

Bimodal Benefits Revealed by Categorical Perception of Lexical Tones in Mandarin-speaking Kindergarteners with a Cochlear Implant and a Contralateral Hearing Aid

Hao Zhang ^{a, b}, Jing Zhang ^a, Gang Peng ^b, Hongwei Ding ^{a, *}, Yang Zhang ^{c, **}

^a Speech-Language-Hearing Center, School of Foreign Languages, Shanghai Jiao Tong University,
Shanghai, China

^b Research Centre for Language, Cognition, and Neuroscience, Department of Chinese and Bilingual
Studies, The Hong Kong Polytechnic University, Hong Kong SAR, China

^c Department of Speech-Language-Hearing Sciences and Center for Neurobehavioral Development,
University of Minnesota, Minneapolis, MN, USA

*** Corresponding author:** Speech-Language-Hearing Center, School of Foreign Languages,
Shanghai Jiao Tong University, Shanghai, China.

**** Corresponding author:** Department of Speech-Language-Hearing Sciences and Center for
Neurobehavioral Development, University of Minnesota, Minneapolis, MN, USA.

E-mail addresses: hwding@sjtu.edu.cn (H. Ding), zhanglab@umn.edu (Y. Zhang).

Conflict of Interest: The authors have declared that no competing interests existed at the time of
publication.

Funding Statement: This study was supported by grants from the Major Program of National Social
Science Foundation of China (No. 18ZDA293). Y. Zhang received additional support from the
University of Minnesota's Brain Imaging Grant and Grand Challenges Exploratory Research Grant
for international collaboration.

20 **Abstract**

21 **Purpose:** Pitch reception poses challenges for individuals with cochlear implants (CIs), and adding a
22 hearing aid (HA) in the nonimplanted ear is potentially beneficial. The current study used fine-scale
23 synthetic speech stimuli to investigate the bimodal benefit for lexical tone categorization in
24 Mandarin-speaking kindergarteners using a CI and an HA in opposite ears.

25 **Method:** The data were collected from 16 participants who were required to complete two classical
26 tasks for speech categorical perception (CP) with CI + HA device condition and CI alone condition.
27 Linear mixed-effects models were constructed to evaluate the identification and discrimination scores
28 across different device conditions.

29 **Results:** The bimodal kindergarteners showed CP for the continuum varying from Mandarin Tone 1
30 and Tone 2. Moreover, the additional acoustic information from the contralateral HA contributes to
31 improved lexical tone categorization, with a steeper slope, a higher discrimination score of between-
32 category stimuli pair, and an improved peakedness score (i.e., an increased benefit magnitude for
33 discriminations of between-category over within-category pairs) for the CI + HA condition than the
34 CI alone condition. The bimodal kindergarteners with better residual hearing thresholds at 250 Hz
35 level in the nonimplanted ear could perceive lexical tones more categorically.

36 **Conclusion:** The enhanced categorical perception results with bimodal listening provide clear
37 evidence for the clinical practice to fit a contralateral HA in the nonimplanted ear in kindergarteners
38 with unilateral CIs with direct benefits from the low-frequency acoustic hearing.

39 **Keywords:** lexical tones, categorical perception, Mandarin-speaking children, cochlear implant,
40 hearing aid, bimodal benefits

41

1 Introduction

Lexical tones are an essential feature of tonal languages such as Mandarin Chinese and Cantonese that use tonal variations to represent different words. Based on differences in pitch height and pitch contour, Mandarin tones are traditionally classified into four categories: Tone 1 (T1) with a high-flat pitch, Tone 2 (T2) with a mid-rising pitch, Tone 3 (T3) with a falling-rising pitch, and Tone 4 (T4) with a high-falling pitch (Chao, 1948). Recognizing the differences in lexical tones primarily depends on the fundamental frequency (F0), commonly referred to as the acoustic correlate for pitch patterns in speech sounds (Whalen & Xu, 1992). This places greater reliance on pitch information in speech perception for tonal language speakers than nontonal language users.

While cochlear implants (CIs) provide significant benefits for the recipients' speech perception and language skills (Peterson, Pisoni, & Miyamoto, 2010; van Wieringen & Wouters, 2015 for reviews), pitch perception poses a unique challenge for this clinical population. A body of prior studies have established CI users' deficits in pitch-related perceptual tasks, including voice emotion perception (e.g., Chatterjee et al., 2015; Jiam, Caldwell, Deroche, Chatterjee, & Limb, 2017; Nakata, Trehub, & Kanda, 2012; Paquette et al., 2018), speech prosody recognition (e.g., Chatterjee & Peng, 2008; S.-C. Peng, Tomblin, & Turner, 2008; Van De Velde et al., 2019), music perception (e.g., Gu et al., 2017; McDermott, 2004; Tao et al., 2015; S. Wang et al., 2012), speaker gender identification (e.g., Cullington & Zeng, 2011; Fu, Chinchilla, Nogaki, & Galvin III, 2005; Meister, Fürsen, Streicher, Lang-Roth, & Walger, 2016; Meister, Landwehr, Pyschny, Walger, & Wedel, 2009), and lexical tone perception (e.g., Y. Chen, Wong, Chen, & Xi, 2014; Gu et al., 2017; Holt, Lee, Dowell, & Vogel, 2018; Morton, Torriane Jr, Throckmorton, & Collins, 2008; S.-C. Peng et al., 2017; S.-C. Peng, Tomblin, Cheung, Lin, & Wang, 2004). The deficits may result from limitations of the contemporary CI multichannel technology. Current CI devices electrically stimulate the surviving auditory nerves of the recipients with severe-to-profound hearing loss by 12 to 22 implanted

66 electrodes to partially restore their hearing sensations. However, only degraded spectral-temporal
67 signals are extracted through the CI device because of the limited number of electrodes, limited depth
68 of the implanted electrodes, channel interactions, and broad analysis filters (Carroll & Zeng, 2007;
69 Dorman et al., 2007; Shannon, Fu, & Galvin III, 2004). The coarse spectral-temporal resolution,
70 therefore, is insufficient in transmitting and representing pitch cues. The primary goal of this study
71 was to examine a widely advocated intervention/practice of fitting an additional hearing aid in the
72 nonimplanted ear (i.e., bimodal hearing configuration) to ameliorate pitch deficits and improve
73 speech perception for individuals with unilateral CIs, via the investigation of the categorical
74 perception (CP) of lexical tones using fine-controlled speech stimuli.

75 In the speech domain, CP represents a fundamental mechanism supporting the abstraction and
76 cognitive organization of phonetic categories, perceptually mapping the infinitely variable spoken
77 signals from different speakers and contexts onto a finite set of discrete phonological symbols, which
78 is strongly influenced by sensory deprivation and language experience (Y. Zhang, 2016 for a review
79 on CP of lexical tones). In theories and models of language development, phonetic categorization
80 constitutes the foundational bedrock for the acquisition of higher-order linguistic structures and
81 communication skills (Bidelman, Moreno, & Alain, 2013; Kuhl, Patricia & Rivera-Gaxiola, 2008;
82 Kuhl, Conboy, Padden, Nelson, & Pruitt, 2005; G. Shen & Froud, 2019; X. Wang, Wang, Fan,
83 Huang, & Zhang, 2017; Y. Zhang, Kuhl, Imada, Kotani, & Tohkura, 2005). Experimentally, the
84 hallmark of CP is captured by better discrimination for stimulus contrasts across different categories
85 than for contrasts within the same category along a phonetic continuum with equal physical intervals
86 (Kuhl, 1991; Liberman, Harris, Hoffman, & Griffith, 1957). Ample evidence has documented a
87 typical CP for the Mandarin tone continuum of T1 and T2 in native normal-hearing (NH) listeners,
88 showing a sharp identification boundary of the two tonal categories, together with a prominent
89 discrimination peak straddling the boundary position (e.g., F. Chen, Peng, Yan, & Wang, 2017; G.

90 Peng et al., 2010; W. S. Y. Wang, 1973; Xu, Gandour, & Francis, 2006). The emergence and
91 maturation of the CP of lexical tones is a remarkable phenomenon over the process of phonetic
92 learning for typically developing children from Mandarin Chinese. In a developmental study, Chen et
93 al. (2017) demonstrated that the CP of Mandarin tones emerges by four years of age or even younger
94 for typically developing children with NH. In addition, the lexical tone categorization matures
95 gradually with the accumulating exposure of the ambient tonal language inputs and achieves
96 saturation comparable to adults by the age of six (F. Chen et al., 2017). In view of deprivations of
97 auditory function in early life and deficits of the contemporary CI technology in pitch encoding, it is
98 of great interest to investigate whether native children with CIs can develop a typical CP of Mandarin
99 tones.

100 Despite the extensive research on the CP of speech sounds including lexical tones, the tonal CP
101 has rarely been investigated among CI recipients. The only exception was the study of Luo et al.
102 (2014), in which Mandarin-speaking CI users and NH listeners were enrolled to identify a tonal
103 continuum ranging from T1 to T2. The results revealed a typical S-shaped identification function for
104 the CI group, which was morphologically similar to the NH control group. However, individual
105 performance needs further inspection in addition to the overall group outcome, considering the
106 inherent heterogeneity in this clinical population. In addition, their study did not use the
107 discrimination task to test whether the CI participants showed enhanced across-category
108 discrimination relative to within-category discrimination. Thus, Luo et al.'s (2014) study did not fully
109 address whether the CI users were able to develop a typical CP of lexical tones. Given the pivotal
110 role of phonetic categorization in language development, it is of theoretical importance to investigate
111 the CP performance in the context of bimodal hearing configuration for the pediatric recipients and to
112 examine its potential value for proper diagnostic evaluation and efficacious intervention.

Rehabilitative practices have demonstrated tangible benefits in auditory rehabilitation by means of bimodal hearing for unilaterally implanted candidates with some residual acoustic hearing. Bimodal hearing refers to the combination of two different hearing configurations—electrical hearing with a CI in one ear and acoustic hearing with a hearing aid (HA) in the contralateral ear. Indeed, successful bimodal listeners could obtain substantial improvements in speech understanding with CI + HA stimulation relative to with a unilateral CI alone (e.g., Ching, Psarros, Hill, Dillon, & Incerti, 2001; Dorman et al., 2007; Morera et al., 2005; Shpak, Most, & Luntz, 2014). The improved perceptual outcome in the bimodal stimulation over the CI alone condition is referred to as the bimodal benefit, and it has been established in a great deal of studies from nontonal languages. Several research groups evaluated the bimodal benefit adopting task-specific tests in the perception of segmental linguistic features that contain low-frequency component, such as voicing, semivowels, and nasals (e.g., Ching et al., 2001; Mok, Grayden, Dowell, & Lawrence, 2006; Most, Gaon-Sivan, Shpak, & Luntz, 2012), or in the perception of suprasegmental linguistic features, such as intonation, emphasis, and stress (e.g., Cullington & Zeng, 2011; Most, Harel, Shpak, & Luntz, 2011; Shpak et al., 2014). It is tempting to attribute the observable group-level benefits of speech perception in these studies to the combination of high-frequency components from the electric stimulation with low-frequency components from the acoustic stimulation. Nevertheless, the great heterogeneity in the CI population reveals enormous individual variability with respect to the degree of benefits from bimodal hearing. Some subjects tend to benefit significantly from the additional HA in the opposite ear, whereas others receive little-to-no bimodal benefit. Meanwhile, some individuals even report an interference effect with the addition of a contralateral HA (Gifford & Dorman, 2019; Neuman et al., 2017). The lack of solid evidence for the contributing factors of the variability in bimodal benefit at the individual level leaves the topic open for further inspection.

