ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Intelligibility of Word-Initial Obstruent Consonants in Mandarin-Speaking Prelingually Deafened Children With Cochlear Implants

Jing Yang,[a] Xianhui Wang,[b] Jue Yu,[c] and Li Xu[b]

[a] Program of Communication Sciences and Disorders, University of Wisconsin–Milwaukee [b] Hearing, Speech and Language Sciences , Ohio University, Athens [c] Center for Speech and Language Processing, School of Foreign Languages, Tongji University, Shanghai, China

ARTICLE INFO

Article History:
Received May 12, 2022
Revision received September 19, 2022
Accepted February 21, 2023

Editor-in-Chief: Peggy B. Nelson
Editor: Rachael Frush Holt

https://doi.org/10.1044/2023_JSLHR-22-00268

ABSTRACT

Purpose: This study assessed the intelligibility of obstruent consonants in prelingually deafened Mandarin-speaking children with cochlear implants (CIs).
Method: Twenty-two Mandarin-speaking children with normal hearing (NH) aged 3.25–10.0 years and 35 Mandarin-speaking children with CIs aged 3.77–15.0 years were recruited to produce a list of Mandarin words composed of 17 word-initial obstruent consonants in different vowel contexts. The children with CIs were assigned to chronological age–matched (CA) and hearing age–matched (HA) subgroups with reference to the NH controls. One hundred naïve NH adult listeners were recruited for a consonant identification task that consisted of a total of 2,663 stimulus tokens through an online research platform. For each child speaker, the consonant productions were judged by seven to 12 different adult listeners. An average percentage of consonants correct was calculated across all listeners for each consonant.
Results: The CI children in both the CA and HA subgroups showed lower intelligibility in their consonant productions than the NH controls. Among the 17 obstruents, both CI subgroups showed higher intelligibility for stops, but they demonstrated major problems with the sibilant fricatives and affricates and showed a different confusion pattern from the NH controls on these sibilants. Of the three places (alveolar, alveolopalatal, and retroflex) in Mandarin sibilants, both CI subgroups showed the lowest intelligibility and the greatest difficulties with alveolar sounds. For the NH children, there was a significant positive relationship between overall consonant intelligibility and chronological age. For the children with CIs, the best fit regression model revealed significant effects of chronological age and age at implantation, with their quadratic terms included.
Conclusions: Mandarin-speaking children with CIs experience major challenges in the three-way place contrasts of sibilant sounds in consonant production. Chronological age and the combined effect of CI-related time variables play important roles in the development of obstruent consonants in the CI children.

For prelingually deafened children, the modern technology of cochlear implants (CIs) enables them to regain partial auditory sensation and to acquire phonetic inventory in a manner similar to typically developing children (Blamey et al., 2001; Bouchard et al., 2007; Serry & Blamey, 1999). Numerous studies have reported improved articulation scores and production skills with extended CI use in children with prelingual deafness (Ertmer & Jung, 2012; Leigh et al., 2013; McCaffrey et al., 1999; Tomblin et al., 2008; Tye-Murray et al., 1995). However, children with CIs still demonstrate perceivable deficits in their speech intelligibility in comparison to normal-hearing (NH) peers (Boonen et al., 2020; Chin et al., 2003; Freeman et al., 2017; Grandon et al., 2020; Poursoroush et al., 2015), which is associated with the impact of delayed access

Correspondence to Jing Yang: jyang888@uwm.edu. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

to auditory input (Moeller & Tomblin, 2015) and the distorted speech information conveyed through the CI devices (Casserly, 2015; Grandon & Vilain, 2020).

Speech intelligibility, defined as the extent to which a talker's speech utterances can be understood by a listener (Yorkston et al., 1996), is a critical measure to assess people's oral communication abilities and skills. Generally, there are two approaches to evaluate speech intelligibility in clinical populations including deaf and hard of hearing talkers (Allison, 2020; Lagerberg et al., 2014; Miller, 2013; Osberger, 1992). One approach is to ask listeners to rate the overall intelligibility of spontaneous speech on an interval scale. This method is fast and convenient to implement, but it is based on subjective judgment that has limited reliability. The other approach is to require listeners to identify collected speech samples including words or sentences in an open-set response format or a closed-set forced-choice format. This approach yields a percentage score that provides a more objective measure for robust statistical analysis. In this study, we adopted a closed-set phoneme identification task to assess consonant intelligibility in Mandarin-speaking children with CIs.

Consonant intelligibility at the segmental level is of interest in this study for several reasons. First, consonant production involves a delicate and complex coordination of respiratory, phonatory, and articulatory gestures. The complicated oromotor configurations are reflected in various manners and places of articulation. For typically developing children, regardless of the languages they use, consonants are acquired later than vowel sounds (Hua & Dodd, 2000; Locke, 1983; Templin, 1957; Vihman, 1996). Critical changes in consonant acquisition occur between the ages of 3 and 9 years (Sander, 1972; Smit et al., 1990). Second, consonants, especially high-frequency sounds such as /s/ and /ʧ/, pose a major difficulty in speech perception in the deaf and hard of hearing population (Chen et al., 2020; Ching et al., 1998; B. C. J. Moore, 2016; Qi et al., 2021; Zeng & Turner, 1990). The lack of accessibility to high-frequency sounds results in a longer delay in the acquisition of consonant phonemes, especially fricatives, than vowel phonemes in deaf and hard of hearing children (Stelmachowicz et al., 2004). Even with the assistance of a hearing prosthesis, CI users still experience challenges in perceiving and producing high-frequency sounds (Grandon & Vilain, 2020; Liker et al., 2007; Munson et al., 2003; Reidy et al., 2017; Summerfield et al., 2002; Todd et al., 2011), which has a detrimental effect on speech intelligibility. Third, speech problems in children are predominately caused by consonant-involved articulation errors and phonological processes. With reference to NH peers, children with CIs show persistent developmental and nondevelopmental phonological processes (Asad et al., 2018; Buhler et al., 2007; Eriks-Brophy et al., 2013).

Consonant Production in Children With CIs

The emergence and development of a consonant inventory in English-speaking children with CIs have been well documented in the literature (Blamey et al., 2001; Chin, 2003; Iyer et al., 2017; Kirk et al., 1995; Serry & Blamey, 1999; Sundarrajan et al., 2020; Tobey et al., 1991). Most of these studies used criteria based on the production accuracy judged by trained listeners to determine the acquisition of individual speech sounds. It was reported that some children with severe-to-profound hearing loss could produce consonants equally well as their NH peers could or even showed a more rapid acquisition of consonant inventories compared with their NH peers (Iyer et al., 2017; Sundarrajan et al., 2020). However, many studies revealed that children with CIs, as a group, still showed lower production accuracy and intelligibility in consonant production compared with their NH peers (Connor et al., 2006; Ertmer & Goffman, 2011; Huang et al., 2005; Tomblin et al., 2008; Warner-Czyz & Davis, 2008). Dillon and colleagues (Dillon, Cleary, et al., 2004; Dillon, Pisoni, et al., 2004) assessed the accuracy percentage of consonants produced by 88 English-speaking children with CIs using a nonword repetition task. These children received implantation at an age younger than 5 years old (M = 3.3) and had an average of 5.6 years (range: 3.8–7.5) of experience in implant use. The production of 20 nonwords was transcribed by two phonetically trained listeners and scored for segment, manner, place, and voicing feature accuracy. In the 76 children who produced more than 15 of the 20 target nonwords, the overall consonant accuracy was only 33%. Of all the examined places of articulation, these children produced coronal sounds more accurately than they did labial or dorsal sounds. Among different manners of articulation, the children demonstrated similar accuracy rates: 52%, 54%, 50%, and 46% for the target stops, fricatives, nasals, and liquids, respectively. As for the voicing feature, these children showed a similar accuracy rate for voiceless target sounds (59%) and voiced target sounds (63%). In the 24 children who repeated all 20 target nonwords, only 5% of nonwords were produced correctly, and the overall accuracy for consonants was 39%.

In a more recent study by Sundarrajan et al. (2020), the authors recruited 129 English-speaking children with CIs who received implantation between the ages of 6 and 38 months (at time points < 12, 12–18, 19–27, and 28–38 months) and were tested at ages 3.5 and 4.5 years. Conversational speech in a play session was elicited for phonetic transcription by speech-language pathology graduate students. The authors found that the children who received implantation before the age of 12 months had acquired the majority of consonants and performed on par with the age-matched NH children. By contrast, the children who received implantation between the ages of 28 and 38 months only mastered /b/ by the age of 3.5 years and /b, p, m, j, w, d, h, ʃ/ by the age of 4.5 years. Among different places, the children with CIs, across all subgroups, showed the highest accuracies for bilabial (> 78% correct) and glottal (> 73% correct) sounds but the lowest accuracies for dental and postalveolar sounds. In particular, the children with CIs implanted between the ages of 28 and 38 months and tested at 3.5 years of age only showed 25% accuracy for dental sounds and 27.3% accuracy for postalveolar sounds. Among different manners, the children with CIs produced stops with the highest accuracies (all > 70% accuracy except for the children implanted between the ages of 28 and 38 months and tested at 3.5 years of age) but produced affricates with the lowest accuracies (< 30% accuracy for children implanted between the ages of 28 and 38 months and tested at both 3.5 and 4.5 years of age).

Different from the studies that recruited trained or experienced listeners to do phonetic transcription, some studies recruited inexperienced or untrained listeners to judge speech performance (e.g., Chin et al., 2003; Huang et al., 2005; McAllister Byun et al., 2015). In such studies, listeners were usually required to rate speech intelligibility on a scale or to write down the words or sentences they heard. Chin et al. (2003) recruited 86 inexperienced adult listeners to judge the connected speech samples from 51 children with CIs and 47 children with NH. The children with CIs received implantation before 6 years of age, with an average of 2.4 years of CI experience. Each child was judged by three listeners, and the listeners were required to write down everything they heard from the child. Intelligibility was defined as a percentage of words correct converted from the average of correctly identified words across the three listeners. The results showed that the CI children were significantly less intelligible than the NH controls. Huang et al. (2005) recruited five inexperienced adult listeners to examine the speech intelligibility of 26 Mandarin-speaking children with CIs and 26 aged-matched NH children. The CI children were between 3.33 and 9.5 years old and received implantation between 1.08 and 7.94 years of age, with the postimplantation period ranging from 1.17 to 3.5 years. The speech material was composed of 37 words produced by each child. Compared to their NH peers (42.8% correct for words, 62.2% correct for consonants, 73.0% correct for vowels, and 75.8% correct for tones), the CI children showed much lower intelligibility on all four categories (18.2% correct for words, 40.4% correct for consonants, 53.6% correct for vowels, and 54.8% correct for tones).

In some studies, researchers recruited both untrained and trained listeners to examine speech intelligibility and production accuracy. In Tobey et al. (2003, 2011), 181 CI participants were tested at 8 or 9 years of age and 110 of them were retested a few years later when they were 15.0–18.5 years old). All children received implantation before the age of 5.33 years. The speech material consisted of 36 sentences composed of three, five, or seven syllables. Three NH adults with limited exposure to speech from the deaf and hard of hearing population were recruited to judge sentence intelligibility, and four trained speech-language pathologists were recruited to transcribe the speech samples for consonant and vowel production accuracy. Results of the Tobey et al. (2003) study showed that the speech intelligibility of the 181 CI participants was only 63.5% correct and that the accuracy of consonant production was 68%. In the follow-up test reported in the Tobey et al. (2011) study, the average speech intelligibility improved to approximately 86% for the high-context sentences and 82% for the low-context sentences. Consonant production improved to 93.8% correct.

Comparison Between Mandarin and English Consonants

Compared to the extensive investigation on speech performance in English-speaking children with CIs, only a few studies examined the speech production and speech development in pediatric CI users from other language backgrounds (e.g., Croatian: Mildner & Liker, 2008; Dutch: Faes & Gillis, 2016, 2018; French: Bouchard et al., 2007; Grandon & Vilain, 2020; Israeli Hebrew: Adi-Bensaid & Ben-David, 2010; Spanish: J. A. Moore et al., 2006). In this study, we attempted to examine the production of consonants, particularly obstruent consonants, in Mandarin-speaking children with CIs. Different from sonorants that are produced with little constrictions, *obstruents* refer to sounds that have airflow obstructed during articulation, such as stops, fricatives, and affricates. Obstruents were of particular interest because their inventory size and manner category are much larger than those of sonorant consonants in most languages (Lindblom & Maddieson, 1988).

Mandarin has a similar size of consonant inventory to that of English (see Table 1 for the Mandarin consonant chart), but Mandarin obstruent consonants are characterized by aspirated versus nonaspirated contrast, whereas English is characterized by voiced versus voiceless contrast. Additionally, Mandarin sibilants have a three-way place contrast of alveolar /s/, /ts/, and /tsʰ/ versus alveolopalatal /ɕ/, /tɕ/, and /tɕʰ/ versus retroflex /ʂ/, /tʂ/, and /tʂʰ/ in both fricatives and affricates (Ladefoged & Wu, 1984; Li & Munson, 2016). Among the three places of the sibilant contrast in Mandarin, alveolar sounds are the most fronted. Alveolopalatal sounds are produced in a relatively front region between alveolar and retroflex sounds (C.-Y. Lee et al., 2014; Li, 2008). The spectral energy of alveolar fricative /s/ can reach

Table 1. Mandarin consonant (pinyin and IPA) phonemes organized by place, manner, and voicing features.

