Acoustic characteristics of stop consonants in Mandarin-speaking children with cochlear implants

Jing Yang[1], Li Xu[2]
[1]Communication Sciences and Disorders, University of Wisconsin-Milwaukee, USA
[2]Hearing, Speech and Language Sciences, Ohio University, USA

ABSTRACT

The present study examined the acoustic features of stop consonants in Mandarin-speaking children with cochlear implants (CIs). The speakers included 22 children with normal hearing (NH) and 35 children with CIs regrouped into chronological age-matched (CA) and hearing age-matched (HA) subgroups. The speech material was Mandarin disyllabic words including six Mandarin stops (/p, t, k, p^h, t^h, k^h/) followed by different vowels. The VOT, ΔVOT, normalized amplitude, and F2 onset were measured and compared between the NH children and each of the CI subgroups. The results revealed that the children with CIs, regardless of CA or HA, showed no difference from the NH controls on the VOT and ΔVOT. However, both CI subgroups produced lower amplitude for aspirated stops than the NH controls. Both CI subgroups followed the NH pattern of F2 onset varying as a function of place of articulation. In sum, Mandarin-speaking children with CIs have developed NH-like temporal and spectral features for stops but their respiratory control for stop production has not fully matured.

Keywords: children, cochlear implant, stop consonants, acoustic features, Mandarin Chinese.

1. INTRODUCTION

Phonological studies on normally developing children suggest that stops emerged very early and most children produce acceptable stops by three years of age [1-3]. However, detailed investigation of stop acoustics in normally developing children reveals that the refinement of articulatory and laryngeal sequences for stop consonants is a gradual process, which shows a continuing development after three years of age [4-6]. For the hearing-impaired population especially children with severe-to-profound hearing loss since birth, accurately producing speech segments is challenging. Regarding the production of stops, deaf children show deviation from the phonological system rather than individual, isolated speech errors in stop productions [7]. Congenital deaf speakers show highly variable and less contrastable voicing features for English stops [8, 9]. Meanwhile, they tend to use alternative features to produce phonological/phonetic contrasts [10, 11].

With the facilitation of auditory prosthesis, children with CIs show an acquisition order for consonant and vowel inventories comparable to NH children [12]. Among different types of consonants, stops are one of the most frequent and predominant manner of articulation in children with CIs [13, 14]. Acoustically, the change in hearing status introduced by CIs exerts an observable influence on the acoustic-phonetic characteristics of stops in CI users. Some researchers reported enlarged voiced-voiceless contrast represented by the VOT change post-implantation [15], adjusted VOT values in accordance with the change of syllable duration [16], increased voicing lead for voiced stops [17], and the NH-like production of VOT in most pediatric CI users [18]. Some researchers found that although early emerged and frequently used, the speech behaviors including articulatory and phonatory gestures of stop consonants in children with CIs still deviate from NH children after a certain period of device use [19, 20].

Extending from the previous studies that focused on the detailed speech characteristics of speech productions in English-speaking children with CIs [18, 21], The present study aims to characterize the acoustic profile of stop consonants in Mandarin-speaking children with CIs. Mandarin stops are produced at bilabial, alveolar, and velar places, similar to English stops. But Mandarin stops are characterized by aspirated-unaspirated distinction, different from the voiced-voiceless contrast in English. Despite the distinctive phonological labelling of the voicing feature, the phonetic representations of word-initial stops in these two languages are both reflected in short-lag vs. long-lag contrast along the VOT continuum. Given the phonological differences and phonetic similarities between Mandarin and English stop systems, we wondered whether the prelingually deafened Mandarin-speaking children with CIs, after a certain period of device use, develop NH-like acoustic features in stop production.