Unlike studies on nontonal language speakers, there is a paucity of research on the benefits of an additional HA in Mandarin-speaking bimodal listeners. A handful of related studies have shown a mixed result (e.g., Y. Li, Zhang, Galvin III, & Fu, 2014; Luo, Chang, Lin, & Chang, 2014; Mok, Holt, Lee, Dowell, & Vogel, 2017; H.-I. Yang & Zeng, 2017; H. Zhang, Zhang, Ding, & Zhang, 2020). In particular, some studies showed that using a contralateral HA in the nonimplanted ear was incapable of improving lexical tone perception in quiet for Mandarin-speaking bimodal listeners (Y. Li et al., 2014; H.-I. Yang & Zeng, 2017). Li et al. (2014) suggested the smaller-than-expected bimodal benefit might attribute to the available duration and amplitude cues that probably weaken the dependence of F0 patterns for lexical tone recognition. Conversely, there have been some reports revealing significant bimodal benefit for tonal perception scenarios that depend more on pitch cues (e.g., Luo et al., 2014; Mok et al., 2017). For example, in the study of Mok et al. (2017), improved Cantonese tone recognition was found in the CI + HA device condition over the CI alone condition for eight native bimodal children. The bimodal benefit is reasonably robust in the perception of Cantonese tones that are represented almost exclusively by pitch cues (Yip, 2002). In our recent study, it was demonstrated that using an HA in the contralateral ear could help Mandarin-speaking children with CIs alleviate the recognition confusion between T2 and T3 (H. Zhang et al., 2020). Given the CI device-related deficits in encoding pitch information, it is reasonable to expect that the additional acoustic signals from an HA would complement the pitch deficits and reflect the subtle acoustic differences between the two tones. Although the definitive cues driving the bimodal benefits remain elusive, the F0 information has been hypothesized to account for the majority of bimodal benefits in speech perception when the acoustic signals are added to the electric stimulation (T. Zhang, Dorman, & Spahr, 2010). The CP paradigm with well-controlled manipulations on acoustic cues (e.g., pitch, duration, and amplitude cues) holds the great promise to allow us to test this hypothesis, which has not been previously applied to bimodal kindergarteners and can invite further exploration on the underlying mechanisms of bimodal benefits.

The main incentive for this study is to employ the classical CP paradigm with fine-grained acoustic control of a synthetic speech continuum to investigate the bimodal benefit for Mandarin tone categorization in native kindergarteners who use CI in one ear and HA in the other. There are three specific aims: (1) To assess the CP of Mandarin tones with both identification and discrimination tasks in the kindergartener participants with hearing impairment, (2) To investigate the bimodal benefit in lexical tone categorization via comparing the outcome of CI +HA with that of CI alone, and (3) To examine the contributing factors for the bimodal hearing performance in the kindergarten-aged participants' lexical tone categorization. In accordance with previous work, three outcomes were predicted: (1) Mandarin-speaking pediatric bimodal listeners might show the CP of lexical tones with the help of their clinical devices; (2) Bimodal hearing could benefit unilaterally implanted children's lexical tone categorization; (3) Audiologic characteristics and demographic factors might contribute to the bimodal listening outcomes in tonal categorization.

2 Methods

2.1 Participants

This study recruited 16 native Mandarin-speaking kindergarteners (7 females and 9 males) with congenitally bilateral hearing impairment from the Shanghai Rehabilitation Center of the Deaf Children. The kindergarteners were involved in this study, mainly because the kindergarten-age range indicates an important period for the development of the CP of Mandarin tones in native child speakers (F. Chen et al., 2017). The mean age of the participants was 5.43 years old ($SD = 0.45$ years), with a range of 4.7 to 6.2 years. According to their medical reports, no participants have any history of psychiatric disorders or brain injuries. All children used spoken Mandarin Chinese as their main mode of communication, and none of them learned any knowledge of sign language. In order to confirm a normal nonverbal intelligence, the Hiskey–Nebraska Test of Learning Aptitude (HNTLA)

was adopted to screen these children's cognitive ability (Hiskey, 1966). All participants scored significantly higher than the passing criteria of 84 for a normal nonverbal intelligence (X. J. Yang, Qu, Sun, Zhu, & Wang, 2011).

The bimodal participants were implanted unilaterally and used a contralateral HA in the nonimplanted ear. They have been using their CI devices for more than two years and have an experience of bilateral bimodal hearing for more than one year. These kindergarteners were regular bimodal listeners who used a combination of the CI and HA devices for at least 75% of waking hours (c.f., Most et al., 2011; Shpak et al., 2014) according to their caregivers' observations. Their demographic information together with the CI and HA details are listed in Table 1. Standard audiometric assessment for pure tones from 125 Hz to 8 kHz was performed to determine the hearing impairment of each kindergartener participant. The aided and unaided hearing thresholds at four frequencies are shown in Figure 1 for the nonimplanted ear of each participant. This study was approved by the Ethics Committee of School of Foreign Languages at Shanghai Jiao Tong University. Informed consents were received from the caregivers of the participants.

[Insert Table 1 around here]

[Insert Figure 1 around here]

2.2 Stimuli

Speech stimuli in the tonal CP test were seven resynthesized monosyllables that constitute a tonal continuum (H. Zhang, Zhang, Ding, & Peng, 2019). As mentioned in the introduction, T1 and T2 in Mandarin were well documented to be perceived categorically by native adults and children with NH, which provided an important reference in tonal CP studies. Mandarin syllable /i/ with T1 and T2 were selected as endpoints for the synthesis of the tonal continuum. The original stimuli, /i/

with T1 and T2, were obtained from a female native adult, with each having a duration of 400 milliseconds (ms). These original samples were recorded using an audio interface (Mbox Mini) coupled with a microphone (AKG C544L), at a sampling rate of 44.1 kHz (16 bit). The F0 of T1 and T2 were measured across the entire duration of their corresponding original samples. The F0 distance between T1 and T2 was equally divided into six intervals, resulting in a seven-stimulus tonal continuum. Based on the T1 syllable token, seven tonal stimuli were generated using the Pitch-Synchronous Overlap Add (PSOLA) implemented in Praat (Boersma & Weenink, 2017). For each tonal stimulus, the F0 transitioned from the onset (ranging from 160 to 250 Hz, with a step of 15 Hz) to the offset (fixed at 250 Hz). Therefore, F0 was the only acoustic cue for manipulation, while keeping amplitude and duration constant across all seven stimuli to 65 dB SPL and 400 ms, respectively. Figure 2 depicts the schematic diagram of the pitch contours for the speech stimuli along the continuum, with stimulus No. 1 representing the prototypical T1 and No. 7 representing the prototypical T2. The feasibility of the current design of tonal continuum to examine the CP of lexical tones in Mandarin-speaking children with CIs was previously validated in our preliminary study (H. Zhang et al., 2019), and the findings indicated that native kindergarteners with CIs can show a typical CP of Mandarin T1 and T2.

[Insert Figure 2 around here]

2.3 Procedures

All kindergarteners performed the measure of CP of lexical tones inside a sound-treated therapy room. The CP measurement was examined in two device conditions including the CI alone condition and the CI + HA condition. The order of the two device conditions was randomized and counterbalanced across all participants. In addition, the test under different device conditions was performed on two lab visits with a one-week separation. The test was implemented in E-Prime 2.0

229 program (Psychology Software Tools Inc., USA) on a Windows-based laptop. All speech stimuli
230 were delivered in the free field via a loudspeaker (JBL CM220) which was placed approximately 1.2
231 m in front of the listeners. The kindergartener participants were required to pay attention to the
232 stimuli and to follow the instructions from the experimenter. All participants were able to understand
233 and follow the requirements of the testing successfully.

234 Two classical tasks for the CP of speech sounds, the identification task and the discrimination
235 task, were presented to the kindergarteners in a random order. The two-alternative forced choice (2
236 AFC) paradigm was adopted in the identification task, which required the participants to recognize
237 each tonal stimulus as T1 or T2. Initially, a procedural learning session was offered with trial-by-trial
238 feedback. During the learning session, the kindergarteners were instructed to match the correct
239 pictures for T1 and T2 stimuli: the picture depicting a car driving on a level road represents T1 and
240 the other picture of a car driving on a rising road represents T2. Two practice blocks were provided
241 after the participants acquired the matching relationship between the tones and their corresponding
242 pictures. The practice stimuli were the two endpoints of the tonal continuum (i.e., stimulus No. 1 and
243 No. 7) with each stimulus repeating four times. During the practice block, feedback was offered to
244 the participant, but not in the formal blocks. The formal blocks were performed after an accuracy of
245 90% was obtained for the practice block. Across all participants, a range of 8 to 16 practice trials was
246 needed to achieve the criterion to start the formal blocks. Two formal blocks were prepared with all
247 seven tonal stimuli repeated five times in each block, resulting in a total of 70-stimulus random
248 presentations. For the discrimination task, AX paradigm was used to instruct the kindergarteners to
249 determine whether the contrastive tonal pairs were the ‘same condition’ or ‘different condition’.
250 Seventeen contrastive pairs were constructed with an inter-stimulus interval (ISI) of 500 ms. Among
251 the 17 pairs, 10 of them were different pairs with two steps separating the two tonal stimuli in each
252 pair (e.g., 1-3, 3-1), and seven were the same pairs with a stimulus pairing with itself (e.g., 1-1, 3-3).

As in the identification task, a procedural learning session and two practice blocks were provided before the formal blocks. The learning session was intended to help the kindergarteners match the picture of two apples with the same condition, while match the picture of an apple and an orange with the different condition. Each practice block contained four contrasts along the tonal continuum (i.e., 1-1, 7-7, 2-4, and 5-3), with each contrast repeating twice. All 17 tonal pairs were presented randomly in the two formal blocks, with all pairs repeating three times in the first block and repeating twice in the second block, which resulted in 85 presentations in total. The participants were required to make a response by pointing to the correct picture on the screen in both identification and discrimination tasks. They were encouraged to make a guess when the stimulus was totally missed, although this situation was avoided as much as possible. The participants' trial-by-trial picture-pointing responses were logged by the experimenter via pressing the corresponding keys on the keyboard.

