Bimodal benefit in categorical perception of lexical tones for Mandarin-speaking children with cochlear implants

Hao Zhang[1,2], Jing Zhang[1], Hongwei Ding[1], Gang Peng[2]

[1]Speech-Language-Hearing Centre, Shanghai Jiao Tong University, Shanghai, China
[2]Research Centre for Language, Cognition, and Neuroscience, Department of Chinese and Bilingual Studies,
The Hong Kong Polytechnic University, Hong Kong SAR
zhang.hao@sjtu.edu.cn, elainezhang@sjtu.edu.cn, hwding@sjtu.edu.cn, gpengjack@gmail.com

ABSTRACT

Bimodal hearing with the combined use of a cochlear implant (CI) and a contralateral hearing aid (HA) has been demonstrated beneficial for deaf individuals in many aspects of speech perception. However, it remains inconsistent whether CI users can obtain bimodal benefit in lexical tone perception. To disentangle this question, Mandarin-speaking children using a CI and an HA in opposite ears were recruited to conduct perceptual tasks with a tonal continuum varying from Tone 1 to Tone 2. All participants were assessed with CI only and CI + HA conditions. Results showed typical S-shaped functions for the identification curves in both device conditions. Moreover, a sharper identification boundary and a higher peakedness score have been exhibited for the CI + HA relative to the CI only condition. The findings suggested that CI children on the whole show categorical perception for Mandarin tones and bimodal hearing could enhance their tonal categorization ability.

Keywords: bimodal hearing, categorical perception, Mandarin tones, cochlear implant, hearing aid.

1. INTRODUCTION

In Mandarin Chinese, tone recognition is crucial for speech perception, since it distinguishes lexical meanings of a word. Mandarin tones can be classified into four categories on the basis of pitch height and pitch contour: Tone 1 (T1, high-flat), Tone 2 (T2, mid-rising), Tone 3 (T3, falling-rising), and Tone 4 (T4, high-falling) [2]. Mandarin normal-hearing listeners can recognize tone categories effortlessly both in isolated and continuous speech [26].

Categorical perception (CP), in speech perception domain, refers to the phenomenon of discrete perceptual warping in a stimulus continuum with equal physical intervals, resulting in a better discrimination for stimuli across different phonemic categories than for stimuli within the same category [13]. CP indicates the fine-grained phonemic categorization ability, mapping the acoustic signals with infinite variability onto a finite set of phonemes.

In terms of lexical tone perception, ample evidence has demonstrated that Mandarin tone perception exhibits a typical CP for NH native adult listeners e.g. [18, 22-23], and for paediatric listeners e.g. [3, 28]. In a developmental study, Chen et al. [3] revealed that CP of Mandarin tones emerges no later than 4-year old for NH native children. Moreover, their tonal categorization ability gets refined gradually with the accumulating perceptual exposure of the tonal information as they grow older. However, it remains unclear whether Mandarin cochlear implanted children with a similar age can perceive Mandarin tones categorically.

Cochlear implant (CI) electrically stimulates the surviving auditory nerves by 12 to 22 implanted electrodes to partially restore hearing sensation for individuals with sever-to-profound hearing loss. However, the fundamental frequency (F0) of incoming sounds is weakly represented with these limited numbers of electrodes [4, 20]. Since the F0 information is the acoustic correlates for pitch, pitch perception poses a unique challenge for CI users. As a result, they perform poorly in pitch-related perceptual tasks, which has been documented in abundant research e.g. [6, 9, 18-19].