Manner	Voicing	Place					
		Bilabial	Labiodental	Alveolar	Alveolopalatal	Retroflex	Velar
Stop	Voiceless unaspirated	b /p/		d /t/			g /k/
	Voiceless aspirated	p /pʰ/		t /tʰ/			k /kʰ/
Affricate	Voiceless unaspirated			z /ts/	j /tɕ/	zh /tʂ/	
	Voiceless aspirated			c /tsʰ/	q /tɕʰ/	ch /tʂʰ/	
Fricative	Voiceless		f /f/	s /s/	x /ɕ/	sh /ʂ/	h /x/
Nasal	Voiced	m /m/		n /n/			ng /ŋ/
Lateral approximant	Voiced			l /l/			
Approximant	Voiced			r /r/			

Note. IPA = International Phonetic Alphabet.

up to 9–10 kHz. The spectral peak of alveolopalatal /ɕ/ is lower than that of /s/ but is also in a relatively high-frequency region between 5 and 9 kHz. By contrast, retroflex sounds are characterized by spectral energy concentration in the lower frequency region between 3 and 5 kHz (C.-Y. Lee et al., 2014; S.-I. Lee, 2011; Li, 2008). With the unaspirated and aspirated distinction in each place of affricates, Mandarin has a total of nine sibilant consonants. By contrast, English has a two-way place contrast of alveolar /s/ versus palatoalveolar /ʃ/ in fricatives but no place contrast in affricates (Ladefoged, 2001). With the voiced and voiceless distinctions in each place of both fricatives and affricates, English has a total of six sibilant sounds /s, z, ʃ, ʒ, ʧ, ʤ/. Sibilant fricatives and affricates are acquired relatively late in both NH and CI children (McLeod & Crowe, 2018; Serry & Blamey, 1999), and these sounds pose audibility challenges to deaf and hard of hearing listeners. The more complex place contrast and a greater number of sibilant consonants in Mandarin provide a valuable source to examine articulation performance and speech development in children with CIs. The findings of the intelligibility and confusion patterns of Mandarin fricatives and affricates will help researchers and clinicians better understand the deficits of production and perception of high-frequency sounds, which will help guide the development of better rehabilitation approaches in prelingually deafened children with CIs.

This Study

The primary goal of this study was to add to our knowledge of consonant intelligibility, particularly the intelligibility of word-initial obstruent consonants, in Mandarin-speaking children with CIs. So far, there have been a few sporadic research reports on consonant production in Mandarin-speaking children with CIs (Y.-S. Lin & Peng, 2003; Peng et al., 2004; Yang et al., 2017). These studies, like most research on speech performance in pediatric CI users, adopted a routinely used approach in examining phonetic and phonological acquisition. The collected speech samples were transcribed (usually in broad transcription) by one or two experienced clinicians or phonetically trained experimenters. Although experienced and trained listeners may capture phonetic details and subtle differences to which inexperienced listeners are not sensitive (Munson et al., 2012), phonetic transcription by professionals does not represent how naïve listeners perceive speech in everyday communication settings. Furthermore, phonetic transcription is a "subjective" assessment that can be influenced by many factors (Shriberg & Kent, 2003). Transcriptions from one or two raters may show bias and be affected by the rater's experiences, even though the raters are well trained and highly experienced.

Alternatively, in this study, we recruited 100 naïve adult NH listeners to evaluate the consonant intelligibility of NH children and children with CIs. Distinct from previous studies that recruited naïve listeners for an open-set task to assess speech intelligibility (e.g., Chin et al., 2003; Huang et al., 2005), we adopted a closed-set consonant identification task. The listeners were asked to choose one of the consonants displayed on a computer screen that best matched what they heard. The identification responses were used to calculate the percentage of consonants correct across all listeners for each consonant in each child. In addition to percentage calculation, the identification responses were used to examine the confusion patterns of individual consonants. Specific research questions included the following.

1. What is the intelligibility of individual obstruent consonants in Mandarin-speaking children with CIs, and how does it differ from that in age-matched NH children?
2. What is the confusion pattern of obstruent consonants in children with CIs, and how does it differ from that in age-matched NH children?
3. Does the consonant intelligibility in children with CIs show a significant relationship with the variables of chronological age, age at implantation, and length of device use?

Based on the above-cited studies on English-speaking children with CIs, it was expected that the intelligibility of consonant production in our Mandarin-speaking CI children would be lower than that of the NH controls. Of the tested consonants, according to the studies on phonological acquisition in both NH and CI children (Iyer et al., 2017; Serry & Blamey, 1999; Smit et al., 1990), those early-acquired sounds such as bilabials and stops would be more intelligible, but the late-acquired fricatives and affricates would pose a greater challenge and show lower intelligibility. Correspondingly, the early-acquired sounds would show fewer confusions with other consonants, whereas the high-frequency sibilants would demonstrate greater confusions. The confusion patterns among the fricatives and affricates were of particular interest in this study. For the third research question, because consonant acquisition mainly occurs in the age range of 3–9 years in typically developing children (Smit et al., 1990) and both age at implantation and length of device use have been identified as two important time-related variables for CI children's speech-language abilities (e.g., Connor et al., 2006; Niparko et al., 2010), we predicted that all three time-related variables would play a role in CI children's consonant production.

Method

Listeners

This study included 100 native Mandarin-speaking listeners (60 women, 40 men) aged between 17 and 44 years (*M* = 23.6, *SD* = 8.0). All listeners grew up in China and spoke Mandarin in their daily life. Nine of the 100 listeners reported not coming from northern dialect (the basic dialect of Mandarin Chinese) regions. The listeners' dialect backgrounds were not a concern of this study because they all learned Mandarin after enrolling in elementary schools at 6 years of age and because all used Mandarin as their primary language in their everyday life. All listeners were naïve to the perception task, with no experience listening to deaf or CI speech. None of the listeners reported having speech-language, hearing, or cognitive problems. The listeners were recruited by personal invitation and word of mouth using chain-referral sampling. This study has been reviewed and approved by the institutional review board of the University of Wisconsin–Milwaukee.

Stimuli for Consonant Identification

The stimuli were speech productions collected from 22 children with NH and 35 children with CIs. All speakers were recruited from the Beijing area in China and spoke Mandarin as their primary language. The children with NH (14 girls, eight boys) were aged between 3.25 and 10.0 years (*M* = 6.19, *SD* = 1.65). The children with CIs (17 girls, 18 boys) were aged between 3.77 and 15.0 years (*M* = 8.22, *SD* = 2.58). The children with CIs were all prelingually deafened and received unilateral implantation at various ages ranging between 1.08 and 8.0 years (*M* = 3.30, *SD* = 1.55), with the length of CI use ranging between 0.08 and 9.49 years (*M* = 5.04, *SD* = 2.50). In the children with CIs, some had no rehabilitation or speech therapy, some received parent-guided home practice, and others participated in formal rehabilitation. Furthermore, those children who had rehabilitation services varied in the content and length of their therapy program. Therefore, the factor of rehabilitation experience was not used as a factor in further analysis.

All tested children were asked to produce 51 Mandarin disyllabic or trisyllabic words (see Appendix A for the full word list) that covered six Mandarin stops /p, pʰ, t, tʰ, k, kʰ/ (pinyin b, p, d, t, g, and k); five fricatives /f, s, ɕ, ʂ, x/ (pinyin f, s, x, sh, and h); and six affricates /ts, tsʰ, tɕ, tɕ , tʂ, tʂʰ/ (pinyin z, c, j, q, zh, and ch). The target consonants were located in the word-initial position, and each consonant occurred in three different vowel contexts /a, i, u/. Due to the phonotactic constraint in Mandarin, not all tested obstruents can occur with all three vowels. In this case, alternative vowel contexts were used. The tone environment of the syllables containing the target consonants was not strictly controlled, except that Tone 3 was avoided in order to reduce the potential confounding effect associated with the deficit in Tone 3 production in children with CIs (e.g., Mao et al., 2020; Zhou & Xu, 2008).

All speech productions were collected in a quiet room using an audiovisual word repetition task. For each word, a picture was presented on a computer screen, followed by an audio prime. Each child was required to repeat the word immediately following the audio prompt. The speech samples were recorded through a digital recorder (Zoom H4n) with a 44.1-kHz sampling rate and a 16-bit quantization rate. All recordings were transferred to a computer hard drive and were segmented into individual words, with the landmarks of syllable onsets and offsets measured manually. For each word, only the first syllable containing the target obstruents was used for the recognition task to avoid the impact of the lexical meaning of the entire word on the recognition performance. Due to missing tokens in some speakers, there were a total of 2,663 tokens used for the recognition test.

Consonant Recognition Procedure

The recognition task was implemented through an online research platform (http://www.gorilla.sc). Given the large number of tokens, the stimuli were divided into 10 subsets. The number of child speakers contained in each subset varied from eight to 10. Tokens from four children (two NH and two CI) were used for all 10 subsets. The responses of all listeners to the same four speakers were compared to ensure consistency among listeners. The other NH and CI children were randomized and roughly equally assigned to each subset. Within each subset, the presenting order of speakers was randomized, and the tokens in each speaker were also randomized. We regarded each speaker as a different block, and tokens from the same speaker were presented in the same block. Each subset was presented to multiple adult listeners, and each listener was randomly assigned to one subset. All listeners were required to conduct the test on a desktop or laptop computer with headphones in a quiet environment. The listeners could adjust the volume to their most comfortable level. Instructions and consent information were first provided at the beginning. No written consent was collected. Then, the listeners were asked to answer a few questions to collect basic demographic information, including gender, age, and language background. During the identification test, a practice session was provided to familiarize the listeners with the experimental procedure. The tokens used for the practice trials were from a typically developing child who was not included in the real test.

In both practice and real tests, a grid with 18 boxes was shown on a computer screen. Seventeen buttons were labeled with Mandarin pinyin, corresponding to the 17 tested obstruents, and one box was labeled "none of the above." Listeners were required to choose the button that matched what they heard immediately after a stimulus token was played. For each token, the listeners can listen no more than 3 times. The average testing time for each listener was approximately 30 min. The participation was voluntary with no incentives provided, and the participants could quit or withdraw at any time during the test. Only data from the listeners who completed the entire subset were used for further analysis. Because the listeners were automatically assigned to a subset test, the number of listeners varied from seven to 12 for individual speakers (except for the four ubiquitously tested children).

Data Analysis

An overall percent-correct score across all stimulus tokens was calculated for each of the four ubiquitously tested children for each listener, which was then converted into z scores. Among the 100 listeners, three listeners consistently showed extremely low accuracies (> 2.58 *SD*s corresponding to a 99% confidence level) for all four children. This indicates that the three listeners failed to perform in good faith for the identification task. Therefore, the three listeners were excluded, which resulted in a total of 97 listeners used for further analysis. Next, an overall percent-correct score across the four ubiquitously tested children was calculated for each obstruent consonant for each listener. These data were used to test the interrater agreement. The 97 listeners were divided into a set of three-listener groups (n = 32), in which each group contained three different randomly selected listeners, and each listener was only tested once. Then, an intraclass correlation coefficient (ICC) was calculated for each group, which indicated the interrater agreement among the three listeners in that group. The average ICC of the 32 groups was .77 (*SD* = 0.12, range: 0.43–0.95, with only two ICCs < .6). This analysis suggested a relatively high agreement among the 97 listeners.