2. METHODS

2.1. Participants

The speakers included 57 Mandarin-speaking children (22 children with NH and 35 children with

CIs) recruited from the Beijing area, China. The children with NH (13 girls and 9 boys) were all aged between 3.25 and 10 years old (M = 6.17 yrs, SD = 1.69 yrs). None of them was reported as having any speech, language, or developmental issues. The children with CIs (16 girls and 19 boys) aged between 3.77 and 15 years old. They were prelingually deafened and all received unilateral implantation at ages between 1.08 and 9.42 years with the length of device use ranging between 0.08 and 9.49 years. To match with the children with NH, the children with CIs were assigned to chronological age-matched (CA) and hearing age-matched (HA) groups. The CA subgroup included 26 children aged between 3.77 and 10.0 years old (M = 7.08 yrs, SD = 1.67 yrs). In the HA subgroup, there were 26 children with the hearing age ranging between 3.33 and 9.49 years (M = 6.12 yrs, SD = 1.88 yrs). Note that the children with CIs were all prelingually-deafened, the influence of pre-surgery acoustic experience from hearing aid, if any, should be very limited. Therefore, we regarded the length of CI use as their hearing age. Some children with CIs were assigned to both subgroups because their chronological age and hearing age were both within the age range of the NH controls. Independent sample t-test revealed no significant group difference in age (both $p > 0.05$) between NH and CA and between NH and HA children.

2.2. Speech materials

All participants were recorded producing a list of 18 Mandarin disyllabic words containing six Mandarin stop consonants /p, t, k, pʰ, tʰ, kʰ/ each followed by three vowels /a, i, u/, respectively. Note that the vowel /i/ was substituted with /ɤ/ for velar stops /k, kʰ/ due to the phonotactic constraints in Mandarin. The target stop sounds all occurred at the word-initial position in the first syllable. The selection of these Mandarin words considered the vocabulary size of young children and the picturability of the target words. By virtue of the same consideration, the tone environment was not strictly controlled.

2.3. Recording

Speech recording was conducted through a visual-auditory paradigm in a quiet room or sound-treated auditory booth. Each participant was seated in front of a laptop computer with which the pictures representing the target words were presented on the screen in random order. An audio prime was then played and the participants were asked to repeat the target word. The speech samples were recorded through a digital recorder with a 16-bit quantization rate and 44.1 kHz sampling rate. All recorded speech samples were transferred to a desktop hard disk with the same quantization rate and sampling rate. The speech samples were then segmented into individual words using a spectrographic analysis program (Adobe Audition 3.0) and saved as separate .wav files.

2.4. Acoustic measurements

The productions that were identified as stop sounds were subject to acoustic analysis. The acoustic measures included VOT, ΔVOT, normalized amplitudes, and F2 onset. The landmark locations of each token including the burst onset, stop offset/vowel onset, and vowel offset were determined by the observation of spectrograms with the assistance of waveform using the digital audio editing program (Adobe Audition 3.0). VOT was measured as the time interval from the release of oral occlusion to the onset of voicing [22]. Because Mandarin stops are all voiceless stops, the onset of voicing was defined as the start of the periodicity of the following vowel. ΔVOT was calculated for the aspirated-unaspirated pair in each place of articulation followed by each vowel. This measure indexes the magnitude of aspirated-unaspirated contrast. Normalized amplitude was calculated to examine the airflow control. To calculate normalized amplitude, the root-mean-square (RMS) amplitude of the stop and the RMS of the following vowel were calculated, then the difference in dB between these two segments was computed. To measure F2 onset, a vowel analysis program TF32 [23] was used to extract formant tracks. A manual correction was implemented when errors occurred in formant extraction. Then, the frequency of F2 onset was measured at the start point of the vowel periodicity. Considering that the participants in the present study varied in chronological age, the F2 values were converted to z scores for each individual speaker [24] to minimize the influence of vocal tract size on the formant values. To ensure ease of interpretation, the z scores were rescaled to Hz-like values following the methods in Thomas and Kendall (2007) [25]. Note that only /a/ and /u/ vowels were used for the analysis of F2 onset because they are the common vowel context used for all six stops.