For CI recipients with some residual acoustic hearing, a potentially beneficial intervention is to fit a hearing aid (HA) in the non-implanted ear (i.e. bimodal hearing). The low frequency provided by an HA may complement the inadequate F0 information from a CI [11, 15, 27]. The bimodal hearing is postulated to benefit speech perception, the pitch perception in particular. The bimodal benefit, the improved performance in the CI + HA device condition in comparison with the CI only condition, is demonstrated in previous studies regarding prosody perception (e.g. [21]) and music perception (e.g. [11]). The reports on bimodal effect for lexical tone perception are, however, scarce and inconsistent. For lexical tone identification task, some studies found no bimodal benefit in quiet [12, 25], while others showed evident benefits [1, 14, 16]. In this investigation, classical tasks of CP were exploited to flesh out the bimodal effect for Mandarin tone perception. Given that the tonal CP tasks depend primarily on pitch cues,

the bimodal effect is hypothesized to be more robust in this study.

In summary, the present study aims to (1) examine the tonal CP in Mandarin children using a CI and a contralateral HA; and (2) delineate bimodal effect via comparing CP outcome of CI +HA with that of CI only. Findings of this study will shed some light on clinical practices for the paediatric deafness.

2. MATERIALS AND METHODS

2.1. Participants

Twelve native Mandarin children between 4.1 and 6.6 years (5 females; M = 5.3, SD = 0.7) who were implanted unilaterally and used a contralateral HA were recruited in this study (see Table 1 for detailed demographic information). The Hiskey–Nebraska Test of Learning Aptitude (HNTLA) [8] was adopted to confirm a normal nonverbal intelligence of each participant.

2.2. Materials

The Mandarin syllable /i/ with T1 and T2 were selected as the endpoints for the synthesis of the tonal continuum. /i/ with T1 recorded from a female native Mandarin-speaking adult served as the template token. Based on the template token, a series of seven tonal stimuli in the continuum were re-synthesized using the Pitch-Synchronous Overlap Add (PSOLA) implemented in Praat. The stimulus No. 1 and No. 7 were the two endpoints of the continuum, which represented T1 and T2 respectively. For each tonal stimulus, the F0 transited from the onset (ranging from 160 to 250 Hz, with a step of 15 Hz) to the offset (fixed at 250 Hz). F0 was the only acoustic cue for manipulation, and the amplitude and duration were kept constant across all stimuli to 70 dB SPL and 400 ms, respectively.

Table 1: Demographic information for the child participants. Ages and durations were in years.

Subject (Sex)	Age at test	Age at CI	CI duration	Age at HA	HA duration
1 (F)	5.4	3	2.4	3.7	1.7
2 (F)	4.7	1.2	3.5	1.1	3.6
3 (M)	5.7	1.7	4	1.5	4.2
4 (F)	5.6	3.3	2.3	2.4	3.2
5 (F)	5.8	1	4.8	3.7	2.1
6 (F)	5	1.4	3.6	3.3	1.7
7 (M)	4.1	0.9	3.2	0.6	3.5
8 (M)	5.6	2.7	2.9	2.2	3.4
9 (M)	5.3	2.7	2.6	3.6	1.7
10 (M)	4.7	1	3.7	0.6	4.1
11 (M)	4.7	1.4	3.3	2.6	2.1
12 (M)	6.6	1.4	5.2	4.6	2

2.3. Procedure

All participants were instructed to complete two tests, including identification task and discrimination task, in a sound-treated therapy room. The E-prime 2.0 software was used to control the tests. The tonal stimulus was delivered via a loudspeaker about one meter away from the participant. Supporting pictures were used for both tasks to ensure the children follow the task requirements.

For the identification task, children were firstly instructed to point to the matching pictures for different tones. The picture depicting a car driving on a level road stands for T1, while the other picture of a car driving on a rising road represents T2. Then a practice block with six stimuli was presented, which contains two endpoints of the tonal continuum. The formal test would start after an accuracy of 90% was obtained. The participant made a response by pointing at the matching picture while the experimenter pressed the corresponding key to record the participant's response. Two sessions were prepared with the seven tonal stimuli repeated five times in each session, resulting in a total of 70-stimulus presentation in a randomized order.