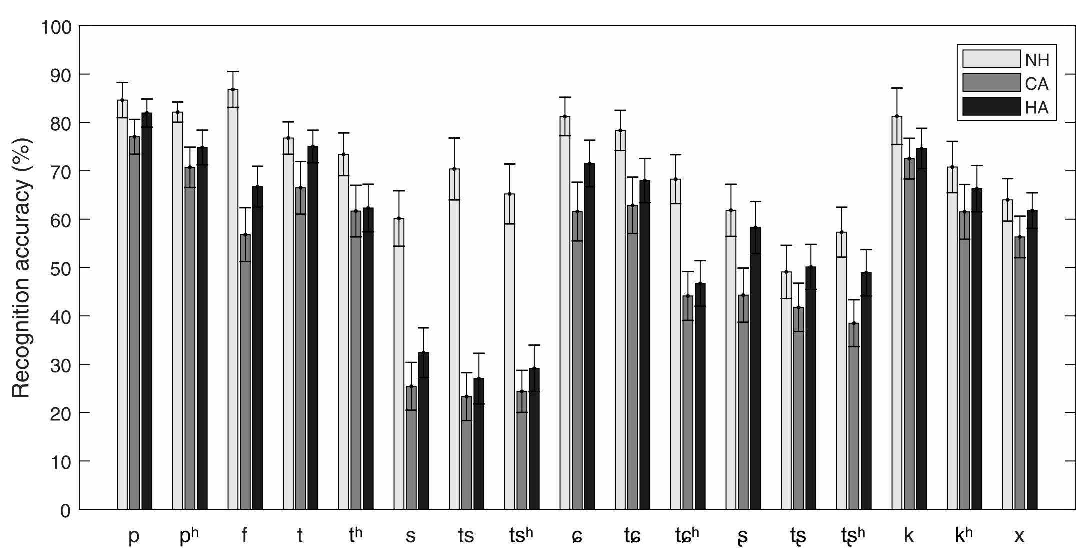
In this study, the children with CIs showed a wide range of chronological age and CI-related time variables. Although the majority of CI studies had the CI participants matched with NH controls based on their chronological age, many researchers examined the language and speech performance of CI children with reference to hearing age peers (e.g., Faes & Gillis, 2016; Geers et al., 2009; Schramm et al., 2010) because the amount of auditory input played a vital role in speech-language development (Nicholas & Geers, 2006; Niparko et al., 2010; Svirsky et al., 2000; Wie et al., 2007). In this study, the children with CIs were assigned into two subgroups. Those whose chronological age fell into the age range of the NH controls (i.e., 3.25–10.0 years) were labeled as the chronological age–matched (CA) subgroup, and those whose electrical hearing age (i.e., duration of CI use) fell into the age range of NH controls were labeled as the hearing age–matched (HA) subgroup. Note that the candidacy criteria for cochlear implantation in China are similar to those of the United States, but the implementation of the criteria tends to be more stringent in China. All of our children with CIs were prelingually deafened. The influence of presurgery acoustic experience from hearing aid, if there was any, should be very limited. Therefore, the length of electrical hearing use was defined as the hearing age of these CI children. The CA subgroup included 26 children aged between 3.77 and 10.0 years (*M* = 7.08, *SD* = 1.67). The HA subgroup included 26 children with hearing age ranging between 3.33 and 9.49 years (*M* = 6.12, *SD* = 1.88). Some children with CIs were both chronological age– and hearing age–matched with the NH children. Therefore, they were assigned to both subgroups. An independent-samples t test was implemented to test group differences in age between the NH and CA subgroups as well as between the NH and HA groups. The results presented no significant group difference in age (both *p*s > .05). None of the NH children were reported as having communicative or cognitive problems. None of the children with CIs were reported as having language problems or cognitive impairments.

For a given consonant sound produced by each child, intelligibility was defined as the percentage correct calculated across all listeners who were assigned to identify the stimulus tokens produced by this child. As the main purpose of this study was to compare consonant intelligibility in CI and NH children and many children with CIs were both chronological age– and hearing age–matched with the NH controls, the percentages were compared between the NH group and the two CI subgroups (CA and HA) separately. No comparison was made between the CA and HA subgroups. The percentages were fitted with a generalized linear mixed model (GLMM) using SPSS. The factors of group (NH and CA or NH and HA) and consonants were defined as fixed effects, and speaker was defined as a random effect, with a random intercept of speakers included. Then, a separate GLMM was used to compare the group differences for each consonant, with the percentage scores defined as the dependent variable, the group effect defined as a fixed effect, and the speaker effect defined as a random effect. Given that there were 17 comparison procedures in this study, the traditional approach of Bonferroni correction that controls the familywise error rate would be too conservative. This study adopted a false discovery rate (FDR) control. FDR provides an alternative procedure to control for a low proportion of false positives (Benjamini & Hochberg, 1995), which has been recommended for multiple comparisons in health studies (Glickman et al., 2014). Based on the listeners' response data, a confusion matrix was generated for each of the NH, CA, and HA groups to visualize the group-level confusion patterns. The confusion matrices were further used in a hierarchical clustering analysis (Ward, 1963). This analysis, using Ward's linkage algorithm in MATLAB, provided perceptual clusters of the consonants in a hierarchical structure that showed the perceptual distance of one consonant to the others. The hierarchical clustering analysis was conducted for each of the NH, CA, and HA groups. Finally, an overall percent-correct score across all 17 consonants was calculated for each child, which was used for regression analysis with the time-related variables, including chronological age, age at implantation, and length of device use.

Results

The percentage correct for each consonant in the NH group and the two CI subgroups is presented in Figure 1. For all tested consonants, the productions of the children with CIs, regardless of subgroup (CA or HA), were perceived with lower intelligibility compared with those of the NH children. The HA subgroup were perceived more similarly to the NH controls, compared with the CA subgroup. Among the three manners of articulation, the intelligibility of stops in the CA and HA subgroups approached that in the NH controls, but the intelligibilities of fricatives and affricates in both CA and HA children were lower than that in the NH children. Among different places of articulation, the percentages of alveolar sounds, especially alveolar fricatives and affricates, in both CA and HA children were dramatically lower than those in the NH controls. Additionally, the labiodental and retroflex sounds in the CA and HA

Figure 1. Bar plots showing the group mean and standard error of the percentage correct of individual obstruent consonants in the normal-hearing (NH) children and children with cochlear implants assigned into chronological age–matched (CA) and hearing age–matched (HA) subgroups.



children showed lower percentages than those in the NH children. For the comparison between the NH and CA groups, the GLMM analysis revealed significant main effects of group, $F(1, 769) = 14.16$, $p < .001$, and consonant, $F(16, 769) = 362.43$, $p < .0001$, and a significant interaction effect between group and consonant, $F(16, 769) = 82.41$, $p < .0001$. For the comparison between the NH and HA groups, the GLMM results revealed significant main effects of group, $F(1, 770) = 8.68$, $p = .003$, and consonant, $F(16, 770) = 335.37$, $p < .0001$, and a significant interaction effect between group and consonant, $F(16, 770) = 80.91$, $p < .0001$. In the subsequent analysis for each consonant (shown in Table 2), with the FDR of .05 being controlled for multiple comparisons, the GLMM results revealed statistically lower intelligibility for the consonants /f, s, ts, tsʰ, tɕʰ/ (all FDR-adjusted $ps < .05$) in the CA subgroup with reference to the NH controls. For the HA subgroup, the same five consonants (i.e., /f, s, ts, tsʰ, tɕʰ/) were produced with lower intelligibility with reference to the NH controls (all FDR-adjusted $ps < .05$).

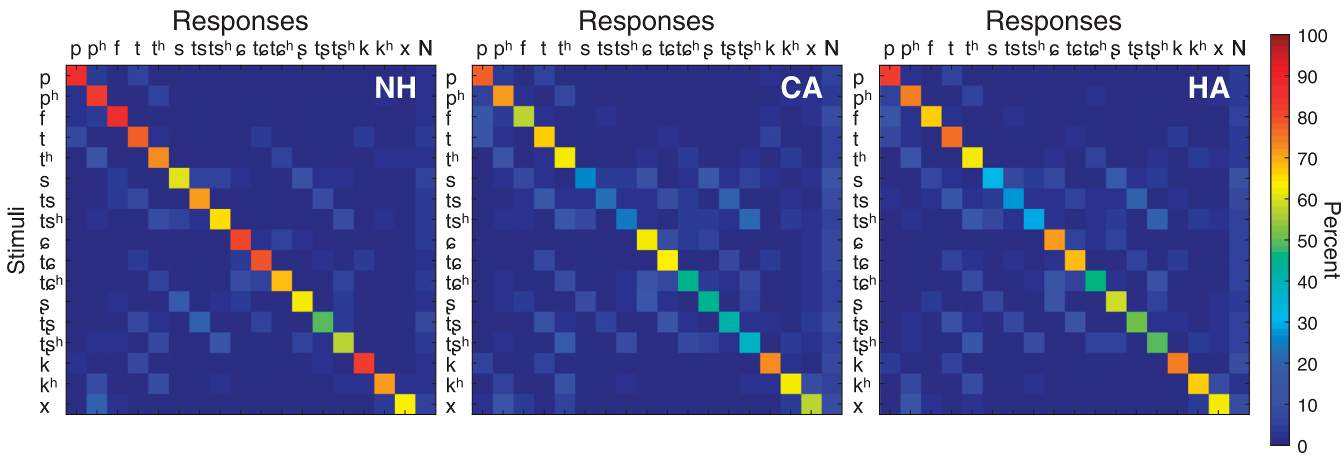
Figure 2 presents the confusion pattern of the 17 obstruent consonants in the NH and CI children (see Appendix B for confusion matrix data). Consistent with the trend shown in Figure 1, the productions of all tested consonants in the children with CIs, regardless of subgroup (CA or HA), were less intelligible and showed a higher degree of confusion than those in the NH group. For all three groups of children especially the two CI subgroups, greater confusion occurred among the three places of sibilants. In the NH children, the greatest confusion occurred between the alveolar and retroflex sounds. Not many confusions occurred between the alveolopalatal sounds and the other two places, although certain confusions occurred across the alveolopalatal fricatives and affricates. Similar to the NH children, the children with CIs showed major confusions between the alveolar and retroflex sounds. However, unlike the NH children who had a higher percentage of retroflex sounds being misperceived as alveolar sounds, the children with CIs, including both the CA and HA subgroups, showed a reverse pattern: Their productions of alveolar sounds were misperceived as retroflex sounds more than retroflex sounds being misperceived as alveolar sounds. Another major difference between the NH group and the two CI subgroups was that both alveolar and retroflex sounds in the two CI subgroups were misperceived as alveolopalatal sounds, which happened less frequently in the children with NH. The confusion patterns of the two CI subgroups also differed from the NH group in that more productions from the children with CIs were perceived as "none of the above" in comparison to the NH controls. These observations suggest that the three places of sibilants produced by the children with CIs were less identifiable in comparison to the NH group. Although the CI subgroups and NH controls demonstrated different features in the confusion patterns, there are some noticeable similarities. Both the NH and CI children showed confusions among the three places for stops and used the alveolar stops to substitute all affricates. Furthermore, the two CI subgroups showed considerable confusions within the three alveolopalatal sounds /ɕ, tɕ, tɕʰ/, just like their NH peers.

Figure 3 displays the hierarchical clustering of the 17 consonants in each group based on the confusion matrix data. For all three groups, the greatest confusion

Table 2. Summary table showing the *F* values, the original *p* values of the generalized linear mixed model, and the false discovery rate (FDR)–adjusted *p* values for the comparisons between the normal-hearing and chronological age–matched groups (left) and between normal-hearing and hearing age–matched groups (right) for individual consonants. Significant differences are indicated by the asterisk.

Consonant	*F*	Original *p*	FDR-adjusted *p*	Consonant	*F*	Original *p*	FDR-adjusted *p*
p	3.75	.059	.091	p	1.20	.280	.433
pʰ	4.10	.049	.087	pʰ	1.71	.198	.374
f	17.95	< .001	< .001*	f	13.52	< .001	< .001*
t	2.27	.139	.179	t	0.16	.696	.789
tʰ	3.30	.076	.108	tʰ	3.65	.062	.176
s	20.61	< .001	< .001*	s	12.79	< .001	< .001*
ts	26.32	< .001	< .001*	ts	22.30	< .001	< .001*
tsʰ	24.21	< .001	< .001*	tsʰ	19.33	< .001	< .001*
ɕ	5.67	.022	.062	ɕ	1.85	.180	.374
tɕ	4.64	.037	.079	tɕ	3.37	.073	.177
tɕʰ	11.20	.002	.007*	tɕʰ	10.31	.002	.007*
ʂ	4.02	.051	.087	ʂ	0.10	.748	.795
tʂ	1.23	.273	.273	tʂ	0.02	.881	.881
tʂʰ	5.12	.029	.070	tʂʰ	0.82	.370	.524
k	2.18	.147	.179	k	1.46	.234	.398
kʰ	1.74	.193	.219	kʰ	0.56	.459	.600
x	1.62	.209	.222	x	0.27	.608	.738

Figure 2. Confusion matrices for the recognition of consonants produced by the normal-hearing (NH; left panel) children and children with cochlear implants subgrouped into chronological age–matched (CA; middle panel) and hearing age–matched (HA; right panel). The stimuli are represented by the ordinate, and the group-pooled responses are represented by the abscissa. In each small square, the color represents the percentage of the stimulus–response pair. The rightmost column labeled "N" represents responses as "none of the above."