2.4 Scoring and Data Analyses

For both tasks, only the formal blocks were involved in scoring and data analyses. The identification scores and discrimination scores were calculated for each participant. The identification score was defined as the average percentage correct for T1 or T2 for each tonal stimulus. Two key parameters of the identification score were computed including the slope (i.e., reflecting how well the endpoint stimuli are differentiated) and boundary position (i.e., the 50% crossover point). Probit analysis (Finney, 1971) was adopted to estimate the identification parameters. The discrimination score was analysed in terms of the sensitivity index d' (Macmillan & Creelman, 2005). The d' takes response bias into consideration, which is calculated as the z-score of the hit rate (correct responses to 'different' pairs) minus that of the false alarm rate (incorrect responses to 'same' pairs) for each contrastive pairs along the continuum per participant. It should be noted that d' reaches infinity when discrimination performance is 0% or 100%, therefore, 0% is adjusted to 1% and 100% is adjusted to

99% to set effective floor ($d' = -4.65$) and ceiling values ($d' = 4.65$) (c.f., Macmillan & Creelman, 2005). In order to obtain a more refined analysis on discrimination performance, the discrimination scores were divided into two trial types (i.e., the between-category type and the within-category type), based on the boundary position in the identification task of each participant (F. Chen et al., 2017; Jiang, Hamm, Lim, Kirk, & Yang, 2012). The between-category score is the averaged d' for tonal pairs spanning two categories, while the within-category score is the averaged d' for pairs falling in the same category. For example, if a kindergartener participant showed a boundary position of 3.87, then for this child, the between-category score was the averaged d' for tonal pairs of 2-4 and 3-5, while the within-category score was the mean d' for the rest three pairs (i.e., 1-3, 4-6, and 5-7). Additionally, the difference between the two trial types was referred to as the peakedness score, which represents the benefit magnitude of discrimination for the between-category type relative to the within-category type (Jiang et al., 2012).

Statistical analyses were performed within R Version 3.6.0. Linear mixed-effect (LME) models were constructed to evaluate the CP of lexical tones for the bimodal kindergarteners. The package of lme4 (Bates, Mächler, Bolker, & Walker, 2015) was used to create the LME models. The models were constructed with device condition (i.e., CI alone and CI + HA) as a fixed factor to analyze the slope and boundary position for identification analysis, as well as the peakedness score for discrimination analysis. In addition, the LME models were created for the discrimination score analysis using trial type (i.e., between-category type and within-category type), device condition, and their interaction as fixed factors. Meanwhile, another LME analysis was performed on the discrimination score, with stimulus pair (i.e., 1-3, 2-4, 3-5, 4-6, and 5-7), device condition, and their interaction as fixed factors. Participant served as a random factor for all models mentioned above. The ANOVA function in lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017) was used to obtain F - and p -values of the fixed factors, which uses the Satterthwaite approximation for the

degree of freedom. Post-hoc pairwise comparisons were performed for the significant fixed factors using the lsmeans package (Lenth, 2016) with Bonferroni adjustment to obtain t -ratios, and p -values.

Moreover, linear regression models were constructed in R to examine the potential variables contributing to the bimodal participants' tonal CP. Hypothesized predictors for the tonal CP performance consisted of audiometric thresholds and demographic factors. Evaluation of the CP of lexical tones in the current report included (1) identification slope in CI alone condition, (2) identification slope in CI + HA condition. Audiometric thresholds included (1) five-frequencies unaided hearing threshold, and (2) five-frequencies aided hearing threshold. The key demographic factors included (1) implanted age and (2) duration of CI use. Separate models were created for each estimate of CP performance with the audiometric variables and demographic variables added as fixed effects. F -statistics, and p -values for the fixed effects were assessed.

3 Results

3.1 Tonal Categorical Perception

The results of tonal CP for the kindergartener participants are collectively arranged in Figures 3, 4, and 5. Figure 3 displays the overall identification curves, the boundary positions, and the identification slopes for all children in the CI alone and CI + HA device conditions. The estimated mean boundary positions for the CI alone and CI + HA device conditions were 3.87 ($SD = 0.30$) and 3.85 ($SD = 0.48$), respectively. In addition, the mean slope of the identification function was 0.96 ($SD = 0.58$) for the CI alone condition, and 1.57 ($SD = 0.65$) for the bimodal condition. LME models revealed significant main effects of device condition for slope analysis ($F(1, 15) = 25.34, p < .001$), but not for boundary position analysis ($F(1, 15) = 0.03, p = .86$). Post-hoc pairwise comparisons showed that the slope for the CI + HA condition was significantly steeper than that for the CI alone condition ($t(15) = 5.03, p < .001$). Moreover, individual identification functions in conjunction with

the corresponding slopes for each participant are shown in Figure 4. Inspection of the individual results revealed a wide range of performances of lexical tone categorization for the bimodal kindergarteners, which coincides with the well-known heterogeneity in children with CIs.

[Insert Figure 3 around here]

[Insert Figure 4 around here]

The group mean discrimination curves of the tonal contrasts for the participants are illustrated in Figure 5(A). For the CI alone condition, the overall mean discrimination peak in d' was 1.11 at the tonal contrast of 3-5, which straddles the mean boundary position at 3.87. Meanwhile, for the CI + HA condition, the average discrimination score across all kindergarteners reached the peak (i.e., 1.51 in d') at 2-4, which also aligns well with the corresponding mean boundary position at 3.85. Figure 5(B) shows the discrimination scores of different stimulus pairs for the CI + HA condition and the CI alone condition. Statistical results revealed significantly higher d' values for the CI + HA condition over the CI alone condition in stimulus pair 2-4 ($t(144) = 2.39, p = .02$), but not in other stimulus pairs ($ps > .4$). The d' values of the between-category type and the within-category type are displayed in Figure 5(C) for the two device conditions. Statistical analysis revealed a significant main effect of the trial type ($F(1, 48) = 26.04, p < .001$). Post-hoc pairwise comparisons with Bonferroni adjustment showed that the between-category score was significantly higher than the within-category score in both the CI + HA condition ($t(48) = 4.54, p < .001$) and the CI alone condition ($t(48) = 2.78, p = .01$). In addition, Figure 5(D) shows the peakedness score of each device condition, which was calculated as the difference between the two trial types within the same device condition. The LME analysis for the discrimination score showed a significant main effect of device condition ($F(1, 15) = 6.56, p = .02$), with significantly higher peakedness scores for the CI + HA condition than the CI alone condition ($t(15) = 2.56, p = .02$).

[Insert Figure 5 around here]

3.2 Linear Regression Result

The linear regression models revealed that the hearing threshold at 250 Hz was significantly associated with the identification slope of tonal CP for the bimodal participants. Specifically, the unaided hearing threshold at 250 Hz was a significant predictor for tonal CP in the CI + HA device condition, with better residual hearing contributing to a steeper slope ($F(1, 15) = 5.73, p = .03$) (see Figure 6). However, neither implanted age nor CI duration revealed a significant relationship with tonal CP in the CI + HA device condition ($ps > .61$) or in the CI alone condition ($ps > .14$). Visual inspection of residual plots revealed that the residuals were normally distributed without any obvious deviations from homoscedasticity for each model. The identification slope was well established in indicating the steepness of the categorical boundary, with a steeper slope representing a more categorical-like perceptual performance. Therefore, the unilaterally implanted kindergartener with better residual acoustic hearing at 250 Hz in the nonimplanted ear tended to show better lexical tone categorization when a contralateral HA was added.

[Insert Figure 6 around here]

4 Discussion

The present study examined bimodal benefits for the CP of lexical tones in native Mandarin-speaking kindergarteners who use a CI in one ear and an HA in the other ear, a topic that has not been systematically investigated with a carefully controlled speech continuum of lexical tones in previous studies. The tonal CP of the participants was measured in the CI + HA condition and the CI alone condition for a direct comparison between the two device conditions. The results demonstrated significant benefits of adding a contralateral HA in lexical tone categorization for the unilaterally

implanted kindergarteners. Moreover, the residual hearing at 250 Hz level could be a contributing factor for the bimodal hearing performance in the CP of lexical tones. The results corroborate the proposition that bimodal benefit arises primarily from the availability of acoustic F0 information, which improves the categorical perception of lexical tones in Mandarin-speaking children with CIs. Moreover, our findings lend support for the clinical practice of fitting a contralateral HA in kindergarteners with unilateral CIs to take advantage of the low-frequency acoustic hearing.

4.1 Categorical Perception of Lexical Tones in Bimodal Kindergarteners

While Mandarin tone recognition has been widely investigated using naturally produced tonal stimuli among the native pediatric CI recipients (Tan, Dowell, & Vogel, 2016 for a review), it is relatively under-studied on the CP of lexical tones with well-controlled synthesized speech stimuli in this clinical population. The CP of speech sounds enhances perceptual sensitivity to between-category contrasts and ignores subtle within-category variations, which is believed essential in supporting and facilitating efficient speech communication (Bidelman et al., 2013; G. Shen & Froud, 2019; X. Wang et al., 2017). Prior developmental studies suggest a strong correlation between early phonetic categorization of native speech sounds during infancy and speech perception and production skills later in childhood (Kuhl, Patricia & Rivera-Gaxiola, 2008; Kuhl et al., 2005, 2006). Therefore, the phonetic categorization of native speech sounds is a remarkable linguistic phenomenon that plays a pivotal role in efficient speech communication and proper language development. However, prior to the present study, it remained an open question as to whether Mandarin-speaking children with CIs can perceive lexical tones categorically.

The data suggest that the bimodal kindergarteners in this study can perceive T1 and T2 categorically. Two major characteristics of CP, as generalized by Liberman et al. (1957), are a sharp identification boundary between two phonetic categories and a prominent discrimination peak around

the identification boundary. In the current tonal CP test, typical S-shaped functions were shown for the identification curves of both the CI alone device condition and the CI + HA condition, indicating relatively sharp categorical boundaries for both device conditions. Meanwhile, prominent discrimination peaks were well aligned with the corresponding categorical boundaries (see Figure 3(A) and Figure 5(A)). In accordance with the hallmark of a typical CP, statistical analyses revealed that the improved discrimination accuracy for the between-category type over the within-category type was significant in both device conditions. The results in this study echo a prior study with a cohort of adolescent and adult CI users (Luo et al., 2014), in which typical S-shaped functions were exhibited in the identification of a tonal continuum varying from T1 to T2. This contributes to the corollary that native kindergarteners with CIs are capable of perceiving Mandarin tones categorically.

Given the well-known limitations of the contemporary CI technology in encoding pitch information, it would be rather implausible that kindergarteners with CIs could show typical CP of lexical tones, the measure of which depends almost solely on pitch cues. Several recent reports validated that the degraded F0 signals provided by the CI device could be exploited by Mandarin-speaking pediatric recipients in lexical tone identification (Deroche, Lu, Lin, Chatterjee, & Peng, 2019; S.-C. Peng et al., 2017) and sentence recognition (L. Zhang et al., 2018). For example, Peng et al. (2017) revealed that Mandarin tone recognition in children with CIs also depends primarily on F0 contour cues, although the clinical individuals may change their cue-weighting patterns with more reliance on duration cues than the NH age mates. Therefore, it is reasonable to postulate that the pediatric CI users could adapt to the impoverished pitch cues for the development of lexical tone categorization via the accumulating experience of tonal language exposure with their CI devices.

4.2 Bimodal Benefits for Lexical Tone Categorization

The available degraded F0 signals from electric hearing combined with the complementary low-frequency components from the acoustic hearing benefit bimodal kindergarteners' lexical tone categorization. Indeed, significant bimodal benefits were demonstrated for the kindergartener participants in the CP of the tonal continuum involving T1 and T2. For the identification analyses, the results revealed a significantly steeper slope for the CI + HA condition relative to the CI alone condition. Additionally, the discrimination analyses found a significantly improved performance for the between-category stimuli pair (i.e., 2-4) and a significantly higher peakedness score in the bimodal hearing condition over the CI alone condition. The identification and discrimination results indicated, respectively, a much sharper identification boundary and a more prominent discrimination peak for the CI + HA device condition than the CI alone condition. All results converged to the conclusion that the kindergarteners could perceive T1 and T2 more categorically with bimodal configuration. It should be noted that the stimuli of the current test for the tonal CP were cued almost exclusively on the F0 feature while keeping amplitude and duration constant. Therefore, the significant benefits observed in this study were compatible with previous studies reported that task-specific tests relying more on pitch cues could receive more tangible bimodal benefits (Luo et al., 2014; Mok et al., 2017; H. Zhang et al., 2020).

