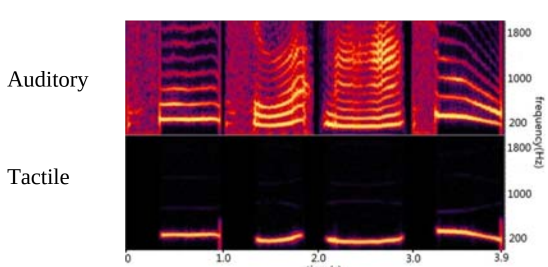


**Figure 2.** Sample spectrum of auditory and tactile stimuli. From the left to the right, “星”, “蛇”, “马”, and “树”

## 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 sat 1 meter away in front of the speaker which was placed at the same azimuth level as their ears. Tactile stimulator is model VBW32 (Tactaid). The stimulator is fixed at a subject’s left index finger tips. The tactile stimulation level was adjusted to the most comfortable level for each subject. Both auditory and tactile signals are delivered by an experimental interface to a subject, with auditory signal from the right channel and tactile signal from the left channel. The tactile signal goes through an amplifier PCA1 (Pyle Pro) and is used to drive a tactile stimulator.

Subjects were asked to vocalizing the syllables following listening to the same syllable while seeing the image of the syllable for CI only condition. For CI+Tactile condition, subjects also felt the vibration at their fingertip corresponding to the same syllable while seeing the picture and listening to the sound. Recordings under the two conditions were done in a random order.

## Data analysis

**Extraction of F0, formants, and sound level**

Recorded subjects’ vocal production signals were manually segmented into single words through the Adobe Audition 3.0 (Adobe systems Inc.) signal processing software. F0, the first formant, the second forman, and the sound level of each word were extracted by using speech processing software Praat (Boersma and Weenink, 2005) for further statistical analysis. Frequency filtering band was set at 80-600 Hz which covers the F0 range of male, female, and children voice pitch (Yang &Chi, 1995).

**Objective evaluation of Chinese tones**

A three-layered feedforward artificial neural network was (ANN) trained to recognize the four tonal patterns of subject’s speech production (Lan, and Zeng, 2004). The correct recognition of Chinese tones by the ANN is 100% for the training speech signals and 80% for speech production of normal hearing adults. Two hundreds Chinese words of various tonal variations were used to train the ANN with four output neurons corresponding to the four tones in Mandarin Chinese. The largest output of ANN corresponding to one out of the four possible tones was taken as the result of the objective evaluation of Chinese tone in the speech production samples.

**Subjective evaluation of Chinese tones**

Five normal hearing Mandarin-speaking undergraduate students (aged 21-23 year, 4 male and 1 female) were recruited to subjectively evaluate the Chinese tone of the speech production signals. Each speech signal is presented to them two times in a random order. Subjects were asked to verify the tone carried by the speech signal. The Clarification Rate (CLR) and Correction Rate (CRR) of the speech production were defined as follows, CLR equals to the percentage of the number of same tone evaluated by all 5 normal hearing students over the total number of words (32). CRR equals to the percentage of the number of same tone that is correct over the total number of each tone (8).

**Statistical Analysis**

The paired t-test was used to detect the difference in the performance between CI only (A) and CI+Tactile (AT) conditions.

### Results

From the objective evaluation data by ANN, 7 out of 12 subjects benefited from tactile stimulation in Chinese tone recognition task. The percentage of correct tone production increased. From subjective evaluation data, 5 out of 12 subjects' correct tone production increased. Paired t-test did not yield significant difference between A and AT conditions (subjective evaluation: p = 0.165; objective evaluation: p = 0.1, see Figure 3).

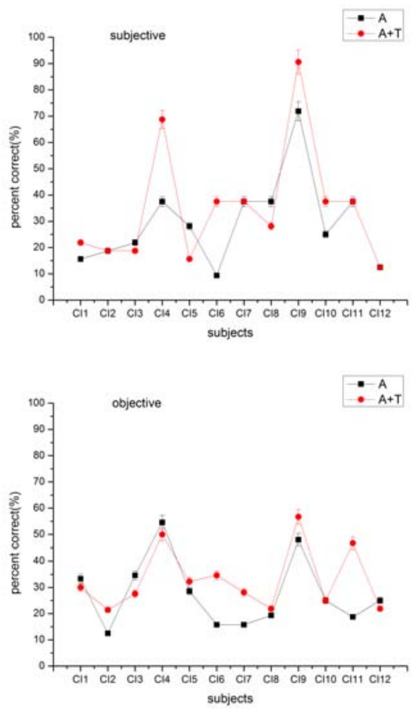


**Figure 3.** The average performance of correct tone production under cochlear implant only (auditory, A) and cochlear implant with tactile stimulation (auditory combined with tactile, AT) conditions from subjective (up panel) and objective (down panel) evaluations.