3. RESULTS

Fig. 1 presents the VOT of NH children and two CI subgroups for each stop. All three groups of children produced longer VOTs for the aspirated stops than the unaspirated stops. Of the three places, the NH children produced longer VOT for both unaspirated

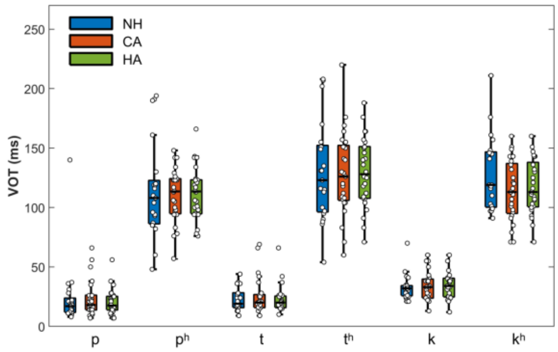

Fig. 1. Box plot showing the VOT of each Mandarin stop in the NH, CA, and HA children.

and aspirated velar stops. The two CI subgroups showed similar place patterns only for the unaspirated stops. A Linear Mixed-Effects Model (LMM) was used to examine VOTs between the NH children and each CI subgroup, respectively. The factors including children's group (NH vs. CA or NH vs. HA), place of articulation, and aspiration were set as fixed effects and the subject was set as a random effect with a random intercept for the subject included. For the NH-CA comparison, the results revealed a significant place difference ($F(2, 223.9) = 4.77$, $p = 0.009$) and aspiration difference ($F(1,223.7) = 1187.3$, $p < 0.0001$) and a place by aspiration interaction ($F(2, 223.4) = 3.78$, $p = 0.024$). For the NH-HA comparison, the results revealed a significant place difference ($F(2, 224.6) = 6.67$, $p = 0.002$) and aspiration difference ($F(1,224.4) = 1523.2$, $p < 0.0001$) and a place by aspiration interaction ($F(2, 224.1) = 5.07$, $p = 0.007$). No significant group difference or group-related interactions were found for both comparisons.

The ΔVOTs across vowel context are shown in Fig. 2. The data were fitted with an LMM in which the children's group (NH vs CA, NH vs HA) and place of articulation were set as the fixed effects, and the subject was set as a random effect with a random intercept for subject included. For the comparison between NH and CA, the results revealed no significant group or place difference. For the comparison between NH and HA, there was a significant place difference ($F(2, 86.8) = 4.3$, $p = 0.016$) but no significant group difference.

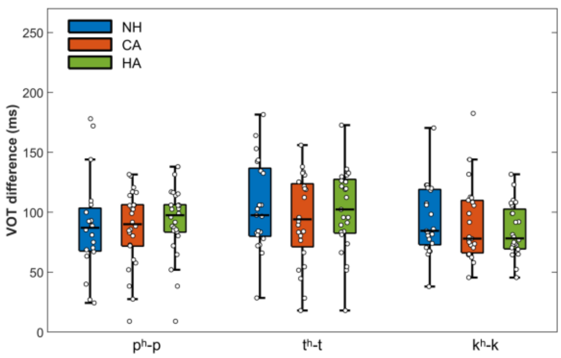

Fig. 2. Box plot showing the ΔVOT for the bilabial, alveolar, and velar stops in the NH, CA, and HA children.

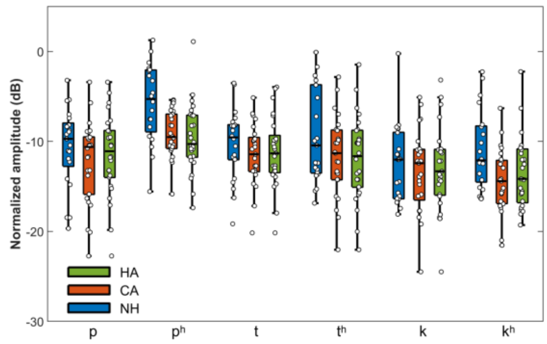

Fig. 3. Box plot showing the normalized amplitude of each Mandarin stop in the NH, CA, and HA children.