However, the current findings are partially inconsistent with a prior report (Luo et al., 2014) that adopted the identification task from the CP paradigm to investigate the bimodal benefits for lexical tone normalization in CI users. In the test without a context, Luo et al. (2014) revealed no significant difference in the identification function slopes between the CI + HA and the CI alone device condition. The inconsistent findings between Luo et al. (2014) and the present study can probably be attributed to several factors including, but not limited to, differences in stimuli features, experimental designs, and subject characteristics. A relatively smaller step size was employed in the present study (i.e., 15 Hz in our study vs. 20 Hz in Luo et al. (2014)). The smaller tested step would

provide more perceptually ambiguous stimuli around the boundary position of T1 and T2, a situation that could be more susceptible to the availability of additional F0 information. As a result, the refined psychoacoustic scale of the current tonal continuum may subject to more fine-grained scrutiny for the differences between the two device conditions in the identification functions of the continuum. The present study adopted a within-subject design, in which all participants were bimodal kindergarteners and were tested in both bimodal condition and CI alone condition. In contrast, Luo et al. (2014) used a between-subject comparison analysis, because not all participants in their study completed the tests in both device conditions. In comparison with the between-subject study design, the within-subject design can minimize confounding variables that may potentially compromise statistical significance when comparing perceptual outcomes of the two device conditions. Moreover, participants in Luo et al.'s study reported relatively poorer unaided thresholds of the nonimplanted ear (e.g., 81 vs. 73 dB HL at 250 Hz) and longer duration of hearing loss of the implanted ear (i.e., age at implantation, 9.5 vs. 2.0 years old). Yang and Zeng (2017) suggested that longer duration and more severe hearing loss could predict poorer bimodal integration efficiency, which in turn led to less bimodal benefits. Accordingly, the kindergarteners in this study could develop better bimodal integration to reap more bimodal benefits in lexical tone perception. On the other hand, kindergartener participants were recruited in this study, who were still under a sensitive period for the development of the CP of Mandarin tones. By contrast, Luo et al.'s study enrolled a cohort of older participants with an age range of 10 to 20 years old. Chen et al. (2017) showed that typically developing children demonstrate CP of Mandarin tones as early as four years old, which becomes adult-like by the age of six years. From the rehabilitative perspective, the older CI subjects in Luo et al.'s study may have reached a plateau during lexical tone rehabilitation and obtained saturated attainment on tonal CP. The ceiling effect encountered by the good performers can limit the degree of bimodal benefits when the acoustic stimulation is added through the contralateral HA (T. Zhang, Spahr, Dorman, & Saoji, 2013). Regardless, Luo et al. (2014) demonstrated significant bimodal benefits in the test with a context,

with a similar lexical tone normalization as NH controls in the CI + HA condition, but not in the CI alone condition. Altogether, the convergent evidence contributes to the corollary that the contralateral HA could be used by the unilateral implantees to improve their lexical tone categorization.

Furthermore, our regression analyses showed that the unaided audiometric threshold at 250 Hz in the nonimplanted ear could be a contributing factor for the lexical tone categorization in the bimodal kindergarteners. The speech stimuli in this study were synthesized from a native female adult, with the F0 range of 160 to 250 Hz. The significant influences of 250 Hz residual hearing support the proposition that the acoustic F0 information could be underlying mechanisms accounting for the bimodal benefits in speech perception when the acoustic stimulation is added to the electric stimulation (T. Zhang et al., 2010). While Zhang et al. (2020) have demonstrated a significant positive relationship between residual hearing at low frequencies and the bimodal benefits in natural-produced tone recognition, the adoption of the classical CP paradigm in this study with well-controlled synthetic speech stimuli maximally avoided the confounding impact of amplitude and duration cues in measuring bimodal benefits. Our results consolidated that unilaterally implanted kindergarteners with better residual acoustic hearing at low frequencies could have better prognostic in bimodal benefits for lexical tone perception when wearing an additional HA in the nonimplanted ear.

4.3 Limitations

One major limitation of our study is that only T1 and T2 were adopted to test the CP of lexical tones. The four Mandarin tones are not uniformly easy or difficult to recognize with T2 and T3 representing the most difficult contrast to discern due to the acoustic similarity of the two tones (X. S. Shen & Lin, 1991; X. S. Shen, Lin, & Yan, 1993). Prior studies demonstrated the identification confusion between T2 and T3 in Mandarin-speaking pediatric CI users (e.g., Y. Chen et al., 2014; S.-

C. Peng et al., 2017; H. Zhang et al., 2020). It is of great interest to employ more challenging tonal contrasts, such as T2 and T3, to evaluate the bimodal benefits for lexical tone categorization. Meanwhile, a smaller step size of the stimuli continuum was suggested to potentially contribute to more robust bimodal benefits in the present study over the Luo et al.' study (2014). Further study with a within-subject experimental design is needed to verify whether differences in stimulus interval steps can exert influences on the magnitudes of the bimodal benefit. In addition, only behavioral protocols were adopted in the present study. A body of evidence has demonstrated the underlying neural correlates of the CP of lexical tones for both native speakers (e.g., Xi, Zhang, Shu, Zhang, & Li, 2010; Zheng, Minett, Peng, & Wang, 2012) and nonnative speakers (e.g., G. Shen & Froud, 2019; Yu et al., 2019), as well as the cortical abnormalities of the tonal CP for children with dyslexia (e.g., Y. Zhang et al., 2012) and autism (e.g., X. Wang et al., 2017). Further investigations with electrophysiological approaches are warranted to flesh out the underlying mechanisms of the bimodal benefits in lexical tone categorization from the neurophysiological perspective.

Another limitation is the relatively small subject sample with a narrow age range and the lack of a control group for comparison. Age at implantation and duration of CI use have been demonstrated as two major demographic factors predicting lexical tone perception for pediatric implantees in a prior report with a large-scale subject pool (Zhou, Huang, Chen, & Xu, 2013). However, our study failed to replicate this predictive relationship. To obtain robust statistical results, a much larger sample size with a broader range of demographic characteristics is desirable. The principal goal of this study was to examine the bimodal benefit of Mandarin tone categorization in native kindergarteners with unilateral CIs. A body of recent studies that aimed to evaluate the bimodal benefits in the perception of various speech features also recruited only bimodal listeners without NH controls (e.g., Cheng et al., 2018; Kessler, Ananthakrishnan, Smith, D'Onofrio, & Gifford, 2020; Liu et al., 2019; Mok et al., 2017; Park, Teagle, Gagnon, Woodard, & Brown, 2019;

Tao, Liu, Yang, Wilson, & Zhou, 2018; H.-I. Yang & Zeng, 2017; H. Zhang et al., 2020). Whether the bimodal kindergarteners can be afforded the opportunity to use the bimodal learning effects to achieve tonal CP performance on par with their NH peers requires further investigation. To this end, future longitudinal research with the target CI users and their NH age-matched controls is necessary to chart and compare the developmental trajectory of lexical tone categorization for the normal and clinical populations.

4.4 Implications

Notwithstanding its limitations, this study presents significant implications for rehabilitative interventions. Although fitting a contralateral HA in the nonimplanted ear for unilaterally implanted children potentially benefits their speech perception at the group level, the prevalence of bimodal configuration is much lower than might be expected (Neuman et al., 2017). In addition, there is a paucity of research on the benefits of an additional HA in Mandarin-speaking bimodal listeners, and even the handful of related studies has shown a mixed result (e.g., Y. Li et al., 2014; Luo et al., 2014; Mok et al., 2017; Tao et al., 2018). This situation confounds the efforts to make guidelines for bimodal candidacy and promote publicity of bimodal configurations due to the shortage of compelling empirical evidence. Furthermore, surveys of audiologists from Mainland China reveal that the lack of nationwide clinical guidelines for bimodal candidacy and management leaves the decision about bimodal use to clinician preference, patient convenience, or even the proverbial coin flip. Overall, our findings of enhanced lexical tone categorization with bimodal listening lend support for the clinical practice of fitting a contralateral HA in the nonimplanted ear for the potential bimodal benefits in lexical tone categorization for kindergarteners with CIs.

The past decades have witnessed remarkable advances in the auditory rehabilitation and cochlear implantation programs in China (J.-N. Li et al., 2017 for a review). More CI recipients have

some residual acoustic hearing in the nonimplanted ear, the situation of which should be taken into consideration while setting up rehabilitation regimens for this clinical population. For instance, computerized speech training paradigms (e.g., Ingvalson, Lee, Fiebig, & Wong, 2013; Miller, Zhang, & Nelson, 2016; T. Zhang, Dorman, Fu, & Spahr, 2012) could be developed to afford the opportunity to optimally accrue the bimodal effects. Additionally, the CP of Mandarin tones could refine the diagnostic measures for the potential bimodal benefit in speech perception for the unilaterally implanted native kindergarteners. The effective reception of pitch information is a prerequisite for correct speech understanding in Mandarin Chinese. Given the fact that an additional HA may enrich the low-frequency components that are beneficial for representing the pitch cues, the bimodal configuration could be a preferable alternative to a second CI basing on the residual acoustic hearing of the recipient's nonimplanted ear. The clinical audiological assessments could incorporate the tonal CP as an evaluation of the bimodal benefits, which can provide empirical support for evidence-based clinical decision of bilateral CI prescription for the pediatric Mandarin-speaking candidates.

5 Conclusions

Despite large individual variability, findings of this study confirmed the categorical perception of the tonal continuum of T1 and T2 in Mandarin kindergarteners using a CI and an HA in opposite ears. The additional acoustic stimulations from the HA in the nonimplanted ear contribute to a more categorical-like perception for Mandarin tones. Moreover, unaided audiometric thresholds at 250 Hz was shown a contributing factor for the varying outcomes of bimodal hearing in lexical tone categorization. The results provide corroborating evidence to the clinical practice of adding an HA in the nonimplanted ear to take advantage of the potential bimodal benefit in kindergarteners with CIs.

6 Acknowledgments

555 We would like to thank Yongqin Li, Rentao Wei, and Renxia Tao from the Shanghai
556 Rehabilitation Center of the Deaf Children for their assistance in implementing this study.