For the discrimination task, an instruction phase and a practice phase were presented before the formal test. In the formal test, 17 contrastive pairs were constructed with an inter-stimulus interval (ISI) of 500 ms. Among the 17 pairs, 10 of them were the different pair with two steps (i.e. 30 Hz) separating the two tonal stimuli in each pair (e.g. 1-3), and the other 7 were the same pair with a tonal stimulus pairing with itself (e.g. 1-1). All contrastive pairs were repeated three times within the first session and twice within the second one. In total, 85 pairs were randomly presented.

Both tasks were examined in two device conditions: CI only and CI + HA. Two conditions were tested in a counterbalanced order across the participants in separated days. Approximately, 35 minutes lasted for each device condition, and child participants were free to have breaks between or within test sessions.

2.4. Data analysis

In order to reveal the tonal CP performance in paediatric CI users with an HA in the non-implanted ear, identification score and discrimination score were calculated for statistical analysis.

The identification score was defined as the averaged percentage response for T1 or T2. Two key parameters of identification score were estimated: boundary position (i.e. 50% cross-over point) and boundary width (i.e. the linear distance between 25%

and 75%). Probit analysis [5] was adopted to assess the two parameters.

The discrimination score (P value) was defined as the accuracy rate in discrimination task, which was calculated according to the formula in Xu et al. [23]. Moreover, following the instructions described specifically in Chen et al. [3], we divided the discrimination accuracies into two types: the between-category accuracy and the within-category accuracy. In addition, the peakedness score was calculated as the difference between the two accuracy types, which represents the magnitude of the benefit for the between-category accuracy relative to the within-category accuracy [10].

3. RESULTS

3.1. Identification and discrimination curves

The identification and discrimination curves for Mandarin CI children using a contralateral HA were shown in Fig. 1 for different device conditions of CI only and CI + HA. As depicted in the figures, typical S shapes were shown for the identification curves. In addition, the categorical boundary and the corresponding discrimination peak were found prominent in both device conditions.

3.2. Position and width of categorical boundary

The estimated average boundary position for the CI only and CI + HA device conditions were 3.86 and 3.93, respectively. Results of paired-samples T test revealed that the categorical boundary positions were not significantly different between the two device conditions (t (11) = -0.69, p = 0.5). The mean boundary width for the CI condition was 2.11, whereas it was 1.51 for the CI + HA condition. A significant difference was found between the two conditions (t (11) = 3.08, p = 0.01). Noted that one outlier (see Fig. 2) was shown in the CI + HA condition. The difference between the two device conditions remains significant even when the outlier was removed (t (10) = 2.82, p = 0.018).
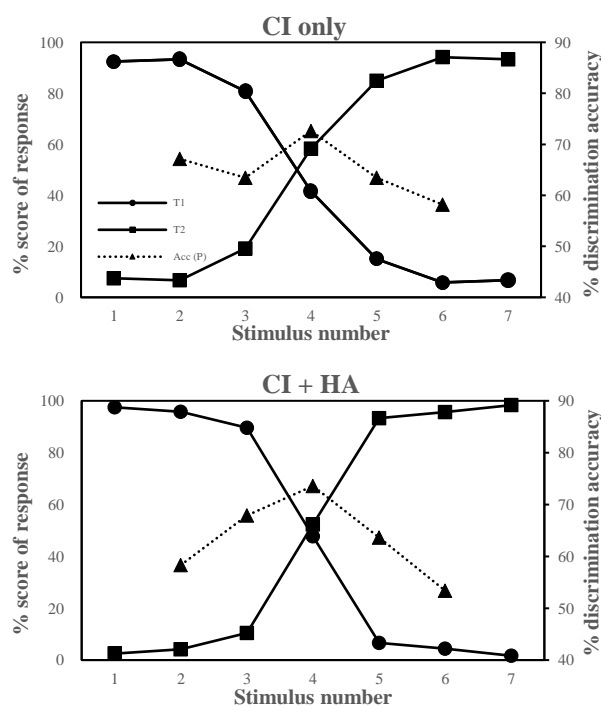
3.3. Discrimination accuracy

For the CI only condition, the discrimination peak was 72.64% at the tonal contrast of 3-5, which straddles the boundary position at 3.86. In addition, for the CI + HA condition, the discrimination accuracy reaches the maximum (i.e. 73.58%) at 3-5, and the discrimination peak also aligns well with the corresponding boundary position at 3.93.