(represented as the shortest dissimilarity distance in each group) always occurred between the alveolar and retroflex sounds, that is, /s/ and /ʂ/, /tsʰ/ and /tʂʰ/, and /ts/ and /tʂ/. Despite this common feature, there were differences in the clustering patterns between the CI and NH children. In the two CI subgroups, the alveolopalatal affricates were clustered with the alveolar and retroflex affricates. Specifically, /tɕʰ/ showed perceptual clustering with /tsʰ/ and /tʂʰ/, and /tɕ/ showed perceptual clustering with /ts/ and /tʂ/. Meanwhile, the HA children showed perceptual clustering of /ɕ/ with /s/ and /ʂ/, which resulted in consistent confusions between alveolopalatal sibilants and the other two places in this group. These observations indicated that the children with CIs did not produce identifiable place contrast for the alveolar, alveolopalatal, and retroflex sounds. By contrast, for the production of the children with NH, there was certain perceptual clustering within the alveolopalatal sounds between /ɕ/ and /tɕʰ/, but these two alveolopalatal sounds in the NH children showed a greater perceptual dissimilarity from the alveolar and retroflex sounds in comparison to the CA and HA children. Also, the unaspirated alveolopalatal affricate /tɕ/ did not show considerable grouping with other consonants in the NH children. These differences suggested that the

Figure 3. Dendrogram plot showing the hierarchical clustering of the 17 consonants based on the confusion matrix data of the children with normal hearing (NH) and cochlear implants. The *y*-axis represents the dissimilarity distance calculated using Ward's algorithm. For a given sound, the nearest neighbor within a cluster shows the greatest perceptual similarity or confusion. Each cluster with a dissimilarity distance shorter than 0.8 shares a color. CA = chronological age–matched subgroup; HA = hearing age–matched subgroup.

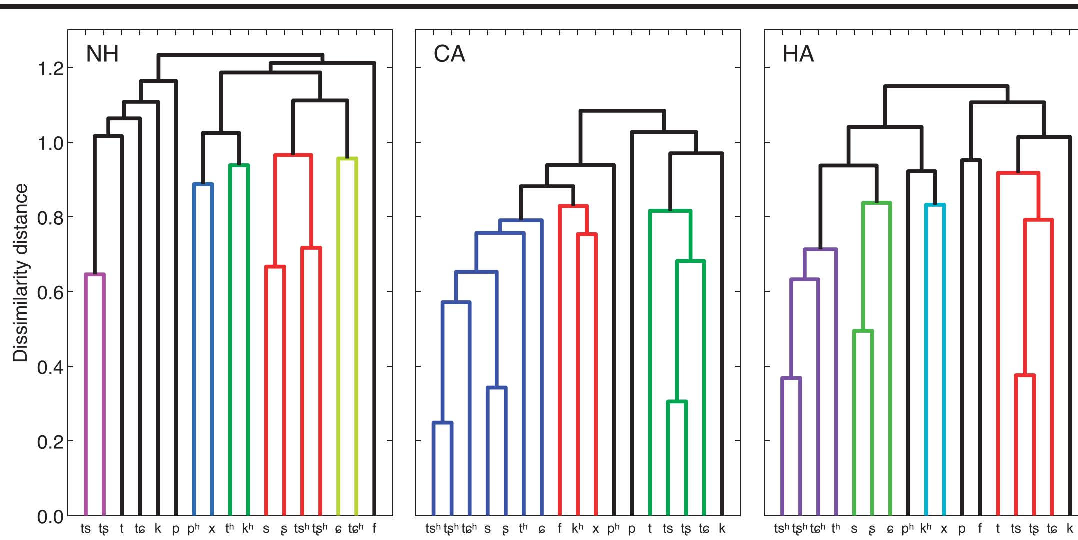
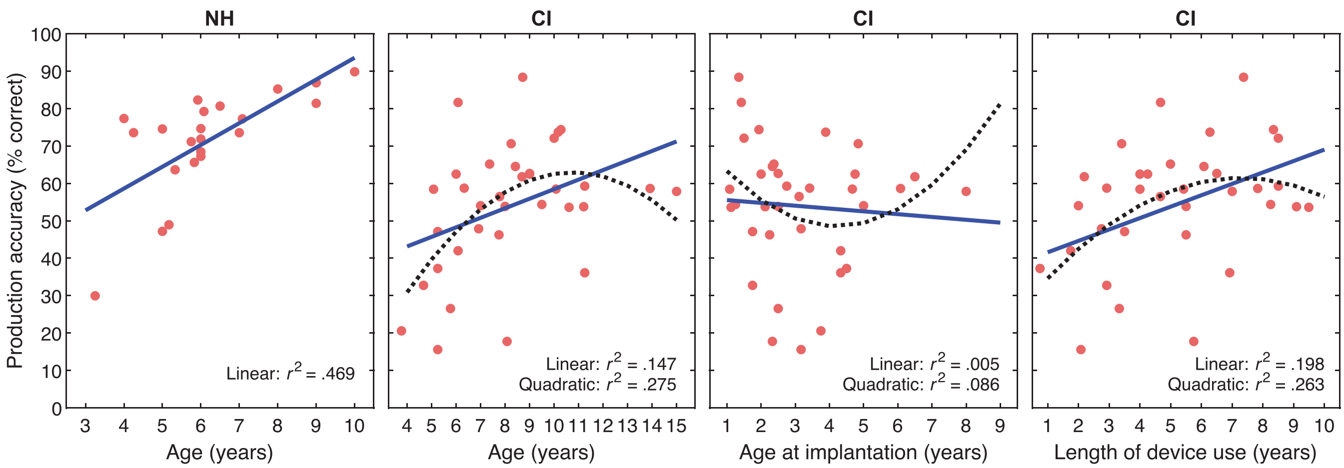


Figure 4. Scatter plots presenting the relations between the overall intelligibility of obstruent consonants and variables including chronological age and cochlear implant (CI)-related time variables for the 22 normal-hearing (NH) children and the 35 children with CIs. For the children with CIs, linear and quadratic relationships between each variable and consonant intelligibility were presented.



alveolopalatal sounds produced by the NH children were perceived as less similar to the alveolar and retroflex sounds in comparison to those produced by the CI subgroups. Another difference between the NH and CI children was that the two CI subgroups showed clustering between /t/ with /ts, tʂ, tɕ/ and /tʰ/ with /tsʰ, tʂʰ, tɕʰ/. However, for the NH children, clustering occurred among the three aspirated stops and the velar fricative /x/.

Our last research question addressed the relationship between time-related variables and consonant intelligibility in the tested children. The age range of both NH and CI children in this study was in the sensitive period of consonant development (Sander, 1972; Smit et al., 1990). A linear regression analysis was conducted for the NH children, which showed that the NH children's chronological age significantly predicted the intelligibility of their consonant production, $r^2 = .469$, $F(1, 20) = 17.66$, $p < .001$. For the 35 children with CIs, three time-related variables were of interest to us. Regression analyses were implemented to examine the contribution of these factors to consonant intelligibility in the CI children. Note that the CI children's chronological age was highly correlated with the factor of length of device use, $r = .786$, $p < .0001$. Therefore, these two factors were not used in the same model to avoid multicollinearity. Meanwhile, a close examination of the data revealed that the time-related variables and consonant intelligibility did not show a simple linear relationship (shown in Figure 4). When running regression models with the tested variables, the linear and quadratic (feature nonlinear relationship) terms were included and compared to find the best fit model. As shown in Table 3, the best fit model included chronological age and age at implantation, with quadratic terms added for both factors, $F(4, 30) = 4.787$, $p = .004$. The coefficients of the best fit model (see Table 4) show that both factors of chronological age and age at implantation, including their quadratic terms, played significant roles in the CI children's consonant intelligibility.

Discussion

The goal of this study was to examine the production of obstruent consonants in Mandarin-speaking children with CIs. Mandarin has a three-way place contrast in sibilants in which the alveolar and alveolopalatal fricatives and

Table 3. Comparison of regression models for the relationship between the overall intelligibility of obstruent consonants and chronological age (Age), age at implantation (AOI), and length of cochlear implant (CI) use (LOU) in the CI children.

Model	Variables	R^2
1 main effect (linear)	Age	.147
	AOI	.005
	LOU	.198
1 main effect (quadratic)	Age + Age_square	.275
	AOI + AOI_square	.086
	LOU + LOU_square	.263
2 main effects (linear)	Age + AOI	.202
	AOI + LOU	.201
2 main effects (quadratic)	Age + Age_square + AOI + AOI_square	.390
	AOI + AOI_square + LOU + LOU_square	.278
2 main effects (linear) + interaction	Age + AOI + Age × AOI	.203
	AOI + LOU + AOI × LOU	.218

Table 4. The parameter estimates of the best fit regression model with the factors of chronological age (Age) and age at implantation (AOI), with the quadratic terms of both factors (Age_square and AOI_square) included.

Parameter	Beta	*SE*	*t*	*p*
Intercept	−11.279	24.712	−0.456	.651
Age	20.595	6.169	3.338	.002
Age_square	−1.078	0.368	−2.931	.006
AOI	−16.230	6.856	−2.367	.025
AOI_square	2.132	0.939	2.271	.030

Note. SE = standard error.

affricates /s, ts, tsʰ, ɕ, tɕ, tɕʰ/ all have acoustic energy concentrated in a high-frequency region, whereas the retroflex fricatives and affricates /ʂ, tʂ, tʂʰ/ have spectral energy concentrated in a lower frequency region (C.-Y. Lee et al., 2014; Li, 2008). The rich number of sibilants in Mandarin provides a good source to examine the production and perception of high-frequency sounds in the CI population. In this study, we recruited 100 naïve listeners, and each listener was assigned a portion of the stimulus set. Each child speaker was rated by multiple listeners. Due to the wide age range of the CI children, they were assigned into CA and HA subgroups with reference to the NH controls.

Intelligibility of Obstruent Consonants in NH and CI Children

The first research question focused on the intelligibility of individual obstruent consonants in children with CIs. In this study, the recruited NH and CI children were all at the age within the crucial period of consonant development (Sander, 1972; Smit et al., 1990). In the NH control group, we observed higher intelligibility for the stops and fricative /f/ but lower intelligibility for the affricates and retroflex sounds. It has been well documented that, in typically developing English-speaking children, stops and front sounds are usually acquired before the age of 4 years but that fricatives (/s/, /z/, /ʃ/, /ʒ/, /θ/, /ð/) and affricates (/ʧ/, /ʤ/) are not fully acquired until the age of 7 or 8 years (using a 90% mastery criterion; Sander, 1972). Hua and Dodd (2000) reported that, in Mandarin-speaking children, retroflex sounds and affricates are among the most difficult and late-acquired sounds. Our observation in this study is consistent with the pattern reported in these studies.

Compared with NH children, the children with CIs showed a similar pattern of higher intelligibility for the six stops than for the other sounds. This outcome conforms to the finding of the early acquisition of stops in children with CIs (Blamey et al., 2001; Peng et al., 2004; Warner-Czyz & Davis, 2008). However, the children with CIs, including both the CA and HA subgroups, produced all consonants with lower intelligibility compared with those produced by the NH controls. The lower intelligibility was mainly reflected in the alveolar sounds /s, ts, tsʰ/. Our data showed that the intelligibility of the three alveolar sounds was higher than 60% correct in the NH children but lower than 26% and 33% correct in the CA and HA children, respectively (see Figure 1). Among all tested consonants, the three alveolar sounds demonstrated the greatest group differences between the NH and CI children. Meanwhile, the percentages of the three alveolar sounds were much lower than those of the other consonants in both CI subgroups. The difficulty with the /s/ sound has been reported in previous acoustic studies that compared the spectral features of /s/ and /ʃ/ between children with CIs and age-matched NH participants (Grandon & Vilain, 2020; Liker et al., 2007; Reidy et al., 2017; Todd et al., 2011). In these studies, the authors found that CI participants had a lower-than-normal spectral peak and center of gravity in the /s/ sound, but not in the /ʃ/ sound. The acoustic deviations in the high-frequency sounds were reflected in the lower intelligibility of these sounds in children with CIs compared with that in children with NH (Peng et al., 2004; Reidy et al., 2017; Yang et al., 2017). For example, Peng et al. (2004) and Yang et al. (2017), both targeting Mandarin-speaking children with CIs, reported the lowest accuracy of the /s/ sound in comparison to other tested sounds.

Researchers proposed that problems with /s/ in children with CIs could be explained by the inaccessibility of high-frequency information (Loizou, 2006; Reidy et al., 2017). Because the upper limit of output frequency through CI devices is approximately 8 kHz but the spectral peak and spectral mean of /s/ could be higher than 8 kHz (Jongman et al., 2000; Reidy et al., 2017; Yang et al., 2018), only partial spectral information of the high-frequency sound is conveyed by the CI processor. The lack of complete acoustic information through the CI device results in a perceptual deficiency in this sound. Given the inherent link between speech perception and production (Casserly & Pisoni, 2010) and the important role of auditory input in speech development in children with CIs (Nicholas & Geers, 2006; Niparko et al., 2010), the inaccessibility of high frequencies is manifested as low intelligibility in their production. By contrast, for the sounds produced at a further back region such as palatal fricative /ʃ/, the acoustic energy is concentrated in a lower frequency region between 4 and 6 kHz (Jongman et al., 2000). The spectral details of the low-frequency sound can be maintained and transmitted well through the CI device. Therefore, the intelligibility of low-frequency sounds is maintained relatively high. In addition to the frequency constraint of the CI processor, other factors may contribute to the deficiency in perceiving and producing high-frequency sounds. One possible factor is that the bandwidth of frequency bands in the CI processor increases as the frequency moves up (Dorman et al., 1998; Loizou, 1999). As a result, the spectral details of the frication noise for high-frequency sounds are less precisely conveyed in comparison to those for low-frequency sounds. Another potential factor is that spiral ganglion cell death may be more severe in the basal half than in the apical half (Hinojosa & Lindsay, 1980; Zimmermann et al., 1995), which induces more difficulty in processing and accessing high-frequency components in people with severe-to-profound sensorineural hearing loss.