557 **References**

- 558 Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models
559 using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- 560 Bidelman, G. M., Moreno, S., & Alain, C. (2013). Tracing the emergence of categorical speech
561 perception in the human auditory system. *Neuroimage*, 79, 201–212.
562 <https://doi.org/10.1016/j.neuroimage.2013.04.093>
- 563 Boersma, P., & Weenink, D. (2017). Praat: Doing phonetics by computer (Computer program,
564 Version 6.0.33). Retrieved from <http://www.praat.org>
- 565 Carroll, J., & Zeng, F.-G. (2007). Fundamental frequency discrimination and speech perception in
566 noise in cochlear implant simulations. *Hearing Research*, 231(1–2), 42–53.
567 <https://doi.org/10.1016/j.heares.2007.05.004>
- 568 Chao, Y. R. (1948). *Mandarin primer: An intensive course in spoken Chinese*. Berkeley, CA:
569 Harvard University Press.
- 570 Chatterjee, M., & Peng, S.-C. (2008). Processing F0 with cochlear implants: Modulation frequency
571 discrimination and speech intonation recognition. *Hearing Research*, 235(1–2), 143–156.
572 <https://doi.org/10.1016/j.heares.2007.11.004>
- 573 Chatterjee, M., Zion, D. J., Deroche, M. L., Burianek, B. A., Limb, C. J., Goren, A. P., ...
574 Christensen, J. A. (2015). Voice emotion recognition by cochlear-implanted children and their

575 normally-hearing peers. *Hearing Research*, 322, 151–162.
576 <https://doi.org/10.1016/j.heares.2014.10.003>

577 Chen, F., Peng, G., Yan, N., & Wang, L. (2017). The development of categorical perception of
578 Mandarin tones in four-to seven-year-old children. *Journal of Child Language*, 44(6), 1413–
579 1434. <https://doi.org/10.1017/S0305000916000581>

580 Chen, Y., Wong, L. L. N., Chen, F., & Xi, X. (2014). Tone and sentence perception in young
581 Mandarin-speaking children with cochlear implants. *International Journal of Pediatric*
582 *Otorhinolaryngology*, 78(11), 1923–1930. <https://doi.org/10.1016/j.ijporl.2014.08.025>

583 Cheng, X., Liu, Y., Wang, B., Yuan, Y., Galvin, J. J., Fu, Q.-J., ... Chen, B. (2018). The Benefits of
584 Residual Hair Cell Function for Speech and Music Perception in Pediatric Bimodal Cochlear
585 Implant Listeners. *Neural Plasticity*, 2018. <https://doi.org/10.1155/2018/4610592>

586 Ching, T. Y. C., Psarros, C., Hill, M., Dillon, H., & Incerti, P. (2001). Should children who use
587 cochlear implants wear hearing aids in the opposite ear? *Ear and Hearing*, 22(5), 365–380.

588 Cullington, H. E., & Zeng, F.-G. (2011). Comparison of bimodal and bilateral cochlear implant users
589 on speech recognition with competing talker, music perception, affective prosody discrimination
590 and talker identification. *Ear and Hearing*, 32(1), 16–30.
591 <https://doi.org/10.1097/AUD.0b013e3181edfbd2>

592 Deroche, M. L. D., Lu, H.-P., Lin, Y.-S., Chatterjee, M., & Peng, S.-C. (2019). Processing of
593 acoustic information in lexical tone production and perception by pediatric cochlear implant
594 recipients. *Frontiers in Neuroscience*, 13, 639. <https://doi.org/10.3389/fnins.2019.00639>

595 Dorman, M. F., Spahr, T., Gifford, R., Loiselle, L., McKarns, S., Holden, T., ... Finley, C. (2007).
596 An electric frequency-to-place map for a cochlear implant patient with hearing in the

597 nonimplanted ear. *Journal for the Association for Research in Otolaryngology*, 8(2), 234–240.
 598 <https://doi.org/10.1007/s10162-007-0071-1>

599 Finney, D. J. (1971). *Probit analysis* (3rd ed.). Cambridge, UK: Cambridge University Press.

600 Fu, Q.-J., Chinchilla, S., Nogaki, G., & Galvin III, J. J. (2005). Voice gender identification by
 601 cochlear implant users: The role of spectral and temporal resolution. *The Journal of the*
 602 *Acoustical Society of America*, 118(3), 1711–1718. <https://doi.org/10.1121/1.1985024>

603 Gifford, R. H., & Dorman, M. F. (2019). Bimodal hearing or bilateral cochlear implants? Ask the
 604 patient. *Ear and Hearing*, 40(3), 501–516. <https://doi.org/10.1097/AUD.00000000000000657>

605 Gu, X., Liu, B., Liu, Z., Qi, B., Wang, S., Dong, R., ... Zhou, Q. (2017). A follow-up study on music
 606 and lexical tone perception in adult Mandarin-speaking cochlear implant users. *Otology &*
 607 *Neurotology*, 38(10), e421–e428. <https://doi.org/10.1097/MAO.0000000000001580>

608 Hiskey, M. S. (1966). *Hiskey-Nebraska test of learning aptitude*. Cambridge, UK: Union College
 609 Press.

610 Holt, C. M., Lee, K. Y. S., Dowell, R. C., & Vogel, A. P. (2018). Perception of Cantonese lexical
 611 tones by pediatric cochlear implant users. *Journal of Speech, Language, and Hearing Research*,
 612 61(1), 174–185. https://doi.org/10.1044/2017_JSLHR-H-17-0027

613 Ingvalson, E. M., Lee, B., Fiebig, P., & Wong, P. C. M. (2013). The effects of short-term
 614 computerized speech-in-noise training on postlingually deafened adult cochlear implant
 615 recipients. *Journal of Speech, Language, and Hearing Research*, 56(1), 81–88.
 616 [https://doi.org/10.1044/1092-4388\(2012/11-0291\)](https://doi.org/10.1044/1092-4388(2012/11-0291))

617 Jiam, N. T., Caldwell, M., Deroche, M. L., Chatterjee, M., & Limb, C. J. (2017). Voice emotion
618 perception and production in cochlear implant users. *Hearing Research*, 352, 30–39.
619 <https://doi.org/10.1016/j.heares.2017.01.006>

620 Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. (2012). Impaired categorical perception of
621 lexical tones in Mandarin-speaking congenital amusics. *Memory & Cognition*, 40(7), 1109–
622 1121. <https://doi.org/10.3758/s13421-012-0208-2>

623 Kessler, D. M., Ananthakrishnan, S., Smith, S. B., D’Onofrio, K., & Gifford, R. H. (2020).
624 Frequency following response and speech recognition benefit for combining a cochlear implant
625 and contralateral hearing aid. *Trends in Hearing*, 24, 2331216520902001.
626 <https://doi.org/10.1177/2331216520902001>

627 Kuhl, Patricia, K., & Rivera-Gaxiola, M. (2008). Neural substrates of language acquisition. *Annual*
628 *Review of Neuroscience*, 31, 511–534. <https://doi.org/10.1146/annurev.neuro.30.051606.094321>

629 Kuhl, P. K. (1991). Human adults and human infants show a “perceptual magnet effect” for the
630 prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, 50(2), 93–107.
631 <https://doi.org/10.3758/BF03212211>

632 Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and
633 later language development: Implications for the “critical period.” *Language Learning and*
634 *Development*, 1(3–4), 237–264. <https://doi.org/10.1080/15475441.2005.9671948>

635 Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a
636 facilitation effect for native language phonetic perception between 6 and 12 months.
637 *Developmental Science*, 9(2), F13–F21. <https://doi.org/10.1111/j.1467-7687.2006.00468.x>

638 Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear
639 mixed effects models. *Journal of Statistical Software*, 82(13), 1–26.
640 <https://doi.org/10.18637/jss.v082.i13>

641 Lenth, R. V. (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software*,
642 69(1), 1–33. <https://doi.org/10.18637/jss.v069.i01>

643 Li, J.-N., Chen, S., Zhai, L., Han, D.-Y., Eshraghi, A. A., Feng, Y., ... Liu, X.-Z. (2017). The
644 advances in hearing rehabilitation and cochlear implants in China. *Ear and Hearing*, 38(6), 647–
645 652. <https://doi.org/10.1097/AUD.0000000000000441>

646 Li, Y., Zhang, G., Galvin III, J. J., & Fu, Q.-J. (2014). Mandarin speech perception in combined
647 electric and acoustic stimulation. *PloS One*, 9(11), e112471.
648 <https://doi.org/10.1371/journal.pone.0112471>

649 Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of
650 speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*,
651 54(5), 358–368. <https://doi.org/10.1037/h0044417>

652 Liu, Y. W., Tao, D. D., Chen, B., Cheng, X., Shu, Y., Galvin, J. J., & Fu, Q. J. (2019). Factors
653 affecting bimodal benefit in pediatric Mandarin-speaking Chinese cochlear implant users. *Ear*
654 *and Hearing*, 40(6), 1316–1327. <https://doi.org/10.1097/AUD.0000000000000712>

655 Luo, X., Chang, Y., Lin, C.-Y., & Chang, R. Y. (2014). Contribution of bimodal hearing to lexical
656 tone normalization in Mandarin-speaking cochlear implant users. *Hearing Research*, 312, 1–8.
657 <https://doi.org/10.1016/j.heares.2014.02.005>

658 Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide*. Psychology press.

- 659 McDermott, H. J. (2004). Music perception with cochlear implants: A review. *Trends in*
660 *Amplification*, 8(2), 49–82. <https://doi.org/10.1177/108471380400800203>
- 661 Meister, H., Fürsen, K., Streicher, B., Lang-Roth, R., & Walger, M. (2016). The use of voice cues for
662 speaker gender recognition in cochlear implant recipients. *Journal of Speech, Language, and*
663 *Hearing Research*, 59(3), 546–556. https://doi.org/10.1044/2015_JSLHR-H-15-0128
- 664 Meister, H., Landwehr, M., Pyschny, V., Walger, M., & Wedel, H. von. (2009). The perception of
665 prosody and speaker gender in normal-hearing listeners and cochlear implant recipients.
666 *International Journal of Audiology*, 48(1), 38–48. <https://doi.org/10.1080/14992020802293539>
- 667 Miller, S., Zhang, Y., & Nelson, P. B. (2016). Efficacy of multiple-talker phonetic identification
668 training in postlingually deafened cochlear implant listeners. *Journal of Speech, Language, and*
669 *Hearing Research*, 59(1), 90–98. https://doi.org/10.1044/2015_JSLHR-H-15-0154
- 670 Mok, M., Grayden, D., Dowell, R. C., & Lawrence, D. (2006). Speech perception for adults who use
671 hearing aids in conjunction with cochlear implants in opposite ears. *Journal of Speech,*
672 *Language, and Hearing Research*, 49(2), 338–351. [https://doi.org/10.1044/1092-](https://doi.org/10.1044/1092-4388(2006/027))
673 [4388\(2006/027\)](https://doi.org/10.1044/1092-4388(2006/027))
- 674 Mok, M., Holt, C. M., Lee, K., Dowell, R. C., & Vogel, A. P. (2017). Cantonese tone perception for
675 children who use a hearing aid and a cochlear implant in opposite ears. *Ear and Hearing*, 38(6),
676 e359–e368. <https://doi.org/10.1097/AUD.0000000000000453>
- 677 Morera, C., Manrique, M., Ramos, A., Garcia-Ibanez, L., Cavalle, L., Huarte, A., ... Estrada, E.
678 (2005). Advantages of binaural hearing provided through bimodal stimulation via a cochlear
679 implant and a conventional hearing aid: A 6-month comparative study. *Acta Oto-Laryngologica*,
680 125(6), 596–606. <https://doi.org/10.1080/00016480510027493>

681 Morton, K. D., Torrione Jr, P. A., Throckmorton, C. S., & Collins, L. M. (2008). Mandarin Chinese
682 tone identification in cochlear implants: Predictions from acoustic models. *Hearing Research*,
683 244(1–2), 66–76. <https://doi.org/10.1016/j.heares.2008.07.008>