The discrimination accuracies were further divided into the between-category type and the within-category type (see Fig. 3). The between-

category and within category accuracies for the CI only and CI + HA conditions, respectively, were 68.06%, 62.93%, 70.73%, and 58.43%. As a result, the averaged peakedness score was 5.13% for the CI only condition and 12.3% for the CI + HA condition. Repeated measures ANOVA with *device condition* and *accuracy type* as within-subject factor showed a main effect of *accuracy type* (F (1, 11) = 18.19, p = 0.001), but not the *device condition*, nor the interaction. In addition, the peakedness score was significantly higher in the CI + HA condition than in the CI only condition (t (11) = 2.56, p = 0.027).

Figure 1: Identification and discrimination curves for different device conditions. The left y-axis indicates the identification score of T1 or T2 response, while the right y-axis indicates the discrimination accuracy (ACC (P)).

Figure 2: Boxplots of boundary width for CI only and CI + HA conditions. The bold line inside the boxes indicates the median, and the upper and lower boundaries of the box mark the upper and lower quartiles.
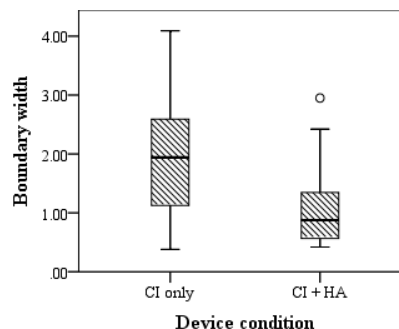
Figure 3: Discrimination accuracy of between-category type and within-category type for the two device conditions.
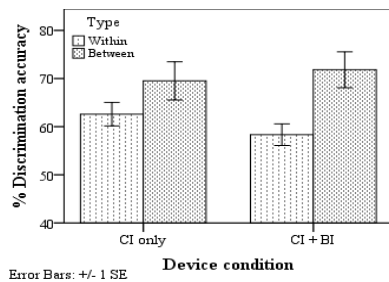
4. DISCUSSION

The current study investigated the tonal CP in Mandarin children using a CI and a contralateral HA. Characteristics of CP were exhibited in the perception of T1 and T2 for both the CI only and the CI +HA device conditions. Moreover, a significantly narrower boundary width and higher peakedness score were revealed for the CI + HA than for the CI only condition, suggesting a bimodal effect in the tonal categorization among Mandarin paediatric CI users.

The major characteristics of CP, as generalized by Liberman et al. [13], were a sharp identification boundary between two phonemic categories and a prominent discrimination peak straddling the identification boundary. In this study, typical S-shaped functions were shown in all identification curves. In addition, prominent discrimination peaks were well aligned with the corresponding categorical boundaries for both device conditions. Moreover, statistical analyses revealed a significantly higher discrimination accuracy for the between-category type than for the within-category type in both device conditions. All the results converged to a conclusion that Mandarin CI children can perceive Mandarin tones categorically. This finding echoes a report recruiting a group of Mandarin CI children with an age range of 4.1 to 5.5 years [7]. Although discrimination task was not involved in that study, the results showed sharp categorical boundaries for both continua of T1-T2 and T1-T4. Therefore, it is reasonable to conclude that Mandarin CI children around 5-year of age, as their NH age mates, can acquire CP ability for Mandarin tones as speech exposure is being gradually accumulated, regardless of the notorious limitations of the CI device in transmitting F0 information.