However, if we separate the data of correct tone production into 4 groups, tone 1, 2, 3, and 4, respectively. We observed that the influence of tactile stimulation on tone production is different across the 4 tones. Tactile stimulation significantly enhanced the production of tone 3 and tone 4, but did not change the performance for tone 1 and tone 2 (see Figure 4). The percentage of correct tone production evaluated by normal hearing subjects and the ANN are: Tone 3, sub-A: 29.1%, sub-AT: 39.5%; obj-A: 1.4%, obj-AT10.3%. Tone 4, sub-A: 17.7%, sub-AT: 30.2%; obj-A: 37.3%, obj-AT: 50%.

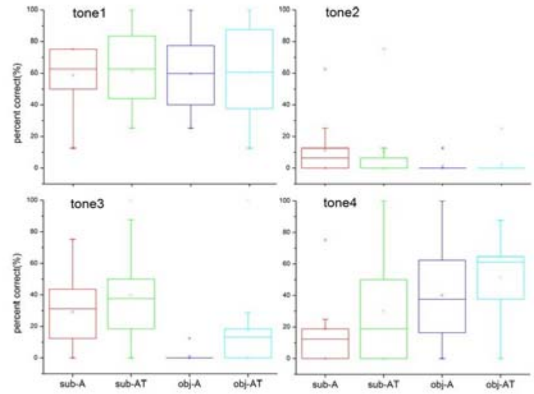


**Figure 4.** The tone production of 4 tones under auditory and auditory combined with tactile stimulation conditions. Sub-A and Sub-AT reprent subjective evaluation of tone production under CI only and CI+Tactile stimulation conditions, respectively. Obj-A and Obj-AT reprent subjective evaluation of tone production under CI only and CI+Tactile stimulation conditions, respectively. The Yaxis reprents the percentage of correct production.

Pearson correlation analysis shows that there is significant positive correlation in Tone 1, Tone 3, and Tone 4, between clarification and correction rates under both CI and CI+Tactile conditions (Table 1).

**Table 1** The correlation analysis between Clarification Rate and Correction Rate under CI only and CI+Tactile conditions.

| | A | AT |
|---|---|---|
| Tone1 | r=0.914 p=3.1e- 5 | r=0.900 p=6.6e- 5 |
| Tone2 | r=0.477 p=0.116 | r=0.410 p=0.185 |
| Tone3 | r=0.710 p=0.009 | r=0.878 p=1.7e- 4 |
| Tone4 | r=0.665 p=0.018 | r=0.792 p=0.002 |

Since the vocal tone production was recorded following both auditory (CI) and visual (picture) presentation of the target tones, we suspect that the visual cue also played a role in enhanceing tone perception in CI children, thus facilitated their tone production. We can define an index of tone perception as Pt Index = Ap/Vp to evaluate the enhancement of tactile stimulation to tone production, where Vp is the percentage of correct production of tones when subjects only see the pictures of the words and Ap is the percentage of correct production of tones when subjects only listen to the words. Data show that tactile stimulation significantly enhanced the Pt Index comparing to CI only condition (See Figure 5).

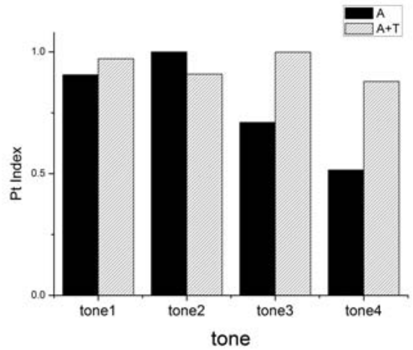


**Figure 5.** The comparision of perception index (Pt) between CI only and CI combined with tactile condition for each tone.

## DISCUSSION

Previous studies have shown that the perception of Tone 2 in Mandarin-speaking CI children is lower than Tone 1, Tone 3, and Tone 4 (Fu et al., 2004). Studies on the prediction of Chinese tones also indicate the vocal production of Chinese tone is not very well, with the correct production of Tone 2 significantly lower than other three tones, but the reason underlying such a phenomenon was not explored (Peng et al. 2004; Han et al., 2007). Our data indicate the same change direction. By adding tactile transformed F0 variation information to assist CI children in Chinese tone perception, the enhancement of tone production was also observed for Tone 3 and Tone 4. The results are not supprising. Tactile stimulation is effective in transducing sound level variation and F0 information. Tone 3 and Tone 4 have large F0 variations, and the sound level for these two tones is covariated with F0 and the direction of change are consistant between sound level and F0. Adding additional information of F0 and sound level variation through the tactile channel, subjects were able to perceive both tones more accurately, thus helped them in the vocal production performance. However, for Tone 2, the F0 increases from the onset to the offset, while the sound level decreases. This creates conflict cues for CI users to use in recognizing the tone. The same conflict is generated to tactile stimulation, which can explain the decreased tone production performance in some subjects when combined tactile stimulation was applied.