684 Most, T., Gaon-Sivan, G., Shpak, T., & Luntz, M. (2012). Contribution of a contralateral hearing aid
685 to perception of consonant voicing, intonation, and emotional state in adult cochlear implantees.
686 *Journal of Deaf Studies and Deaf Education*, 17(2), 244–258.
687 <https://doi.org/10.1093/deafed/enr046>

688 Most, T., Harel, T., Shpak, T., & Luntz, M. (2011). Perception of suprasegmental speech features via
689 bimodal stimulation: Cochlear implant on one ear and hearing aid on the other. *Journal of*
690 *Speech, Language, and Hearing Research*, 54(2), 668–678. [https://doi.org/10.1044/1092-](https://doi.org/10.1044/1092-4388(2010/10-0071))
691 4388(2010/10-0071)

692 Nakata, T., Trehub, S. E., & Kanda, Y. (2012). Effect of cochlear implants on children’s perception
693 and production of speech prosody. *The Journal of the Acoustical Society of America*, 131(2),
694 1307–1314. <https://doi.org/10.1121/1.3672697>

695 Neuman, A. C., Waltzman, S. B., Shapiro, W. H., Neukam, J. D., Zeman, A. M., & Svirsky, M. A.
696 (2017). Self-reported usage, functional benefit, and audiologic characteristics of cochlear
697 implant patients who use a contralateral hearing aid. *Trends in Hearing*, 21, 1–14.
698 <https://doi.org/10.1177/2331216517699530>

699 Paquette, S., Ahmed, G. D., Goffi-Gomez, M. V, Hoshino, A. C. H., Peretz, I., & Lehmann, A.
700 (2018). Musical and vocal emotion perception for cochlear implants users. *Hearing Research*,
701 370, 272–282. <https://doi.org/10.1016/j.heares.2018.08.009>

702 Park, L. R., Teagle, H. F. B., Gagnon, E., Woodard, J., & Brown, K. D. (2019). Electric-acoustic
 703 stimulation outcomes in children. *Ear and Hearing*, 40(4), 849–857.
 704 <https://doi.org/10.1097/AUD.0000000000000658>

705 Peng, G., Zheng, H.-Y., Gong, T., Yang, R.-X., Kong, J.-P., & Wang, W. S.-Y. (2010). The influence
 706 of language experience on categorical perception of pitch contours. *Journal of Phonetics*, 38(4),
 707 616–624. <https://doi.org/10.1016/j.wocn.2010.09.003>

708 Peng, S.-C., Lu, H.-P., Lu, N., Lin, Y.-S., Deroche, M. L. D., & Chatterjee, M. (2017). Processing of
 709 acoustic cues in lexical-tone identification by pediatric cochlear-implant recipients. *Journal of*
 710 *Speech, Language, and Hearing Research*, 60(5), 1223–1235.
 711 https://doi.org/10.1044/2016_JSLHR-S-16-0048

712 Peng, S.-C., Tomblin, J. B., Cheung, H., Lin, Y.-S., & Wang, L.-S. (2004). Perception and
 713 production of Mandarin tones in prelingually deaf children with cochlear implants. *Ear and*
 714 *Hearing*, 25(3), 251–264. <https://doi.org/10.1097/01.AUD.0000130797.73809.40>

715 Peng, S.-C., Tomblin, J. B., & Turner, C. W. (2008). Production and perception of speech intonation
 716 in pediatric cochlear implant recipients and individuals with normal hearing. *Ear and Hearing*,
 717 29(3), 336–351. <https://doi.org/10.1097/AUD.0b013e318168d94d>

718 Peterson, N. R., Pisoni, D. B., & Miyamoto, R. T. (2010). Cochlear implants and spoken language
 719 processing abilities: Review and assessment of the literature. *Restorative Neurology and*
 720 *Neuroscience*, 28(2), 237–250. <https://doi.org/10.3233/RNN-2010-0535>

721 Shannon, R. V., Fu, Q.-J., & Galvin III, J. J. (2004). The number of spectral channels required for
 722 speech recognition depends on the difficulty of the listening situation. *Acta Oto-Laryngologica*,
 723 124(0), 50–54. <https://doi.org/10.1080/03655230410017562>

724 Shen, G., & Froud, K. (2019). Electrophysiological correlates of categorical perception of lexical
725 tones by English learners of Mandarin Chinese: An ERP study. *Bilingualism: Language and*
726 *Cognition*, 22(2), 253–265. <https://doi.org/10.1017/S136672891800038X>

727 Shen, X. S., & Lin, M. (1991). A perceptual study of Mandarin tones 2 and 3. *Language and Speech*,
728 34(2), 145–156. <https://doi.org/10.1177/002383099103400202>

729 Shen, X. S., Lin, M., & Yan, J. (1993). F0 turning point as an F0 cue to tonal contrast: A case study
730 of Mandarin tones 2 and 3. *The Journal of the Acoustical Society of America*, 93(4), 2241–2243.
731 <https://doi.org/10.1121/1.406688>

732 Shpak, T., Most, T., & Luntz, M. (2014). Fundamental frequency information for speech recognition
733 via bimodal stimulation: Cochlear implant in one ear and hearing aid in the other. *Ear and*
734 *Hearing*, 35(1), 97–109. <https://doi.org/10.1097/AUD.0b013e3182a2c814>

735 Tan, J., Dowell, R., & Vogel, A. (2016). Mandarin lexical tone acquisition in cochlear implant users
736 with prelingual deafness: A review. *American Journal of Audiology*, 25(3), 246–256.
737 https://doi.org/10.1044/2016_AJA-15-0069

738 Tao, D., Deng, R., Jiang, Y., Galvin III, J. J., Fu, Q.-J., & Chen, B. (2015). Melodic pitch perception
739 and lexical tone perception in Mandarin-speaking cochlear implant users. *Ear and Hearing*,
740 36(1), 102–110. <https://doi.org/10.1097/AUD.0000000000000086>

741 Tao, D., Liu, J.-S., Yang, Z.-D., Wilson, B. S., & Zhou, N. (2018). Bilaterally combined electric and
742 acoustic hearing in Mandarin-speaking listeners: The population with poor residual hearing.
743 *Trends in Hearing*, 22, 1–13. <https://doi.org/10.1177/2331216518757892>

- 744 Van De Velde, D. J., Schiller, N. O., Levelt, C. C., Van Heuven, V. J., Beers, M., Briare, J. J., &
745 Frijns, J. H. M. (2019). Prosody perception and production by children with cochlear implants.
746 *Journal of Child Language*, 46(1), 111–141. <https://doi.org/10.1017/S0305000918000387>
- 747 van Wieringen, A., & Wouters, J. (2015). What can we expect of normally-developing children
748 implanted at a young age with respect to their auditory, linguistic and cognitive skills? *Hearing*
749 *Research*, 322, 171–179. <https://doi.org/10.1016/j.heares.2014.09.002>
- 750 Wang, S., Liu, B., Dong, R., Zhou, Y., Li, J., Qi, B., ... Zhang, L. (2012). Music and lexical tone
751 perception in Chinese adult cochlear implant users. *The Laryngoscope*, 122(6), 1353–1360.
752 <https://doi.org/10.1002/lary.23271>
- 753 Wang, W. S. Y. (1973). The Chinese language. *Scientific American*, 228(2), 50–63.
754 <https://doi.org/10.1038/scientificamerican0273-50>
- 755 Wang, X., Wang, S., Fan, Y., Huang, D., & Zhang, Y. (2017). Speech-specific categorical perception
756 deficit in autism: An event-related potential study of lexical tone processing in Mandarin-
757 speaking children. *Scientific Reports*, 7, 43254. <https://doi.org/10.1038/srep43254>
- 758 Whalen, D. H., & Xu, Y. (1992). Information for Mandarin tones in the amplitude contour and in
759 brief segments. *Phonetica*, 49(1), 25–47. <https://doi.org/10.1159/000261901>
- 760 Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in
761 Chinese revealed by mismatch negativity. *Neuroscience*, 170(1), 223–231.
762 <https://doi.org/10.1016/j.neuroscience.2010.06.077>
- 763 Xu, Y., Gandour, J. T., & Francis, A. L. (2006). Effects of language experience and stimulus
764 complexity on the categorical perception of pitch direction. *The Journal of the Acoustical*
765 *Society of America*, 120(2), 1063–1074. <https://doi.org/10.1121/1.2213572>

766 Yang, H.-I., & Zeng, F.-G. (2017). Bimodal benefits in Mandarin-speaking cochlear implant users
 767 with contralateral residual acoustic hearing. *International Journal of Audiology*, 56(sup2), S17–
 768 S22. <https://doi.org/10.1080/14992027.2017.1321789>

769 Yang, X. J., Qu, C. Y., Sun, X. B., Zhu, F., & Wang, J. P. (2011). Norm revision of H-NTLA for
 770 children from 3 to 7 years old in China. *Chinese Journal of Clinical Psychology*, 19(2), 195–
 771 197.

772 Yip, M. (2002). *Tone*. Cambridge, UK: Cambridge University Press.

773 Yu, K., Li, L., Chen, Y., Zhou, Y., Wang, R., Zhang, Y., & Li, P. (2019). Effects of native language
 774 experience on Mandarin lexical tone processing in proficient second language learners.
 775 *Psychophysiology*, 56(11), e13448. <https://doi.org/10.1111/psyp.13448>

776 Zhang, H., Zhang, J., Ding, H., & Peng, G. (2019). Bimodal benefit in categorical perception of
 777 lexical tones for Mandarin-speaking children with cochlear implants. In *Proceedings of the*
 778 *International Congress of Phonetics Science* (Vol. 278).