In addition to the three alveolar sounds /s, ts, tsʰ/, our data revealed that the children with CIs, including both the CA and HA subgroups, showed significantly lower percent-correct scores compared with the NH group for the labiodental fricative /f/ and the alveolopalatal aspirated affricate /tɕʰ/ (see Figure 1). Although the spectrum of /f/ is relatively flat in comparison to the sibilants, the articulation of /f/ is in the front region of the oral cavity. The low intelligibility of fricatives and affricates produced in the relatively front region in CI children was reported in other studies (e.g., Peng et al., 2004). It is noteworthy that although the alveolopalatal aspirated affricate /tɕʰ/ showed significantly lower intelligibility than that in the NH children, the percentages of the alveolopalatal fricative /ɕ/ and the unaspirated affricate /tɕ/ in the children with CIs, especially in the HA subgroup, were close to those in the NH children, who showed very high intelligibility for these two sounds (see Figure 1). The performance gap among the three alveolopalatal sounds /ɕ, tɕ, tɕʰ/ was also found in Peng et al. (2004), in which the percent-correct score of the aspirated affricate /tɕʰ/ was 48.3% correct, much lower than that of /ɕ/ (55.4% correct) and /tɕ/ (69.17% correct). The higher intelligibility in alveolopalatal sounds /ɕ/ and /tɕ/ compared with that in the alveolar and retroflex sounds in the CI children might be due to the phonetic environment of the alveolopalatal sounds being different from that of the other two places. According to the description of the Mandarin vowel system (T. Lin & Wang, 2001), Mandarin has six single vowels, namely, a, i, u, ü, e, and o (International Phonetic Alphabet: /a/, /i/, /u/, /y/, /ɤ/, and /o/), and 13 compound vowels (not including nasal finals), namely, ai /aɪ/, ao /ɑʊ/, ie /iɛ/, ia /iA/, uo /uo/, ua /uA/, üe /yɛ/, ei /eɪ/, ou /oʊ/, iao /iɑʊ/, iou /ioʊ/, uai /uaɪ/, and uei /ueɪ/. The vowel phoneme /i/ has three allophonic variants, namely, [i], [ɿ] (high front apical vowel), and [ʅ] (high back apical vowel), which are all written as "i" in pinyin. Of the three places of sibilant consonants, alveolopalatal sounds are followed by the variant [i], whereas the alveolar and retroflex sounds are followed by the apical variants (high front apical [ɿ] after alveolar sounds and high back apical [ʅ] after retroflex sounds). For the other vowels, the alveolopalatal sounds can be followed by vowels ü, üe, ia, ie, iao, and iou. By contrast, the alveolar and retroflex fricatives and affricates are followed by vowels a, o, e, u, ao, ai, ei, ua, uo, ou, uai, and uei. The phonetic environment of the alveolopalatal sibilants is complementary to that of the other two places. This phonotactic constraint is associated with the distinct tongue shape of a bunched and raised tongue dorsum during the articulation of alveolopalatal sounds in comparison to the other two places (Li, 2008; Li & Munson, 2016). As the perceptual stimuli in this study contained the entire consonant–vowel syllable, the distinct vowel contexts following the alveolopalatal sounds helped improve listeners' recognition. However, for the significantly lower intelligibility of /tɕʰ/ in the CI subgroups compared to that in the NH controls, we speculated that this might be associated with the developmental trend that aspirated sounds are stabilized later than unaspirated sounds (Hua & Dodd, 2000; McLeod & Crowe, 2018). Meanwhile, the involvement of aspiration features may pose additional difficulty for CI children to produce.

With regard to the retroflex sounds, both CI subgroups showed relatively low intelligibility (an average of 41% correct for the CA subgroup and an average of 52% correct for the HA subgroup) but did not show differences from the NH controls (an average of 56% correct) as great as for the alveolar and alveolopalatal sounds. It is noteworthy that the three retroflex sounds produced by the NH children had the lowest intelligibility among all tested consonants. The low intelligibility of the retroflex sounds in both NH and CI children and the lack of NH–CI difference in the retroflex sounds suggest that the difficulty in producing these sounds in the children with CIs was associated with the more complex articulatory gestures involved in producing retroflex sounds compared with those involved in producing the other sounds.

Confusion Patterns of Obstruents in Children With CIs

The second research question addressed the features of confusion patterns in the CI and NH groups. Although for all three groups of children, the greatest confusion occurred in the s–ʂ, ts–tʂ, and tsʰ–tʂʰ pairs (see Figures 2 and 3), the children with NH tended to substitute /ʂ, tʂ, tʂʰ/ with /s, ts, tsʰ/ because the retroflex sounds involve a more complex articulatory configuration than the alveolar sounds. By contrast, the two CI subgroups showed much more substitutions of /s, ts, tsʰ/ with /ʂ, tʂ, tʂʰ/, which indicates that the high-frequency alveolar sounds were more difficult to produce than the retroflex sounds for the CI users. In addition, there were more confusions between the alveolopalatal sounds and the other two places in the two CI subgroups but not in the children with NH. Another noteworthy point shown in the confusion matrices was that both NH and CI children (including CA and HA subgroups) used alveolar stops to substitute the six affricates, but this confusion occurred more often in the CI subgroups. Mandarin has three pairs of affricates, and they all share the common component of alveolar stop. The consistent pattern of using an alveolar stop for all six affricates in both NH and CI children supported the finding that affricates are more difficult and are acquired later than the other types (Hua & Dodd, 2000; McLeod & Crowe, 2018). This is likely because affricates involve two articulatory targets and a complex articulatory coordination that require mature and precise speech motor control. Note that both CI subgroups, just like the NH controls, showed confusions in place and manner features but rarely in the aspiration feature. They did not use unaspirated sounds to substitute aspirated sounds or vice versa. For example, they used the unaspirated alveolar stop /t/ to substitute the unaspirated affricates /ts, tɕ, tʂ/ but not the aspirated sounds. They used the aspirated alveolar stop /tʰ/ to substitute the aspirated affricates /tsʰ, tɕʰ, tʂʰ/ but not the unaspirated sounds. Among the six stops, confusions occurred among /p, t, g/ and /pʰ, tʰ, kʰ/ but not across the aspiration feature. Among the six affricates, confusions occurred among /ts, tɕ, tʂ/ and /tsʰ, tɕʰ, tʂʰ/ but not across the aspiration feature. This observation indicates that the unaspirated–aspirated contrast is likely mastered well by the CI children.

Relationship Between Obstruent Intelligibility and Relevant Time Variables

The third research question addressed the relationship between the intelligibility of obstruent consonants and demographic features. Our NH children were aged between 3 and 10 years. According to previous studies on consonant acquisition in typically developing children (McLeod & Crowe, 2018; Smit et al., 1990), some critical changes in consonant acquisition occur in this age range. In this study, the regression analysis revealed a significant relationship between chronological age and the overall intelligibility of obstruent production in NH children. This finding suggests that consonant acquisition is a long-term process, which reflects the gradual development of speech motor control and articulatory coordination (Green et al., 2000; Smith & Zelaznik, 2004).

For children with CIs, age at implantation and length of CI use have been identified as two key factors in predicting CI children's speech-language abilities and perceptual performance (e.g., Connor et al., 2006; Gao et al., 2021; Kirk et al., 2002; Liu et al., 2013; Miyamoto et al., 1996; Nikolopoulos et al., 1999; Niparko et al., 2010; Svirsky et al., 2000). In Peng et al. (2004), which also targeted Mandarin-speaking children with CIs, the authors found a significant negative relationship between the children's age at implantation and their overall score of correct consonant production as well as a significant positive correlation between the children's length of CI use and consonant scores. In this study, the best fit regression model yielded significant factors of chronological age and age at implantation. One important difference between Peng et al.'s study and this study was that the CI children in Peng et al.'s study were all older than 6 years (average = 9.25). The children with CIs in this study had a younger average age, and a few were younger than 6 years. The significant effect of chronological age reflected the important role of this factor in the consonant acquisition of children with CIs, especially young children with CIs, as in NH children. It is noteworthy that significant effects of chronological age and age at implantation did not exclude the importance of the length of CI use. According to this model, children with a younger implantation age (e.g., < 3 years old) and an older chronological age tended to have higher consonant intelligibility. Because the factor of length of CI use was determined by chronological age and age at implantation, children with a younger implantation age and an older chronological age indicated that they had a longer duration of CI use. Also, because the factor of chronological age was highly correlated with the factor of length of CI use, the effect of chronological age included the contribution of the length of CI use. Another point shown in our data set was the nonlinear relationship between the time variables and consonant production in children with CIs. As shown in Figure 4, consonant intelligibility increased as chronological age increased before 10 years of age. After that, consonant intelligibility was at a plateau (approximately 60% correct), which did not show further improvement as chronological age increased. For the variable of length of CI use, the positive effect of this factor on the CI children's consonant intelligibility was lessened after 8 years of CI use. With regard to the variable of implantation age, the children who received implantation at a later age did not always show lower intelligibility than the children who received it at a younger age. However, the few top performers were the children who received cochlear implantation at the youngest age (see Figure 4). In general, the regression findings highlighted that the production performance of the CI children was not determined by one single factor in a linear way. Instead, it was the outcome of the combined effect of multiple factors, and these factors contributed to consonant development in a complex manner.

Other Points, Limitations, and Future Study

Finally yet importantly, this study provided insights into the methodology adoption in phonetic studies. Different from the commonly used approach of recruiting a few trained clinicians/listeners to conduct phonetic transcription and measure production accuracy, we recruited a large number of inexperienced listeners for a consonant identification task to evaluate consonant intelligibility through an online research platform. Instead of having all speech samples transcribed by the same one or two individuals, this study had every stimulus identified by multiple novel listeners. In the meantime, unlike other speech intelligibility studies that typically have novel listeners rated the overall intelligibility on a scale (e.g., Speech Intelligibility Rating) or use an open-set task in which listeners write down the words or sentences they hear from a speaker, this study adopted a closed-set task to examine the intelligibility of speech sounds. Many of our key findings conform to the outcomes reported in previous studies, especially those that also targeted the Mandarin-speaking CI population (e.g., Peng et al., 2004). The confusion matrices generated from the response data provided valuable information regarding the error patterns of speech production in both NH and CI children. The results of this study provided empirical evidence on the reliability and effectiveness of the approach of recruiting naïve listeners for a consonant identification task through an online research platform.

Although we highlighted the advantages of naïve listeners, this study does not advocate naïve listeners for all perception tasks. Some researchers found a significant effect of professional experience on listeners' responses in certain perception tasks (Munson et al., 2012; Wolfe et al., 2003). Furthermore, although the online research platform demonstrates superiority in ease of sampling, cost, flexibility, and so forth, compared to traditional behavioral studies conducted in research laboratories, it likely introduces additional distracting factors that may affect data quality (Peer et al., 2021). In this study, we randomly selected four child speakers and used their speech samples in all stimulus subtests for all listeners. This procedure helped us identify and exclude invalid data, but it increased the number of testing trials and the time cost of the listeners.

Compared to previous large-scale studies that included a greater number of child speakers, this study had a relatively small sample size. As the CI population shows substantial individual differences in outcome and benefit, it is necessary to recruit more children with CIs to improve the generalizability of the research findings. In this study, there was a strong correlation between the CI children's chronological age and length of device use, which caused data-based multicollinearity. Because of this issue, the two highly correlated factors were not tested in the same regression model. With an increased sample size and more diverse CI participants included, multicollinearity can be better controlled, and the relationship between consonant production and all three factors can be addressed. Furthermore, although we recruited 100 listeners for the consonant identification task, all of the participants were young adults. It is unknown whether there will be an age-related difference in the response data between younger and older NH listeners. For future studies, listeners of different ages should be recruited. Finally, although this perception study addressed the deficits in consonant production in Mandarin-speaking children with CIs, for future studies, an acoustic analysis should be conducted to identify the spectral and temporal features of consonants produced by the children with CIs. With the acoustic data showing the deviations of fine-grained speech characteristics from typically developing targets for individual sounds, researchers and clinicians can better understand the true deficits and difficulties of speech production in children with CIs and can design a more targeted plan for oral rehabilitation to improve their speech intelligibility.