Fig. 3 presents the normalized amplitudes of the six Mandarin stops collapsed across vowel contexts in the NH and two CI subgroups. The children with NH showed a higher amplitude for aspirated stops than unaspirated stops and the velar stops were produced with lower amplitudes than the other two places. Both CA and HA subgroups followed the NH pattern varying with the place of articulation, but they produced lower normalized amplitude than the NH controls, especially for the aspirated stops. An LMM was implemented to compare the normalized amplitude between the NH and each CI subgroup, respectively. For the NH-CA comparison, the results revealed a significant group effect ($F(1, 46.6) = 6.83$, $p = 0.012$), place effect ($F(2, 222.8) = 23.65$, $p < 0.001$), and aspiration effect ($F(1, 222.6) = 14.18$, $p < 0.001$). Meanwhile, the results revealed a significant group by aspiration interaction ($F(1,222.6) = 5.42$, $p = 0.021$) and place by aspiration interaction ($F(2, 222.4) = 9.12$, $p < 0.001$). For the NH-HA comparison, the results revealed a significant group effect ($F(1, 46.5) = 4.10$, $p = 0.049$), place effect ($F(2, 223.4) = 20.83$, $p < 0.001$), and aspiration effect ($F(1, 223.2) = 11.51$, $p < 0.001$). Meanwhile, the results revealed a significant group by aspiration interaction ($F(1,223.2) = 8.29$, $p = 0.004$) and place by aspiration interaction ($F(2, 223.0) = 6.82$, $p = 0.001$).

Fig. 4 shows the F2 onset of the six stops followed by the vowels /a/ and /u/ in the NH group and two CI subgroups. All children, NH or CI alike, produced higher F2 for both /a/ and /u/ following the unaspirated alveolar stop /t/. An LMM was used to compare the group difference in F2 onset in /a/ and /u/ contexts, respectively. In the /a/ context, the NH-CA comparison yielded a significant place ($F(2, 203.5) = 14.130$, $p < 0.001$) and group effect ($F(1, 47.0) = 6.76$, $p = 0.012$) and a significant place by aspiration interaction ($F(2, 202.8) = 10.79$, $p < 0.001$). The NH-HA comparison yielded a significant place effect ($F(2, 206.9) = 15.18$, $p < 0.001$) and place by aspiration interaction ($F(2, 206.04) = 8.50$, $p < 0.001$). In the /u/ context, the NH-CA comparison yielded a

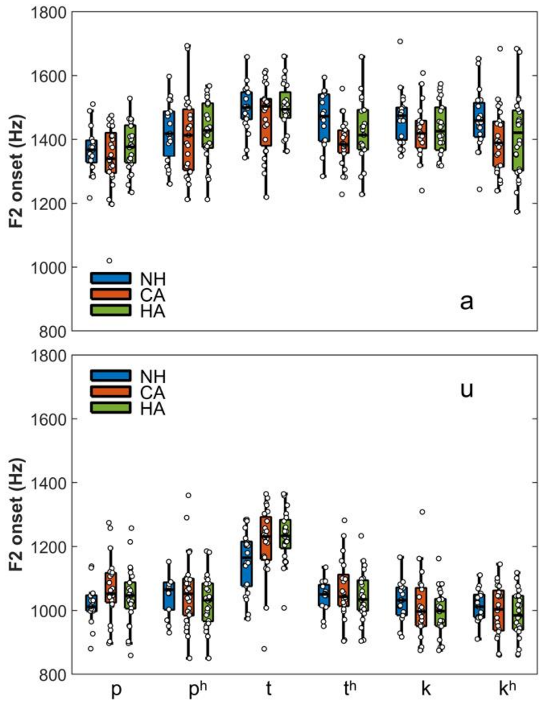

Fig. 4. Box plot showing F2 onset for each Mandarin stop followed by /a/ (top) and /u/ (bottom) in the NH, CA, and HA children.

significant place ($F(2,188.4) = 36.04$, $p < 0.0001$) and aspiration effect ($F(1,190.7) = 19.23$, $p < 0.0001$) and a significant place by aspiration interaction ($F(2,187.7) = 16.66$, $p < 0.001$) but no significant group effect or group related interactions. The NH-HA comparison yielded significant place ($F(2, 194.7) = 58.91$, $p < 0.0001$) and aspiration effects ($F(1,196.4) = 35.74$, $p < 0.0001$) but no significant group effect. All two-way interactions were significant (all $p < 0.05$).