779 Zhang, H., Zhang, J., Ding, H., & Zhang, Y. (2020). Bimodal benefits for lexical tone recognition:
 780 An investigation on Mandarin-speaking preschoolers with a cochlear implant and a contralateral
 781 hearing aid. *Brain Sciences*, 10(4), 238. <https://doi.org/10.3390/brainsci10040238>

782 Zhang, L., Wang, J., Hong, T., Li, Y., Zhang, Y., & Shu, H. (2018). Mandarin-speaking,
 783 kindergarten-aged children with cochlear implants benefit from natural F0 patterns in the use of
 784 semantic context during speech recognition. *Journal of Speech, Language, and Hearing*
 785 *Research*, 61(8), 2146–2152. https://doi.org/10.1044/2018_JSLHR-H-17-0327

- 786 Zhang, T., Dorman, M. F., Fu, Q.-J., & Spahr, A. J. (2012). Auditory training in patients with
787 unilateral cochlear implant and contralateral acoustic stimulation. *Ear and Hearing*, 33(6), e70–
788 e79. <https://doi.org/10.1097/AUD.0b013e318259e5dd>
- 789 Zhang, T., Dorman, M. F., & Spahr, A. J. (2010). Information from the voice fundamental frequency
790 (F0) region accounts for the majority of the benefit when acoustic stimulation is added to
791 electric stimulation. *Ear and Hearing*, 31(1), 63–69.
792 <https://doi.org/10.1097/AUD.0b013e3181b7190c>
- 793 Zhang, T., Spahr, A. J., Dorman, M. F., & Saoji, A. (2013). Relationship between auditory function
794 of nonimplanted ears and bimodal benefit. *Ear and Hearing*, 34(2), 133–141.
795 <https://doi.org/10.1097/AUD.0b013e31826709af>
- 796 Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M., & Tohkura, Y. (2005). Effects of language experience:
797 Neural commitment to language-specific auditory patterns. *NeuroImage*, 26(3), 703–720.
798 <https://doi.org/10.1016/j.neuroimage.2005.02.040>
- 799 Zhang, Y., Zhang, L., Shu, H., Xi, J., Wu, H., Zhang, Y., & Li, P. (2012). Universality of categorical
800 perception deficit in developmental dyslexia: An investigation of Mandarin Chinese tones.
801 *Journal of Child Psychology and Psychiatry*, 53(8), 874–882. [https://doi.org/10.1111/j.1469-](https://doi.org/10.1111/j.1469-7610.2012.02528.x)
802 [7610.2012.02528.x](https://doi.org/10.1111/j.1469-7610.2012.02528.x)
- 803 Zheng, H.-Y., Minett, J. W., Peng, G., & Wang, W. S. Y. (2012). The impact of tone systems on the
804 categorical perception of lexical tones: An event-related potentials study. *Language and*
805 *Cognitive Processes*, 27(2), 184–209. <https://doi.org/10.1080/01690965.2010.520493>

806 Zhou, N., Huang, J., Chen, X., & Xu, L. (2013). Relationship between tone perception and
807 production in prelingually-deafened children with cochlear implants. *Otology & Neurotology*,
808 34(3), 499–509. <https://doi.org/10.1097/MAO.0b013e318287ca86>

809

810 Table 1. Demographic information of the bimodal participants.

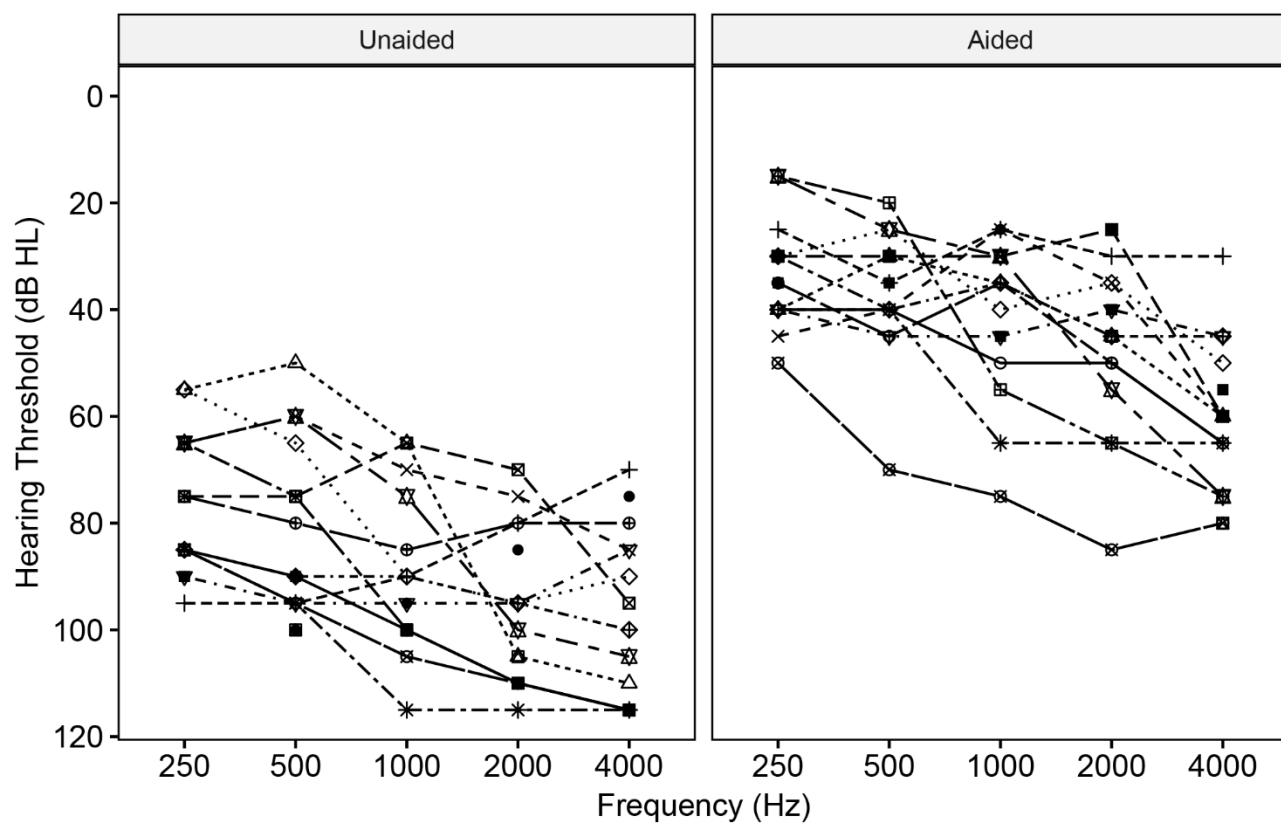
Subject (Sex)	CA (yrs)	CI (ear)	Speech strategy	HA	Age at CI	CI duration	Age at HA	HA duration	PTA
					(yrs)	(yrs)	(yrs)	(yrs)	
S1 (F)	5.6	Nucleus6 (L)	ACE	Phonak Q90 SP	3.3	2.3	2.4	3.2	77
S2 (F)	5.5	Nucleus6 (R)	ACE	Phonak SKY Q90-RIC	2.8	2.7	2.6	2.9	79
S3 (M)	5.3	Nucleus6 (L)	ACE	ReSound AL777	2.7	2.6	3.6	1.7	81
S4 (M)	5.7	OPUS2 (R)	FS4-P	Widex C3-FS	1.7	4	1.5	4.2	93
S5 (F)	6.2	Nucleus5 (R)	ACE	Phonak Naida S IX UP	1.1	5.1	1.2	5	71
S6 (M)	5.6	Naida (R)	HiRes-Optima	Phonak Naida S IX UP	2.7	2.9	2.2	3.4	80
S7 (F)	5.8	Nucleus5 (L)	ACE	Phonak Q90 UP	3.2	2.6	3.3	2.5	76
S8 (M)	5.9	OPUS2 (R)	FS4-P	Widex C4-FS	0.9	5	0.6	5.3	91
S9 (M)	5.6	OPUS2 (R)	FS4-P	Phonak Naida S IX UP	1	4.6	0.6	5	102
S10 (M)	4.8	Nucleus6 (R)	ACE	Phonak Naida S IX UP	1.5	3.3	1.4	3.4	101
S11 (F)	5.8	Neptune (R)	HiRes-Optima	Phonak Q70 SP	1	4.8	3.7	2.1	91
S12 (F)	4.7	Freedom (R)	ACE	Phonak Naida S V SP	1.2	3.5	1.1	3.6	105
S13 (M)	4.9	OPUS2 (R)	FS4-P	Phonak Bolero Q50 SP	0.9	4	0.6	4.3	101
S14 (M)	4.7	Nucleus5 (R)	ACE	Phonak Naida S IX UP	1.4	3.3	2.6	2.1	101
S15 (M)	5.3	Nucleus5 (L)	ACE	Widex C4-FS	3.2	2.1	0.5	4.8	89
S16 (F)	5.4	Nucleus6 (R)	ACE	Phonak Naida S IX UP	3	2.4	3.7	1.7	86

811 CA: chronological age; CI: cochlear implant; HA: hearing aid; yrs: years; PTA: unaided four-frequency pure tone average
812 across 250Hz, 500Hz, 1 kHz, 2 kHz, and 4 kHz for the nonimplanted ear; M: male; F: female.