The binaural bimodal fitting is supposed to be beneficial for Mandarin tone perception, with the integration of electric hearing from the CI and the residual low-frequency acoustic hearing from the HA [15]. The bimodal benefit in Mandarin tone recognition has been verified evident in noise [12], [24]. In quiet, however, there are limited research with inconsistent results on whether CI users can receive bimodal benefit in lexical tone recognition. For instance, in the studies of Yuen et al. [25] and Li et al. [12], no bimodal effect was found. On the contrary, Chang et al. [1] and Luo et al. [14] both revealed a significant bimodal effect. Apparently, differences in testing materials, experimental paradigms, and subjects could account for the inconsistencies among previous studies.

Implementing a refined manipulation of F0 information, the present study showed a significantly narrower boundary width for the CI + HA than the CI only condition, indicating a sharper categorical boundary for the bimodal-hearing condition. The sharper boundary came from the higher identification acuity for the ambiguous stimuli in the middle part of the continuum, which suggested a more categorical-like perception. The bimodal benefit found in this study indicated that the binaural bimodal hearing could improve some Mandarin CI children's phonemic categorization ability in lexical tone identification. Noted that the tonal stimuli used in this study were manipulated only on the F0 information, while keeping all other acoustic cues constant. CI participants in this study were unable to take full advantage of the secondary tonal cues, such as duration and amplitude envelope, which are supposed to facilitate their tonal perceptual performance [18]. Therefore, the improved tonal categorization was due to the supplemental F0 information provided by the contralateral HA, which complemented the inadequate low-frequency resolution from the CI.

5. CONCLUSION

In conclusion, paediatric Mandarin CI users are able to extract useful pitch information and perceive lexical tones categorically, despite the well-known degraded spectral and temporal signals provided by the CI device. Moreover, the additional HA in the non-implanted ear evoked more categorical-like perception for Mandarin tones, suggesting an improved tonal categorization with the integration of residual low-frequency acoustic hearing and the electric hearing. The findings support the clinical trial of fitting a contralateral HA in the non-implanted ear for potential benefits in lexical tone perception among native Mandarin-speaking CI children.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] Chang, Y., Chang, R. Y., Lin, C.-Y., Luo, X. 2016. Mandarin tone and vowel recognition in cochlear implant users: Effects of talker variability and bimodal hearing. *Ear Hear.* 37, 271.

[2] Chao, Y. R. 1948. *Mandarin primer: An intensive course in spoken Chinese*. Harvard University Press.

[3] Chen, F., Peng, G., Yan, N., Wang, L. 2017. The development of categorical perception of Mandarin tones in four-to seven-year-old children. *J. Child Lang.* 44, 1413–1434.

[4] Croghan, N. B. H., Duran, S. I., Smith, Z. M. 2017. Re-examining the relationship between number of cochlear implant channels and maximal speech intelligibility. *J. Acoust. Soc. Am.* 142, EL537-EL543.

[5] Finney, D. J. 1971. *Probit analysis*. New York, Ny, 10022, 32.

[6] Good A. *et al.* 2017. Benefits of music training for perception of emotional speech prosody in deaf children with cochlear implants. *Ear Hear.* 38, 455.

[7] Gu, W., Yin, J., Mahshie, J. 2017. Production of sustained vowels and categorical perception of tones in Mandarin among cochlear-implanted children. *Proc. Interspeech 2017*. 1869–1873.

[8] Hiskey, M. S. 1966. *Hiskey-Nebraska test of learning aptitude*. Union College Press.

[9] Holt, C. M., Lee, K. Y. S., Dowell, R. C., Vogel, A. P. 2018. Perception of Cantonese Lexical Tones by Pediatric Cochlear Implant Users. *J. Speech, Lang. Hear. Res.* 61, 174–185.

[10] Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., Yang, Y. 2012. Impaired categorical perception of lexical tones in Mandarin-speaking congenital amusics. *Mem. Cognit.* 40, 1109–1121.