4. DISCUSSION

VOT is the "acoustic result of laryngeal-oral coordination reflected in the temporal domain" (p.1336, [26]). Lane and colleagues (1994) [16] reported that postlingually deafened CI users modified the VOT values for English voiced and voiceless stops in accordance with the change in their hearing status caused by CIs. Uchanski and Geers (2003) [18] found that although the CI children demonstrated greater inter-subject variations in the VOT for English /t/ and /d/, over 60% of CI children who used either total communication or oral communication mode had VOTs and ΔVOT values within the normal-hearing range. In the present study, our data revealed that the CI children, regardless of CA or HA, showed no significant difference from the NH children on the VOT for unaspirated or aspirated stops or the ΔVOT for the aspirated-unaspirated contrasts. These results suggested that the children with CIs approximated the NH children and could use the temporal features to effectively distinguish aspiration contrasts in their stop production.

Among the tested acoustic measures, the normalized amplitude is associated with respiratory and airflow control during the process of occlusion and release. Lane and Perkell (2005) [26] proposed that the reduced aerodynamic distinction between voiced and voiceless stops in prelingually deafened speakers might be reflected as having higher than normal peak flow for voiced stops while lower than normal peak flow for voiceless stops. Other researchers also found that deaf children produced lower expiratory airflow than NH peers did [27]. In this study, we observed that both CI subgroups produced significantly lower normalized amplitude than the NH control for the aspirated stops but showed no significant difference from the NH controls for the unaspirated stops. The production of aspirated stops involves a strong airflow immediately following the release burst, which therefore has a higher amplitude than the unaspirated stops. Our data suggested that the aerodynamic distinction of amplitude between unaspirated and aspirated stops in the children with CIs was reduced than the NH controls.

F2 onset contains information regarding the placement of articulation [28, 29]. Of the three places, the alveolar stops have spectral prominence at a higher frequency than the bilabial stops [28]. The F2 onset following the alveolar stops is normally higher than that following bilabial stops. Our data showed significant place difference, in particular, higher F2 onset following the alveolar stops in both /a/ and /u/ contexts in the NH group, which was consistent with the pattern reported in previous studies. The two CI subgroups generally followed the NH pattern of F2 onset varying as a function of place of articulation. This finding suggested that children with CIs could differentiate the place of articulation in their stop production. While F2 onset reflects the articulatory configuration of the preceding consonant, it is also determined by the vowel target. The group difference between the CA and NH children might be partially accounted for by less definable vowel productions in children with CIs [30].

In summary, our acoustic data indicated that Mandarin-speaking children with CIs had not fully developed NH-like respiratory control for stop production, but they could produce NH-like temporal features for the aspirated-unaspirated contrast and produce distinguishable place features in their stops. All three groups of children showed considerable variabilities in the tested measures, which indicated that the acoustic development of stop consonants is a long-term process.

7. REFERENCES

[1] Dodd, B., Holm, A., Hua, Z., Crosbie, S. 2003. Phonological development: a normative study of British English-speaking children. *Clin Linguist. Phon.* 17(8), 617-643.

[2] Smit, A., Hand, L., Freilinger, J., Bernthal, J. Bird, A. 1990. The Iowa articulation norms project and its Nebraska replication. *J. Speech. Hear. Disord.*, 55, 779–798.

[3] Vihman, M., 1996. Phonological Development (Oxford: Blackwell).

[4] Imbrie, A. K. K. 2005. *Acoustical study of the development of stop consonants in children* (Doctoral dissertation, Massachusetts Institute of Technology).

[5] Yang, J. 2018. Development of stop consonants in 3-to-6-year-old Mandarin-speaking children. *J. Child Lang. 45*, 1091-1115.

[6] Liu, C. T. 2021. A first step toward the clinical application of landmark-based acoustic analysis in child Mandarin. *Children*, *8*(2), 159.

[7] Monsen, R. B. 1976. Normal and reduced phonological space: The production of English vowels by deaf adolescents. *J. Phon. 4*(3), 189-198.

[8] Harris, K. S., Rubin-Spitz, J., McGarr, N. S. 1985. The role of production variability in normal and deviant developing speech. In J. L. Lauter (Ed.), *Proceedings of the Conference on the Planning and Production of Speech in Normal and Hearing-Impaired Individuals: A Seminar in Honor of S. Richard Silverman* (pp. 50-57). Rockville, MD: American Speech-Language-Hearing Association.