Conclusions

In summary, our perceptual data revealed that the Mandarin-speaking children with CIs, regardless of subgroup (CA or HA), showed lower intelligibility for obstruent consonants with reference to the NH controls. Among the 17 tested obstruent consonants, the children with CIs showed higher intelligibility for stops but demonstrated major problems with the fricatives and affricates involved in the three-way place contrast. Among the three places of the sibilant contrast, the children with CIs showed the lowest intelligibility and the greatest difficulty with the alveolar sounds. The confusion pattern in the children with CIs was similar to the NH pattern in that they all showed the greatest confusion between the alveolar and retroflex sounds. However, compared with the NH controls, the children with CIs, including both the CA and HA subgroups, tended to use the retroflex sounds to replace the alveolar sounds and showed confusions between the alveolopalatal sounds and the other two places, which indicated the joint influence of auditory impairment and speech development on consonant production in children with CIs. Finally, the regression analysis revealed a strong positive relationship between the children's chronological age and the overall intelligibility in NH children. For the children with CIs, the best fit regression model revealed significant effects of chronological age and age at implantation, with their quadratic terms included. These results suggested that chronological age and the combined effect of the time variables played important roles in consonant development in the CI children.

Data Availability Statement

The response data from listeners are available upon request from the authors.

References

Adi-Bensaid, L., & Ben-David, A. (2010). Typical acquisition by atypical children: Initial consonant cluster acquisition by Israeli Hebrew–acquiring children with cochlear implants. *Clinical Linguistics & Phonetics, 24*(10), 771–794. https://doi.org/10.3109/02699206.2010.498932

Allison, K. M. (2020). Measuring speech intelligibility in children with motor speech disorders. *Perspectives of the ASHA Special Interest Groups, 5*(4), 809–820. https://doi.org/10.1044/2020_PERSP-19-00110

Asad, A. N., Purdy, S. C., Ballard, E., Fairgray, L., & Bowen, C. (2018). Phonological processes in the speech of school-age children with hearing loss: Comparisons with children with normal hearing. *Journal of Communication Disorders, 74,* 10–22. https://doi.org/10.1016/j.jcomdis.2018.04.004

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological), 57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Blamey, P. J., Barry, J. G., & Jacq, P. (2001). Phonetic inventory development in young cochlear implant users 6 years postoperation. *Journal of Speech, Language, and Hearing Research, 44*(1), 73–79. https://doi.org/10.1044/1092-4388(2001/007)

Boonen, N., Kloots, H., & Gillis, S. (2020). Is the spontaneous speech of 7-year-old cochlear implanted children as intelligible as that of their normally hearing peers? *International Journal of Pediatric Otorhinolaryngology, 133,* 109956. https://doi.org/10.1016/j.ijporl.2020.109956

Bouchard, M.-E. G., Le Normand, M.-T., & Cohen, H. (2007). Production of consonants by prelinguistically deaf children with cochlear implants. *Clinical Linguistics & Phonetics, 21*(11–12), 875–884. https://doi.org/10.1080/02699200701653634

Buhler, H. C., DeThomasis, B., Chute, P., & DeCora, A. (2007). An analysis of phonological process use in young children with cochlear implants. *The Volta Review, 107*(1), 55–74. https://doi.org/10.17955/tvr.107.1.569

Casserly, E. D. (2015). Effects of real-time cochlear implant simulation on speech production. *The Journal of the Acoustical Society of America, 137*(5), 2791–2800. https://doi.org/10.1121/1.4916965

Casserly, E. D., & Pisoni, D. B. (2010). Speech perception and production. *WIREs Cognitive Science, 1*(5), 629–647. https://doi.org/10.1002/wcs.63

Chen, X., You, Y., Yang, J., Qian, J., Lu, E., Kuehnel, V., Rehmann, J., Liu, B., & Xu, L. (2020). Effects of nonlinear frequency compression on Mandarin speech and sound-quality perception in hearing-aid users. *International Journal of Audiology, 59*(7), 524–533. https://doi.org/10.1080/14992027.2020.1761035

Chin, S. B. (2003). Children's consonant inventories after extended cochlear implant use. *Journal of Speech, Language, and Hearing Research, 46*(4), 849–862. https://doi.org/10.1044/1092-4388(2003/066)

Chin, S. B., Tsai, P. L., & Gao, S. (2003). Connected speech intelligibility of children with cochlear implants and children with normal hearing. *American Journal of Speech-Language Pathology, 12*(4), 440–451. https://doi.org/10.1044/1058-0360(2003/090)

Ching, T. Y., Dillon, H., & Byrne, D. (1998). Speech recognition of hearing-impaired listeners: Predictions from audibility and the limited role of high-frequency amplification. *The Journal of the Acoustical Society of America, 103*(2), 1128–1140. https://doi.org/10.1121/1.421224

Connor, C. M., Craig, H. K., Raudenbush, S. W., Heavner, K., & Zwolan, T. A. (2006). The age at which young deaf children receive cochlear implants and their vocabulary and speech-production growth: Is there an added value for early implantation? *Ear and Hearing, 27*(6), 628–644. https://doi.org/10.1097/01.aud.0000240640.59205.42

Dillon, C. M., Cleary, M., Pisoni, D. B., & Carter, A. K. (2004). Imitation of nonwords by hearing-impaired children with cochlear implants: Segmental analyses. *Clinical Linguistics & Phonetics, 18*(1), 39–55. https://doi.org/10.1080/0269920031000151669

Dillon, C. M., Pisoni, D. B., Cleary, M., & Carter, A. K. (2004). Nonword imitation by children with cochlear implants: Consonant analyses. *Archives of Otolaryngology—Head & Neck Surgery, 130*(5), 587–591. https://doi.org/10.1001/archotol.130.5.587

Dorman, M. F., Loizou, P. C., Fitzke, J., & Tu, Z. (1998). The recognition of sentences in noise by normal-hearing listeners using simulations of cochlear-implant signal processors with 6–20 channels. *The Journal of the Acoustical Society of America, 104*(6), 3583–3585. https://doi.org/10.1121/1.423940

Eriks-Brophy, A., Gibson, S., & Tucker, S.-K. (2013). Articulatory error patterns and phonological process use of preschool children with and without hearing loss. *The Volta Review, 113*(2), 87–125.

Ertmer, D. J., & Goffman, L. (2011). Speech production accuracy and variability in young cochlear implant recipients: Comparisons with typically developing age-peers. *Journal of Speech, Language, and Hearing Research, 54*(1), 177–189. https://doi.org/10.1044/1092-4388(2010/09-0165)

Ertmer, D. J., & Jung, J. (2012). Prelinguistic vocal development in young cochlear implant recipients and typically developing infants: Year 1 of robust hearing experience. *Journal of Deaf Studies and Deaf Education, 17*(1), 116–132. https://doi.org/10.1093/deafed/enr021

Faes, J., & Gillis, S. (2016). Word initial fricative production in children with cochlear implants and their normally hearing peers matched on lexicon size. *Clinical Linguistics & Phonetics, 30*(12), 959–982. https://doi.org/10.1080/02699206.2016.1213882

Faes, J., & Gillis, S. (2018). Intraword variability in children with cochlear implants: The long-term development up to 5 years of age and a comparison with children with normal hearing. *American Journal of Speech-Language Pathology, 27*(2), 706–720. https://doi.org/10.1044/2017_AJSLP-16-0130

Freeman, V., Pisoni, D. B., Kronenberger, W. G., & Castellanos, I. (2017). Speech intelligibility and psychosocial functioning in deaf children and teens with cochlear implants. *Journal of Deaf Studies and Deaf Education, 22*(3), 278–289. https://doi.org/10.1093/deafed/enx001

Gao, Q., Wong, L. L. N., & Chen, F. (2021). A review of speech perception of Mandarin-speaking children with cochlear implantation. *Frontiers in Neuroscience, 15,* 773694. https://doi.org/10.3389/fnins.2021.773694

Geers, A. E., Moog, J. S., Biedenstein, J., Brenner, C., & Hayes, H. (2009). Spoken language scores of children using cochlear implants compared to hearing age-mates at school entry. *Journal of Deaf Studies and Deaf Education, 14*(3), 371–385. https://doi.org/10.1093/deafed/enn046

Glickman, M. E., Rao, S. R., & Schultz, M. R. (2014). False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. *Journal of Clinical Epidemiology, 67*(8), 850–857. https://doi.org/10.1016/j.jclinepi.2014.03.012

Grandon, B., Martinez, M. J., Samson, A., & Vilain, A. (2020). Long-term effects of cochlear implantation on the intelligibility of speech in French-speaking children. *Journal of Child Language, 47*(4), 881–892. https://doi.org/10.1017/S0305000919000837

Grandon, B., & Vilain, A. (2020). Development of fricative production in French-speaking school-aged children using cochlear implants and children with normal hearing. *Journal of Communication Disorders, 86,* 105996. https://doi.org/10.1016/j.jcomdis.2020.105996

Green, J. R., Moore, C. A., Higashikawa, M., & Steeve, R. W. (2000). The physiologic development of speech motor control: Lip and jaw coordination. *Journal of Speech, Language, and Hearing Research, 43*(1), 239–255. https://doi.org/10.1044/jslhr.4301.239

Hinojosa, R., & Lindsay, J. R. (1980). Profound deafness: Associated sensory and neural degeneration. *Archives of Otolaryngology, 106*(4), 193–209. https://doi.org/10.1001/archotol.1980.00790280001001

Hua, Z., & Dodd, B. (2000). The phonological acquisition of Putonghua (Modern Standard Chinese). *Journal of Child Language, 27*(1), 3–42. https://doi.org/10.1017/S030500099900402X

Huang, C.-Y., Yang, H.-M., Sher, Y.-J., Lin, Y.-H., & Wu, J.-L. (2005). Speech intelligibility of Mandarin-speaking deaf children with cochlear implants. *International Journal of Pediatric Otorhinolaryngology, 69*(4), 505–511. https://doi.org/10.1016/j.ijporl.2004.10.017

Iyer, S. N., Jung, J., & Ertmer, D. J. (2017). Consonant acquisition in young cochlear implant recipients and their typically developing peers. *American Journal of Speech-Language Pathology, 26*(2), 413–427. https://doi.org/10.1044/2016_AJSLP-16-0073

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America, 108*(3), 1252–1263. https://doi.org/10.1121/1.1288413

Kirk, K. I., Diefendorf, E., Riley, A., & Osberger, M. J. (1995). Consonant production by children with multichannel cochlear implants or hearing aids. *Advances in Otorhinolaryngology, 50,* 154–159. https://doi.org/10.1159/000424452

Kirk, K. I., Ying, E., Miyamoto, R. T., O'Neill, T., Lento, C. L., & Fears, B. (2002). Effects of age at implantation in young children. *Annals of Otology, Rhinology & Laryngology, 111*(Suppl. 5), 69–73. https://doi.org/10.1177/00034894021110S515

Ladefoged, P. (2001). *A course in phonetics* (4th ed.). Harcourt.

Ladefoged, P., & Wu, Z. (1984). Places of articulation: An investigation of Pekingese fricatives and affricates. *Journal of Phonetics, 12*(3), 267–278. https://doi.org/10.1016/S0095-4470(19)30883-6

Lagerberg, T. B., Åsberg, J., Hartelius, L., & Persson, C. (2014). Assessment of intelligibility using children's spontaneous speech: Methodological aspects. *International Journal of Language & Communication Disorders, 49*(2), 228–239. https://doi.org/10.1111/1460-6984.12067

Lee, C.-Y., Zhang, Y., & Li, X. (2014). Acoustic characteristics of voiceless fricatives in Mandarin Chinese. *Journal of Chinese Linguistics, 42*(1), 150–171.

Lee, S. I. (2011). Spectral analysis of Mandarin Chinese sibilant fricatives. In W. S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 1178–1181). City University of Hong Kong.

Leigh, J., Dettman, S., Dowell, R., & Briggs, R. (2013). Communication development in children who receive a cochlear implant by 12 months of age. *Otology & Neurotology, 34*(3), 443–450. https://doi.org/10.1097/MAO.0b013e3182814d2c

Li, F. (2008). *The phonetic development of voiceless sibilant fricatives in English, Japanese and Mandarin Chinese* [Doctoral dissertation, The Ohio State University]. OhioLINK Electronic Theses and Dissertations Center.

Li, F., & Munson, B. (2016). The development of voiceless sibilant fricatives in Putonghua-speaking children. *Journal of Speech, Language, and Hearing Research, 59*(4), 699–712. https://doi.org/10.1044/2016_JSLHR-S-14-0142

Liker, M., Mildner, V., & Šindija, B. (2007). Acoustic analysis of the speech of children with cochlear implants: A longitudinal study. *Clinical Linguistics & Phonetics, 21*(1), 1–11. https://doi.org/10.1080/02699200400026991

Lin, T., & Wang, L. (2001). *A course in phonetics*. Peking University Press.

Lin, Y.-S., & Peng, S.-C. (2003). Acquisition profiles of syllable-initial consonants in Mandarin-speaking children with cochlear implants. *Acta Oto-Laryngologica, 123*(9), 1046–1053. https://doi.org/10.1080/00016480410016171

Lindblom, B., & Maddieson, I. (1988). Phonetic universals in consonant systems. In L. H. Hyman & C. N. Li (Eds.), *Language, speech and mind* (pp. 62–78). Routledge.