As shown in other studies, the three-layered feedforward artificial neural network used in the study provides a standard and objective evaluation method to measure tone production performance in CI children (Zhou et al., 2008). The positive correlation between clarification rate and correction rate in objective and subjective evaluation results provides the evidence of the effectiveness of the ANN. The interesting result of no correlation between clarification rate and correction rate in Tone 2 production further shows the conflicts between the variations of F0 and sound for the inconsistence, in which normal hearing people can recognize the tone better than the aritifical neural network.

The calculated perception index indicates the perception of tone under vision and vision combined with CI conditions. Our data again show that adding tactile stimulation enhances subject's perception of Tone 1, Tone 3, and Tone 4, which is in accordance with previous studies (Huang et al., 2008, 2009). In these studies, researchers show that tactile stimulation can enhance pitch related performance, familiar melody recognition, mandarin tone recognition, and speechd percpetion in noise, in CI users. This can explain the enhancement that tactile stimulation to tone production we observed in Tone 1, Tone 3, and Tone 4.

The present study has shown that tactilly presented F0 information can enhance Chinese tone perception, thus moify tone production in children with cochlear implant.

## ACKNOWLEDGEMENT

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

## REFERENCES

1 Boersma, P. and Weenink, D., "Praat: doing phonetics by computer (Version 5.1.26)," 2008 http://www.praat.org/.
2 Fu, Q. J., Hsu, C. J. and Horng, M. J., "Effects of Speech Processing Strategy on Chinese Tone Recognition by Nucleus-24 Cochlear Implant Users," *Ear & Hearing*, 2004, vol. 25, pp. 501–508.
3 Han, D., Zhou, N., Li, Y. X., Chen, X. W., Zhao, X. Y. and Xu, Li., "Tone production of Mandarin Chinese speaking children with cochlear implants," *Int J Pediatr Otorhinolaryngol*, 2007, vol. 71(6), pp. 875–880.
4 Huang, J., Chang, J., Zeng, FG. "Using vibrotactile stimulation to enhance Mandarin tone recognition in normal-hearing and cochlear-implant subjects," *31st MidWinter Meeting of the Association for Research in Otolaryngology* (2008).
5 Kong, Y. Y. and Zeng, F. G., "Temporal and spectral cues in Mandarin tone recognition," *J. Acoust. Soc. Am*, 2006, vol. 120, pp. 2830-2840.
6 Lan, N., Nie, K. B., Gao, S. K. and Zeng, F. G., "A Novel Speech-Processing Strategy Incorporating Tonal Information for Cochlear Implants," *IEEE Trans. Bio. Med. Eng*, 2004, vol. 51, pp. 752-760.
7 Liang, Z. A., "Tonal discrimination of Mandarin Chinese," *Acta Physiol. Sinica*, 1963, vol. 26, pp. 85–91.
8 Peng, S. C., Tomblin, J. B., Cheung, H., Lin, Y. S. and Wang, L. S., "Perception and Production of Mandarin Tones in Prelingually Deaf Children with Cochlear Implants," *Ear & Hearing*, 2004, vol. 25, pp. 251–264.
9 Xu, L., Tsai, Y. and Pfingst, B. E., "Features of stimulation affecting tonal-speech perception: Implications for cochlear prostheses," *J. Acoust. Soc. Am*, 2002, vol. 112, pp. 247–258.
10 X. J. Yang and H. S. Chi, Speech Signal Processing. *Beijjing: The Electronic Industry Press of China, 1995.*
11 Zeng, F. G., Rebscher, S., Harrison, W., Sun, X. and Feng, H., "Cochlear Implants: System Design, Integration, and Evaluation," *IEEE Rev. Bio. Med. Eng*, 2008, vol. 1, pp. 105-142.
12 Zhou, N., Zhang, W., Lee, C. Y., Xu, Li., "Lexical Tone Recognition with an Artificial Neural Network," *Ear and Hearing*, 2008, vol. 29(3), pp. 326-335.
13 Wilson, B.S., Finley, C.C., Lawson, D.T., Wolford, R.D. (1991) Eddington DK, Rabinowitz WM. Better speech recognition with cochlear implants. Nature 18;352(6332):236-238.
14 Avery, Peter and Susan Ehrlich. (1992) Teaching American English Pronunciation, Oxford University Press: Oxford.