[11] Kong, Y.-Y., Stickney, G. S., Zeng, F.-G., 2005. Speech and melody recognition in binaurally combined acoustic and electric hearing. *J. Acoust. Soc. Am.* 117, 1351–1361.

[12] Li, Y., Zhang, G., Galvin III, J., Fu, Q.-J. 2014. Mandarin speech perception in combined electric and acoustic stimulation. *PLoS One*. 9, e112471.

[13] Liberman, A. M., Harris, K. S., Hoffman, H. S., Griffith, B. C. 1957. The discrimination of speech sounds within and across phoneme boundaries. *J. Exp. Psychol.* 54, 358-368.

[14] Luo, X., Chang, Y., Lin, C.-Y., Chang, R. Y. 2014. Contribution of bimodal hearing to lexical tone normalization in Mandarin-speaking cochlear implant users. *Hear. Res.* 312, 1–8.

[15] Luo X., Fu, Q.-J. 2006. Contribution of low-frequency acoustic information to Chinese speech recognition in cochlear implant simulations. *J. Acoust. Soc. Am.* 120, 2260–2266.

[16] Mok, M., Holt, C. M., Lee, K., Dowell, R. C., Vogel, A. P. 2017. Cantonese tone perception for children who use a hearing aid and a cochlear implant in opposite ears. *Ear Hear*. 38, e359–e368.

[17] Peng, G., Zheng, H.-Y., Gong, T., Yang, R.-X., Kong, J.-P., Wang, W. S.-Y. 2010. The influence of language experience on categorical perception of pitch contours. *J. Phon.* 38, 616–624.

[18] Peng, S.-C., Lu, H.-P., Lu, N., Lin, Y.-S., Deroche, M. L. D., Chatterjee, M. 2017. Processing of acoustic cues in lexical-tone identification by pediatric cochlear-implant recipients. *J. Speech, Lang. Hear. Res.* 60, 1223–1235.

[19] Peng, S.-C., Tomblin, J. B., Cheung, H., Lin, Y.-S., Wang, L.-S. 2004. Perception and production of Mandarin tones in prelingually deaf children with cochlear implants. *Ear Hear.* 25, 251–264.

[20] Shannon, R. V., Fu, Q.-J., Galvin III, J. 2004. The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. *Acta Otolaryngol.* 124, 50–54.

[21] Shpak, T., Most, T., Luntz, M. 2014. Fundamental frequency information for speech recognition via bimodal stimulation: Cochlear implant in one ear and hearing aid in the other. *Ear Hear.* 35, 97–109.

[22] Wang, W. S.-Y. 1973. The Chinese language. *Sci. Am.* 228, 50–63.

[23] Xu, Y., Gandour, J. T., Francis, A. L. 2006. Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *J. Acoust. Soc. Am.* 120, 1063–1074.

[24] Yang H.-I., Zeng, F.-G. 2017. Bimodal benefits in Mandarin-speaking cochlear implant users with contralateral residual acoustic hearing. *Int. J. Audiol.* 56, S17–S22.

[25] Yuen, K. C. P., Cao, K., Wei, C., Luan, L., Li, H., Zhang, Z. 2009. Lexical tone and word recognition in noise of Mandarin-speaking children who use cochlear implants and hearing aids in opposite ears. *Cochlear Implants Int.* 10, 120–129.

[26] Zhang, C., Peng, G., & Wang, W. S.-Y. 2013. Achieving constancy in spoken word identification: Time course of talker normalization. *Brain and language*, *126*(2), 193-202.

[27] Zhang, T., Dorman, M. F., Spahr, A. J. 2010. Information from the voice fundamental frequency (F0) region accounts for the majority of the benefit when acoustic stimulation is added to electric stimulation. *Ear Hear.* 31, 63.

[28] Zhang Y. *et al.* 2012. Universality of categorical perception deficit in developmental dyslexia: an investigation of Mandarin Chinese tones. *J. Child Psychol. Psychiatry*, 53, 874–882.