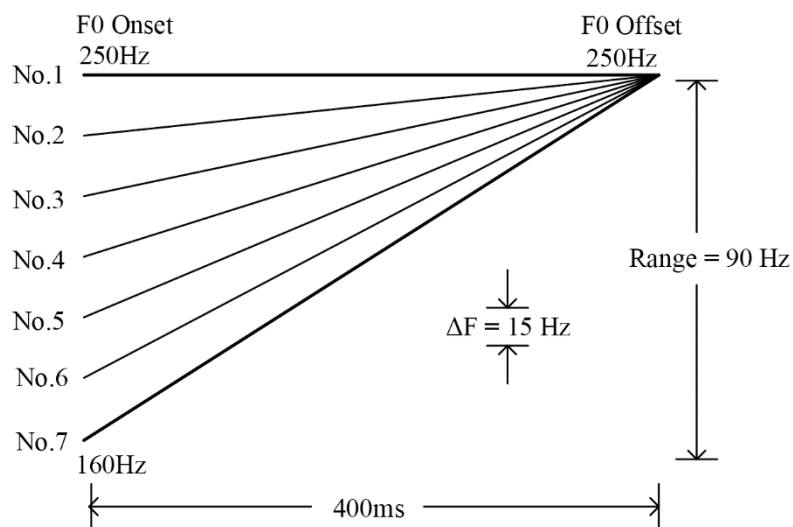
813

814

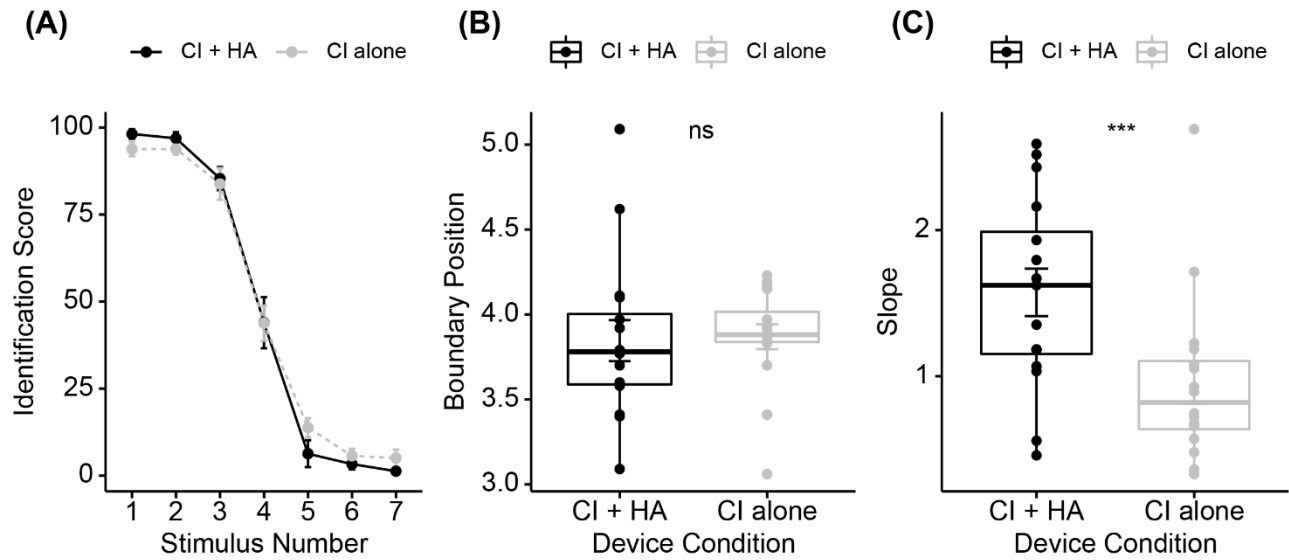
815



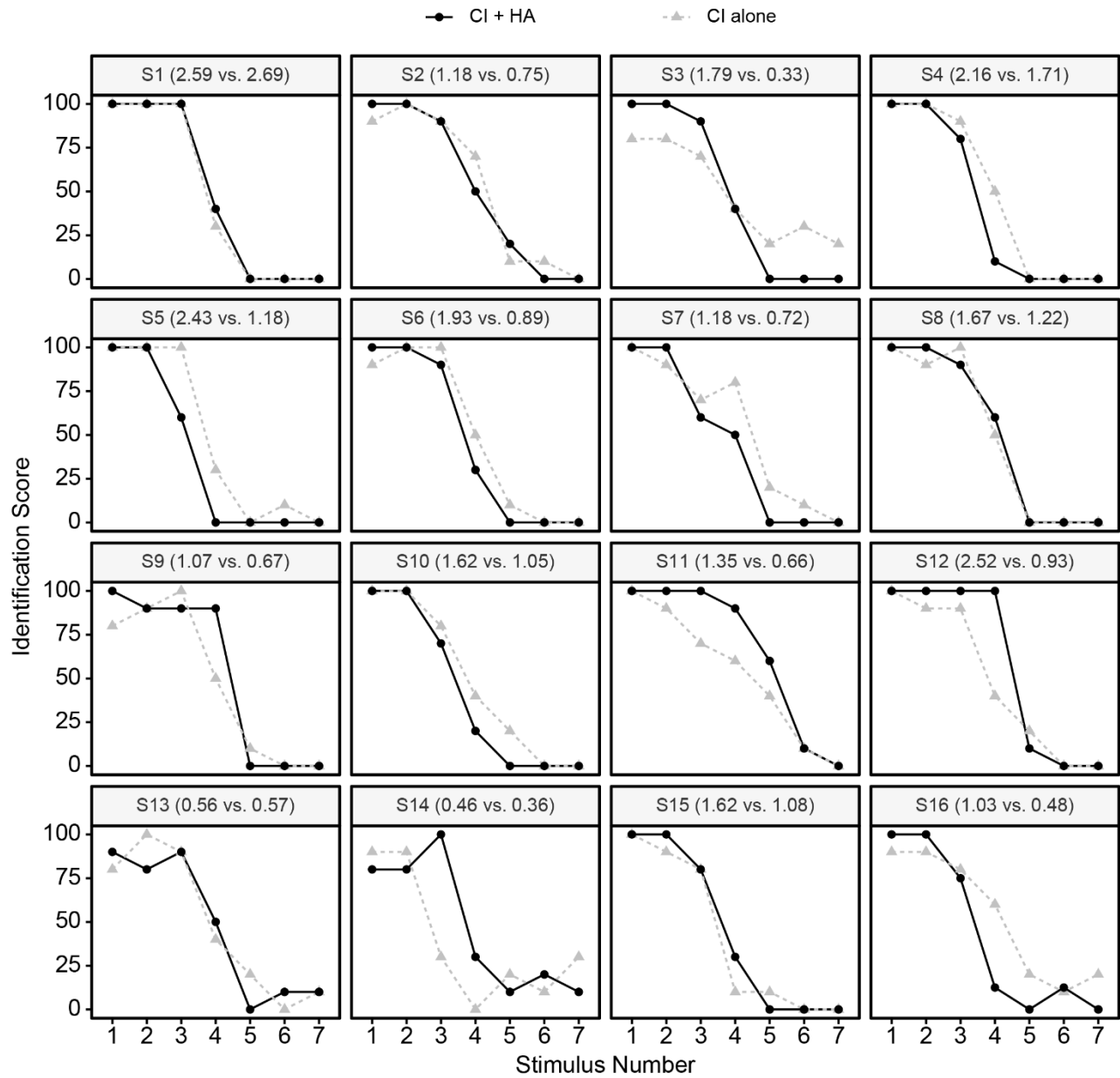
817 Figure 1. Unaided and aided hearing thresholds for the nonimplanted ear of each kindergartener
 818 participant in 250, 500, 1k, 2k, and 4k Hz.



820 Figure 2. Schematic diagram of the pitch contours for the stimuli in tonal categorical perception.

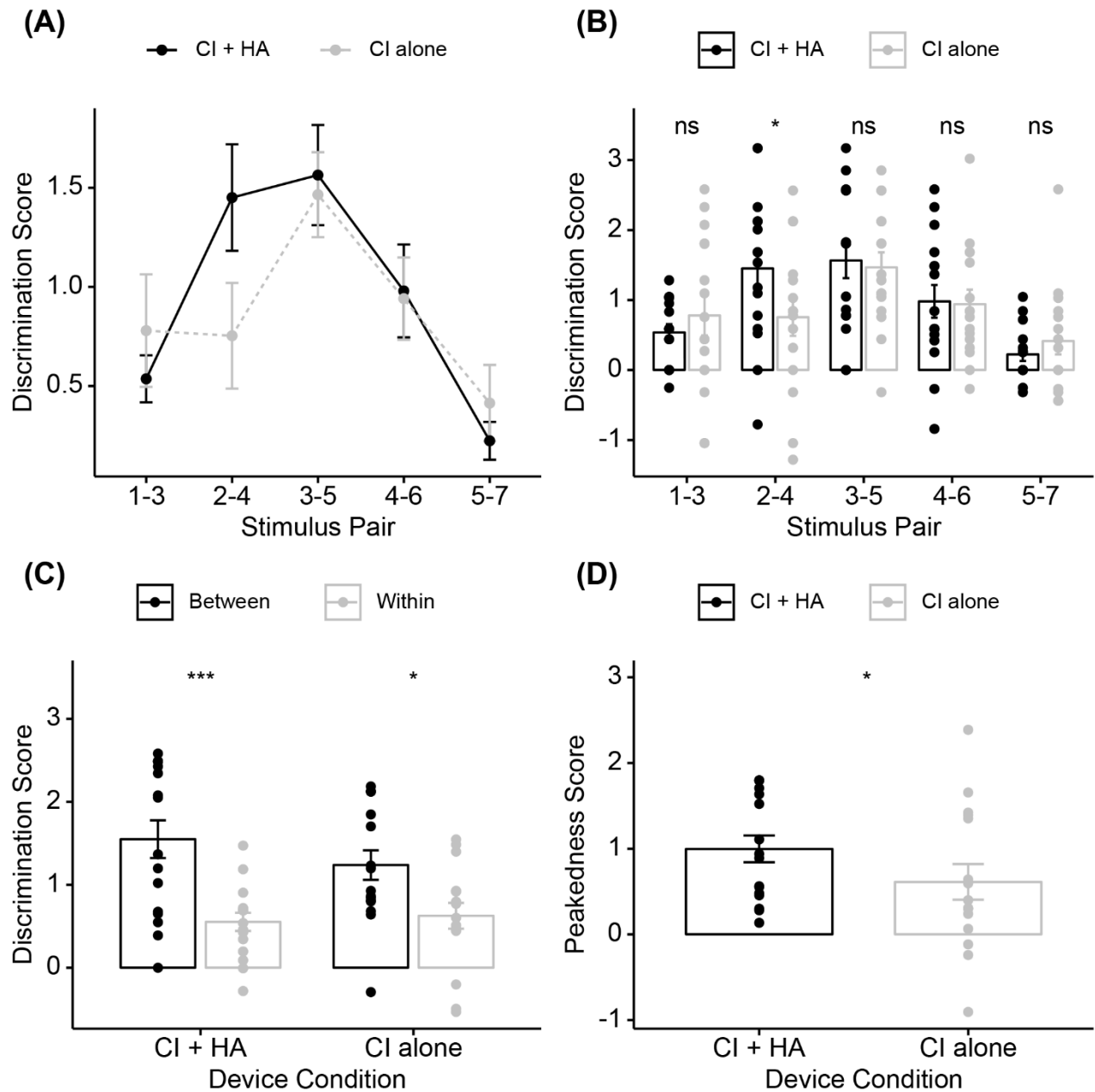


822 Figure 3. (A) Identification functions of T1 for CI + HA and CI alone device conditions; (B) Mean
823 boundary positions for CI +HA condition and CI alone condition; (C) Mean identification slopes for
824 CI +HA condition and CI alone condition. Points in (A) indicate the group mean data, whereas points
825 in (B) and (C) indicate the individual data of each participant. Error bars represent the standard errors
826 across all participants. Statistical significances are provided: three *asterisks* represent $p < 0.001$,
827 while *ns* represents $p > 0.05$.



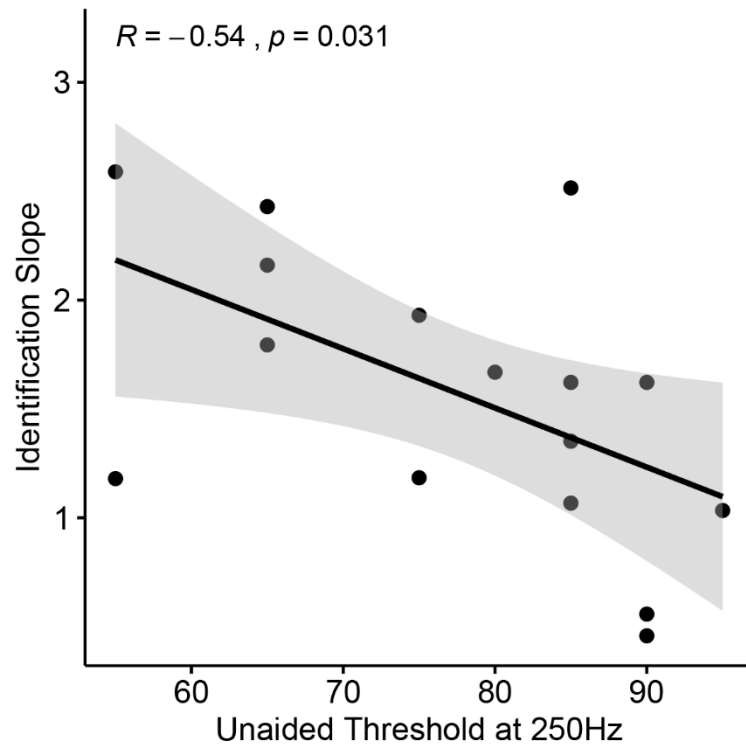
828

829 Figure 4. Identification functions of T1 for CI + HA and CI alone device conditions for each
 830 participant. Data in brackets indicate slopes for the two device conditions (CI + HA vs. CI alone).



832 Figure 5. (A) Discrimination functions for CI + HA and CI alone device conditions; (B) Mean
833 discrimination scores of different stimulus pairs for CI + HA condition and CI alone condition; (C)
834 Mean discrimination scores of between-category type and within-category type for CI +HA condition
835 and CI alone condition; (D) Mean peakedness score for each device condition. Points in (A) indicate
836 the group mean d' , whereas points in (B), (C), and (D) indicate the individual d' of each participant.

837 Error bars represent the standard errors across all participants. Statistical significances are provided:
838 three *asterisks* represent $p < 0.001$, one *asterisk* represent $p < 0.05$, while *ns* represents $p > 0.05$.



840 Figure 6. Relationship between unaided hearing threshold at 250 Hz and identification slope for the
841 bimodal kindergarteners. The grey shades indicate 95% confidence interval.