Liu, Q., Zhou, N., Berger, B., Huang, D., & Xu, L. (2013). Mandarin consonant contrast recognition among children with cochlear implants or hearing aids and normal-hearing children. *Otology & Neurotology, 34*(3), 471–476. https://doi.org/10.1097/MAO.0b013e318286836b

Locke, J. L. (1983). *Phonological acquisition and change*. Academic Press.

Loizou, P. C. (1999). Introduction to cochlear implants. *IEEE Engineering in Medicine and Biology Magazine, 18*(1), 32–42. https://doi.org/10.1109/51.740962

Loizou, P. C. (2006). Speech processing in vocoder-centric cochlear implants. *Advances in Otorhinolaryngology, 64,* 109–143. https://doi.org/10.1159/000094648

Mao, Y., Chen, H., Xie, S., & Xu, L. (2020). Acoustic assessment of tone production of prelingually-deafened Mandarin-speaking children with cochlear implants. *Frontiers in Neuroscience, 14,* 592954. https://doi.org/10.3389/fnins.2020.592954

McAllister Byun, T., Halpin, P. F., & Szeredi, D. (2015). Online crowdsourcing for efficient rating of speech: A validation study. *Journal of Communication Disorders, 53,* 70–83. https://doi.org/10.1016/j.jcomdis.2014.11.003

McCaffrey, H. A., Davis, B. L., MacNeilage, P. F., & von Hapsburg, D. (1999). Multichannel cochlear implantation and the organization of early speech. *The Volta Review, 101*(1), 5–28.

McLeod, S., & Crowe, K. (2018). Children's consonant acquisition in 27 languages: A cross-linguistic review. *American Journal of Speech-Language Pathology, 27*(4), 1546–1571. https://doi.org/10.1044/2018_AJSLP-17-0100

Mildner, V., & Liker, M. (2008). Fricatives, affricates, and vowels in Croatian children with cochlear implants. *Clinical Linguistics & Phonetics, 22*(10–11), 845–856. https://doi.org/10.1080/02699200802130557

Miller, N. (2013). Measuring up to speech intelligibility. *International Journal of Language & Communication Disorders, 48*(6), 601–612. https://doi.org/10.1111/1460-6984.12061

Miyamoto, R. T., Kirk, K. I., Robbins, A. M., Todd, S., & Riley, A. (1996). Speech perception and speech production skills of children with multichannel cochlear implants. *Acta Oto-Laryngologica, 116*(2), 240–243. https://doi.org/10.3109/00016489609137832

Moeller, M. P., & Tomblin, J. B. (2015). An introduction to the Outcomes of Children with Hearing Loss study. *Ear and Hearing, 36*(Suppl. 1), 4S–13S. https://doi.org/10.1097/AUD.0000000000000210

Moore, B. C. J. (2016). A review of the perceptual effects of hearing loss for frequencies above 3 kHz. *International Journal of Audiology, 55*(12), 707–714. https://doi.org/10.1080/14992027.2016.1204565

Moore, J. A., Prath, S., & Arrieta, A. (2006). Early Spanish speech acquisition following cochlear implantation. *The Volta Review, 106*(3), 321–341. https://doi.org/10.17955/tvr.106.3.m.578

Munson, B., Donaldson, G. S., Allen, S. L., Collison, E. A., & Nelson, D. A. (2003). Patterns of phoneme perception errors by listeners with cochlear implants as a function of overall speech perception ability. *The Journal of the Acoustical Society of America, 113*(2), 925–935. https://doi.org/10.1121/1.1536630

Munson, B., Johnson, J. M., & Edwards, J. (2012). The role of experience in the perception of phonetic detail in children's speech: A comparison between speech-language pathologists and clinically untrained listeners. *Journal of Speech, Language, and Hearing Research, 21*(2), 124–139. https://doi.org/10.1044/1058-0360(2011/11-0009)

Nicholas, J. G., & Geers, A. E. (2006). Effects of early auditory experience on the spoken language of deaf children at 3 years of age. *Ear and Hearing, 27*(3), 286–298. https://doi.org/10.1097/01.aud.0000215973.76912.c6

Nikolopoulos, T. P., O'Donoghue, G. M., & Archbold, S. (1999). Age at implantation: Its importance in pediatric cochlear implantation. *The Laryngoscope, 109*(4), 595–599. https://doi.org/10.1097/00005537-199904000-00014

Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N.-Y., Quittner, A. L., & Fink, N. E. (2010). Spoken language development in children following cochlear implantation. *JAMA, 303*(15), 1498–1506. https://doi.org/10.1001/jama.2010.451

Osberger, M. J. (1992). Speech intelligibility in the hearing impaired: Research and clinical implications. In R. D. Kent (Ed.), *Speech intelligibility in speech disorders: Theory, measurement and management* (pp. 233–264). John Benjamins. https://doi.org/10.1075/sspcl.1.07osb

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2021). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods, 54*(4), 1643–1662. https://doi.org/10.3758/s13428-021-01694-3

Peng, S.-C., Weiss, A. L., Cheung, H., & Lin, Y.-S. (2004). Consonant production and language skills in Mandarin-speaking children with cochlear implants. *Archives of Otolaryngology—Head & Neck Surgery, 130*(5), 592–597. https://doi.org/10.1001/archotol.130.5.592

Poursoroush, S., Ghorbani, A., Soleymani, Z., Kamali, M., Yousefi, N., & Poursoroush, Z. (2015). Speech intelligibility of cochlear-implanted and normal-hearing children. *Iranian Journal of Otorhinolaryngology, 27*(82), 361–367.

Qi, S., Chen, X., Yang, J., Wang, X., Tian, X., Huang, H., Rehmann, J., Kuehnel, V., Guan, J., & Xu, L. (2021). Effects of adaptive non-linear frequency compression in hearing aids on Mandarin speech and sound-quality perception. *Frontiers in Neuroscience, 15,* 722970. https://doi.org/10.3389/fnins.2021.722970

Reidy, P. F., Kristensen, K., Winn, M. B., Litovsky, R. Y., & Edwards, J. R. (2017). The acoustics of word-initial fricatives and their effect on word-level intelligibility in children with bilateral cochlear implants. *Ear and Hearing, 38*(1), 42–56. https://doi.org/10.1097/AUD.0000000000000349

Sander, E. K. (1972). When are speech sounds learned? *Journal of Speech and Hearing Disorders, 37*(1), 55–63. https://doi.org/10.1044/jshd.3701.55

Schramm, B., Bohnert, A., & Keilmann, A. (2010). Auditory, speech and language development in young children with cochlear implants compared with children with normal hearing. *International Journal of Pediatric Otorhinolaryngology, 74*(7), 812–819. https://doi.org/10.1016/j.ijporl.2010.04.008

Serry, T. A., & Blamey, P. J. (1999). A 4-year investigation into phonetic inventory development in young cochlear implant users. *Journal of Speech, Language, and Hearing Research, 42*(1), 141–154. https://doi.org/10.1044/jslhr.4201.141

Shriberg, L. D., & Kent, R. D. (2003). *Clinical phonetics* (3rd ed.). Pearson.

Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A. (1990). The Iowa Articulation Norms Project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55*(4), 779–798. https://doi.org/10.1044/jshd.5504.779

Smith, A., & Zelaznik, H. N. (2004). Development of functional synergies for speech motor coordination in childhood and adolescence. *Developmental Psychobiology, 45*(1), 22–33. https://doi.org/10.1002/dev.20009

Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., Lewis, D. E., & Moeller, M. P. (2004). The importance of high-frequency audibility in the speech and language development of children with hearing loss. *Archives of Otolaryngology—Head & Neck Surgery, 130*(5), 556–562. https://doi.org/10.1001/archotol.130.5.556

Summerfield, A. Q., Nakisa, M. J., McCormick, B., Archbold, S., Gibbin, K. P., & O'Donoghue, G. M. (2002). Use of vocalic information in the identification of /s/ and /ʃ/ by children with cochlear implants. *Ear and Hearing, 23*(1), 58–77. https://doi.org/10.1097/00003446-200202000-00005

Sundarrajan, M., Tobey, E. A., Nicholas, J., & Geers, A. E. (2020). Assessing consonant production in children with cochlear implants. *Journal of Communication Disorders, 84,* 105966. https://doi.org/10.1016/j.jcomdis.2019.105966

Svirsky, M. A., Robbins, A. M., Kirk, K. I., Pisoni, D. B., & Miyamoto, R. T. (2000). Language development in profoundly deaf children with cochlear implants. *Psychological Science, 11*(2), 153–158. https://doi.org/10.1111/1467-9280.00231

Templin, M. C. (1957). *Certain language skills in children: Their development and interrelationships*. University of Minnesota Press. https://doi.org/10.5749/j.ctttv2st

Tobey, E. A., Geers, A. E., Brenner, C., Altuna, D., & Gabbert, G. (2003). Factors associated with development of speech production skills in children implanted by age five. *Ear and Hearing, 24*(1), 36S–45S. https://doi.org/10.1097/01.AUD.0000051688.48224.A6

Tobey, E. A., Geers, A. E., Sundarrajan, M., & Lane, J. (2011). Factors influencing elementary and high-school aged cochlear implant users. *Ear and Hearing, 32*(1), 27S–38S. https://doi.org/10.1097/AUD.0b013e3181fa41bb

Tobey, E. A., Pancamo, S., Staller, S. J., Brimacornbe, J. A., & Beiter, A. L. (1991). Consonant production in children receiving a multichannel cochlear implant. *Ear and Hearing, 12*(1), 23–31. https://doi.org/10.1097/00003446-199102000-00003

Todd, A. E., Edwards, J. R., & Litovsky, R. Y. (2011). Production of contrast between sibilant fricatives by children with cochlear implants. *The Journal of the Acoustical Society of America, 130*(6), 3969–3979. https://doi.org/10.1121/1.3652852

Tomblin, J. B., Peng, S.-C., Spencer, L. J., & Lu, N. (2008). Long-term trajectories of the development of speech sound production in pediatric cochlear implant recipients. *Journal of Speech, Language, and Hearing Research, 51*(5), 1353–1368. https://doi.org/10.1044/1092-4388(2008/07-0083)

Tye-Murray, N., Spencer, L., & Woodworth, G. G. (1995). Acquisition of speech by children who have prolonged cochlear implant experience. *Journal of Speech and Hearing Research, 38*(2), 327–337. https://doi.org/10.1044/jshr.3802.327

Vihman, M. M. (1996). *Phonological development: The origins of language in the child*. Blackwell.

Ward, J. H., Jr. (1963). Hierarchical grouping to optimize an objective function. *The Journal of the American Statistical Association, 58*(301), 236–244. https://doi.org/10.1080/01621459.1963.10500845

Warner-Czyz, A., & Davis, B. (2008). The emergence of segmental accuracy in young cochlear implant recipients. *Cochlear Implants International, 9*(3), 143–166. https://doi.org/10.1179/cim.2008.9.3.143

Wie, O. B., Falkenberg, E. S., Tvete, O., & Tomblin, B. (2007). Children with a cochlear implant: Characteristics and determinants of speech recognition, speech-recognition growth rate, and speech production. *International Journal of Audiology, 46*(5), 232–243. https://doi.org/10.1080/14992020601182891

Wolfe, V., Martin, D., Borton, T., & Youngblood, H. C. (2003). The effect of clinical experience on cue trading for the /r-w/ contrast. *American Journal of Speech-Language Pathology, 12*(2), 221–228. https://doi.org/10.1044/1058-0360(2003/068)

Yang, J., Qian, J., Chen, X., Kuehnel, V., Rehmann, J., von Buol, A., Li, Y., Ren, C., Liu, B., & Xu, L. (2018). Effects of nonlinear frequency compression on the acoustic properties and recognition of speech sounds in Mandarin Chinese. *The Journal of the Acoustical Society of America, 143*(3), 1578–1590. https://doi.org/10.1121/1.5027404

Yang, J., Vadlamudi, J., Yin, Z., Lee, C.-Y., & Xu, L. (2017). Production of word-initial fricatives of Mandarin Chinese in prelingually deafened children with cochlear implants. *International Journal of Speech-Language Pathology, 19*(2), 153–164. https://doi.org/10.3109/17549507.2016.1143972

Yorkston, K. M., Strand, E. A., & Kennedy, M. R. T. (1996). Comprehensibility of dysarthric speech: Implications for assessment and treatment planning. *American Journal of Speech-Language Pathology, 5*(1), 55–66. https://doi.org/10.1044/1058-0360.0501.55

Zeng, F.-G., & Turner, C. W. (1990). Recognition of voiceless fricatives by normal and hearing-impaired subjects. *Journal of Speech and Hearing Research, 33*(3), 440–449. https://doi.org/10.1044/jshr.3303.440

Zhou, N., & Xu, L. (2008). Development and evaluation of methods for assessing tone production skills in Mandarin-speaking children with cochlear implants. *The Journal of the Acoustical Society of America, 123*(3), 1653–1664. https://doi.org/10.1121/1.2832623

Zimmermann, C. E., Burgess, B. J., & Nadol, J. B., Jr. (1995). Patterns of degeneration in the human cochlear nerve. *Hearing Research, 90*(1–2), 192–201. https://doi.org/10.1016/0378-5955(95)00165-1

Appendix A

Word List Used for Speech Sample Collection Through an Audiovisual Word Repetition Task

Consonant IPA	Pinyin	English translation	Consonant IPA	Pinyin	English translation	Consonant IPA	Pinyin	English translation
p	bā zì	character eight	f	fā shāo	fever	ts	zá jì	acrobatics
p	bí zi	nose	f	fú zì	character fu	ts	zú qiú	soccer
p	bù wá wa	doll	f	fēi jī	airplane	ts	zì lái shǔi	tap water
p^h	pā xià	lie face down	x	hā mì guā	cantaloupe	ts^h	cā dì	wipe the floor
p^h	pí qiú	ball	x	hú dié	butterfly	ts^h	cù píng	vinegar bottle
p^h	pú táo	grape	x	hēi sè	black	ts^h	cì wèi	hedgehog
t	dā jī mù	play with blocks	s	să shǔi chē	sprinkler truck	tʂ	zhá jiàng miàn	Zhajiang noodle
t	dī tóu	lower head	s	sù liào dài	plastic bag	tʂ	zhū bā jiè	Zhu Bajie (Pigsy)
t	dú shū	read a book	s	sī jīn	silk scarf	tʂ	zhī máo yī	knit a sweater
t^h	tà bǎn	pedal	ʂ	shā fā	sofa	tʂ^h	chā zuò	outlet
t^h	tī zi	ladder	ʂ	shū zhuō	desk	tʂ^h	chú fáng	kitchen
t^h	tù zi	rabbit	ʂ	shī zi	lion	tʂ^h	chī fàn	have a meal
k	gā li	curry	ɕ	xiā zi	shrimp	tɕ	jiā fēi māo	Garfield
k	gē zi	pigeon	ɕ	xuē zi	boots	tɕ	jú zi	orange
k	gū niáng	girls	ɕ	xī guā	watermelon	tɕ	jī dàn	egg
k^h	kā fēi	coffee				tɕ^h	qiā shǒu	pinch the arm
k^h	kè běn	textbook				tɕ^h	qū qū	cricket
k^h	kū bí zi	cry				tɕ^h	qī zì	character seven

Note. IPA = International Phonetic Alphabet.

Appendix B (p. 1 of 3)

Confusion Matrix Data

Table B1. Confusion matrix of children with normal hearing.

Consonant	p	p^h	f	t	t^h	s	ts	ts^h	ɕ	tɕ	$tɕ^h$	ʂ	tʂ	$tʂ^h$	k	k^h	x	None
p	84.6	4.4	0.7	5.3	0.0	0.2	0.0	0.0	0.0	0.4	0.0	0.0	0.2	0.2	0.6	0.0	0.4	3.1
p^h	2.1	82.1	0.1	0.2	6.0	0.1	0.0	1.4	0.0	0.2	1.3	0.2	0.0	0.3	0.2	2.7	0.8	2.2
f	1.8	3.4	86.8	0.2	0.0	0.5	1.4	0.1	0.0	0.2	0.0	0.8	0.0	0.3	0.2	0.2	1.5	2.5
t	7.1	0.5	0.2	76.8	1.8	0.0	0.8	0.0	0.0	3.4	0.0	0.0	1.1	0.0	4.6	0.2	0.4	3.3
t^h	0.0	11.1	0.0	0.0	73.4	0.2	0.5	0.5	0.5	0.5	5.5	0.4	0.0	1.0	0.1	2.7	1.9	1.7
s	0.3	0.2	3.6	0.4	0.5	60.1	5.3	4.8	2.6	0.2	1.1	10.1	2.2	1.8	0.0	0.2	0.6	5.8
ts	0.8	0.4	3.5	6.4	1.2	0.8	70.4	1.4	0.0	1.1	0.0	0.0	9.3	0.6	0.1	0.2	0.0	3.9
ts^h	0.2	1.7	0.6	0.1	8.9	5.1	1.8	65.2	0.3	0.2	0.6	0.5	0.8	8.1	0.0	2.0	0.5	3.3
ɕ	0.0	1.2	0.0	0.0	1.0	0.7	0.9	0.1	81.2	1.7	6.1	2.3	0.6	0.3	0.5	0.0	0.0	3.2
tɕ	0.0	1.4	0.0	3.2	1.5	0.3	1.2	0.0	3.1	78.3	1.1	0.2	2.7	0.4	3.3	1.4	0.0	1.9
$tɕ^h$	0.0	0.4	0.2	0.0	5.0	0.0	0.7	1.1	9.0	5.3	68.3	0.4	0.5	4.8	0.2	0.2	0.2	3.7
ʂ	0.0	0.7	2.7	0.2	0.0	18.5	0.6	2.3	3.8	0.0	1.4	61.8	1.1	4.1	0.0	0.0	0.2	2.6
tʂ	0.3	0.6	1.1	7.0	2.8	0.1	19.9	1.4	0.0	5.0	0.2	1.5	49.1	3.2	1.1	0.3	0.0	6.3
$tʂ^h$	0.0	2.2	2.6	0.2	7.7	0.8	1.1	13.6	0.5	0.2	1.2	5.5	2.8	57.3	0.0	0.6	0.2	3.4
k	3.1	1.0	0.2	6.5	0.5	0.0	0.6	0.0	0.0	0.2	0.0	0.0	3.0	0.2	81.3	0.7	0.0	2.8
k^h	0.0	7.9	0.1	0.0	9.3	0.0	0.2	2.8	0.2	0.0	0.3	0.2	0.0	2.9	1.8	70.8	1.8	1.7
x	1.0	20.1	2.4	0.0	1.1	0.0	0.0	0.3	0.4	0.0	0.4	0.4	0.2	0.0	0.4	3.2	64.0	6.2

Appendix B (p. 2 of 3)

Confusion Matrix Data

Table B2. Confusion matrix of the chronological age–matched cochlear implant subgroup.

Consonant	p	p^h	f	t	t^h	s	ts	ts^h	ɕ	tɕ	$tɕ^h$	ʂ	tʂ	$tʂ^h$	k	k^h	x	None
p	77.0	4.1	1.2	5.9	0.0	0.0	0.2	0.4	0.2	0.9	0.1	0.3	0.0	0.5	2.4	0.4	1.2	5.3
p^h	6.0	70.7	1.7	0.0	7.3	0.3	0.7	0.6	0.4	0.2	1.4	0.1	0.4	0.5	1.4	2.3	2.7	3.2
f	15.0	3.7	56.8	2.1	0.6	0.4	2.0	1.5	0.6	1.2	0.3	0.9	0.6	1.2	1.6	0.6	3.1	7.9
t	12.1	2.4	0.5	66.5	0.5	0.1	0.6	0.1	0.3	1.8	0.1	0.3	1.2	0.4	5.7	0.5	1.2	5.5
t^h	0.0	11.5	0.3	1.9	61.7	0.1	1.0	1.5	1.7	1.2	3.8	1.5	0.1	2.7	0.3	4.0	1.5	5.1
s	0.4	1.8	7.6	1.6	1.2	25.4	4.5	3.9	9.4	1.1	2.6	16.0	3.0	5.2	1.9	0.7	2.3	11.4
ts	1.0	1.1	1.7	17.2	1.7	2.6	23.3	2.0	3.0	9.1	3.6	1.3	18.8	1.1	0.8	1.8	1.3	8.5
ts^h	0.7	1.9	2.4	1.5	16.6	6.5	1.2	24.4	2.0	1.0	3.6	3.3	1.9	20.8	1.1	3.3	2.7	5.1
ɕ	0.4	1.2	0.5	1.0	1.0	1.5	1.1	0.6	61.6	9.2	3.2	5.9	1.5	0.7	0.3	1.3	2.0	7.1
tɕ	0.7	0.6	0.3	6.7	1.2	0.3	1.4	0.3	1.1	62.9	1.3	0.4	7.2	0.7	4.5	0.1	2.6	7.7
$tɕ^h$	0.3	1.6	0.0	0.4	8.9	0.6	0.9	2.0	12.2	6.4	44.1	4.6	1.7	7.1	1.4	0.8	2.1	4.9
ʂ	0.9	2.4	4.4	0.8	0.2	8.3	1.1	1.6	15.8	1.8	1.3	44.3	0.9	6.2	0.3	0.7	3.0	5.9
tʂ	0.4	1.4	0.0	10.3	3.0	0.3	5.7	0.6	1.1	12.6	1.8	1.1	41.8	3.1	2.7	0.8	3.9	9.4
$tʂ^h$	0.1	2.4	0.1	0.2	18.4	1.1	0.4	8.7	1.5	0.8	7.4	4.9	2.0	38.5	0.3	4.2	3.0	5.9
k	5.3	1.5	0.2	5.8	0.4	0.3	0.6	0.3	0.0	1.1	0.6	0.0	1.5	0.3	72.5	1.7	1.1	6.9
k^h	0.6	5.7	0.7	0.3	10.4	0.1	0.3	1.9	0.5	0.5	0.4	0.3	0.0	1.6	1.8	61.5	8.2	5.1
x	2.8	11.8	3.3	0.0	5.1	0.6	0.6	1.4	0.6	0.7	0.2	1.1	0.1	1.4	1.3	4.2	56.3	8.3

Appendix B (p. 3 of 3)

Confusion Matrix Data

Table B3. Confusion matrix of the hearing age–matched cochlear implant subgroup.

Consonant	p	p^h	f	t	t^h	s	ts	ts^h	ɕ	tɕ	$tɕ^h$	ʂ	tʂ	$tʂ^h$	k	k^h	x	None
p	81.9	2.1	0.8	4.9	0.2	0.4	0.1	0.2	0.1	0.8	0.1	0.6	0.0	0.4	1.2	0.1	1.4	4.5
p^h	4.8	74.8	1.5	0.2	5.8	0.4	0.6	0.4	0.3	0.5	2.1	0.1	0.2	0.1	0.4	1.1	3.1	3.5
f	13.4	2.1	66.7	2.0	0.8	0.4	2.0	0.6	0.6	1.0	0.4	0.9	0.9	0.3	0.4	0.9	1.5	5.3
t	7.0	1.8	0.5	75.0	0.3	0.2	1.0	0.4	0.5	2.0	0.4	0.3	1.0	0.4	4.0	0.5	0.9	3.8
t^h	0.0	11.9	0.3	0.9	62.3	0.1	1.2	1.2	2.4	1.3	4.7	0.6	0.4	2.8	0.3	3.7	1.8	4.0
s	1.0	1.3	7.5	1.9	0.9	32.4	6.4	4.5	7.3	0.6	1.4	17.3	1.7	3.9	0.0	0.6	1.5	9.8
ts	2.0	1.5	1.9	15.2	1.7	4.1	27.0	1.8	2.6	6.0	1.9	1.9	20.2	0.6	2.8	0.6	0.5	7.6
ts^h	0.8	2.3	1.4	0.2	16.6	7.7	2.0	29.2	1.7	0.6	2.2	3.9	1.4	19.6	0.0	3.3	2.2	4.9
ɕ	0.3	0.4	0.1	0.7	0.8	1.2	1.2	0.8	71.5	6.2	2.9	5.2	1.1	0.4	0.5	0.6	1.1	5.1
tɕ	0.5	0.6	0.3	5.9	2.0	0.0	1.8	0.3	1.6	68.0	1.4	0.6	5.0	0.5	3.6	0.4	1.5	6.1
$tɕ^h$	0.2	1.4	0.2	0.1	8.8	0.5	0.9	2.2	12.8	6.2	46.7	2.0	1.3	9.4	0.2	1.1	0.8	5.0
ʂ	0.8	1.6	3.8	0.0	0.1	6.4	1.4	1.3	12.9	1.1	0.9	58.3	0.1	4.9	0.4	0.7	2.0	3.4
tʂ	0.5	1.7	0.1	10.2	1.6	0.5	6.3	0.6	0.8	11.4	0.6	1.0	50.1	3.0	2.5	1.2	2.1	5.9
$tʂ^h$	0.4	2.3	0.5	0.1	14.4	0.5	0.3	8.5	1.0	1.0	4.7	4.8	1.1	48.9	0.3	4.1	2.3	4.7
k	4.2	0.7	0.4	4.7	0.6	0.4	0.4	0.2	0.0	0.8	0.3	0.0	1.1	0.4	74.6	1.5	1.0	8.8
k^h	0.5	6.3	1.0	0.1	8.0	0.4	0.2	1.8	0.7	0.5	0.4	0.3	0.0	1.3	1.1	66.3	7.8	3.1
x	1.6	13.7	2.8	0.0	2.8	0.5	0.8	0.6	0.5	0.3	0.4	0.5	0.2	0.9	0.9	3.7	61.8	8.0