[9] Metz, D. E., Samar, V. J., Schiavetti, N., Sitler, R. W. 1990. Acoustic dimensions of hearing-impaired speakers' intelligibility: Segmental and suprasegmental characteristics. *J. Speech Lang. Hear. Res*. *33*(3), 476-487.

[10] Slis, I. H., Cohen, A. 1969. On the complex regulating the voiced-voiceless distinction I. *Lang. Speech*. 12(2), 80-102.

[11] Higgins, M. B., Carney, A. E., McCleary, E., Rogers, S. 1996. Negative intraoral air pressures of deaf children with cochlear implants: Physiology, phonology, and treatment. *J. Speech. Lang. Hear. Res*. *39*(5), 957-967.

[12] Serry, T. A., Blamey, P. J. 1999. A 4-year investigation into phonetic inventory development in young cochlear implant users. *J. Speech. Lang. Hear. Res*. *42*(1), 141-154.

[13] Tobey, E. A., Pancamo, S., Staller, S. J., Brimacombe, J. A., Beiter, A. L. 1991. Consonant production in children receiving a multichannel cochlear implant. *Ear. Hear*. 12, 23–31.

[14] Grogan, M. L., Barker, E. J., Dettman, S. J., Blamey, P. J. 1995. Phonetic and phonological changes in the connected speech of children using a cochlear implant. *Ann. Otol. Rhinol. Laryngol. Suppl. 166, 390–393.*

[15] Economou, A., Tartter, V. C., Chute, P. M., Hellman, S. A. 1992. Speech changes following reimplantation from a single-channel to a multichannel cochlear implant. *J. Acoust. Soc. Am*. *92*(3), 1310-1323.

[16] Lane, H., Wozniak, J., Perkell, J. 1994. Changes in voice-onset time in speakers with cochlear implants. *J. Acoust. Soc. Am*. *96*(1), 56-64.

[17] Kishon-Rabin, L., Taitelbaum, R., Tobin, Y., Hildesheimer, M. 1999. The effect of partially restored hearing on speech production of postlingually deafened adults with multichannel cochlear implants. *J. Acoust. Soc. Am*. *106*(5), 2843-2857.

[18] Uchanski, R. M., Geers, A. E. 2003. Acoustic characteristics of the speech of young cochlear implant users: A comparison with normal-hearing age-mates. *Ear. Hear*. *24*(1), 90S-105S.

[19] Higgins, M. B., McCleary, E. A., Carney, A. E., Schulte, L. 2003. Longitudinal changes in children's speech and voice physiology after cochlear implantation. *Ear. Hear*. *24*(1), 48-70.

[20] Yu. J., Xia, X. 2019. Production of Mandarin stop consonants in prelingually deaf children with cochlear implants. In S. Calhoun, P.Escudero, M. Tabain & P. Warren (eds.) *Proc. 19th ICPhS Melbourne, Australia*

[21] Joy, D. A., Sreedevi, N. 2019. Temporal characteristics of stop consonants in pediatric cochlear implant users. *Cochlear. Implants. Int*. *20*(5), 242-249.

[22] Lisker L. Abramson A.S. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384-422.

[23] Milenkovic, P. (2003). TF32 software program. Madison: University of Wisconsin

[24] Lobanov, B. M. 1971. Classification of Russian vowels spoken by different listeners. *J. Acoust. Soc. Am*. 49, 606-08.

[25] Thomas, Erik R. Tyler Kendall. 2007. NORM: The vowel normalization and plotting suite. Retrieved November 14, 2022, from http://lingtools.uoregon.edu/norm/about_norm1.php#scaling.

[26] Lane, H., Perkell, J. S. 2005. Control of voice-onset time in the absence of hearing. *J. Speech Lang. Hear. Res*. 48(6), 1334-1343.

[27] Żebrowska, A., Zwierzchowska, A., Manowska, B., Przybyła, K., Krużyńska, A., Jastrzębski, D. 2016. Respiratory function and language abilities of profoundly deaf adolescents with and without cochlear implants. In *Prospect in Pediatric Diseases Medicine* (pp. 73-81). Springer, Cham.

[28] Delattre, P. C., Liberman, A. M., Cooper, F. S. 1955. Acoustic loci and transitional cues for consonants. *J. Acoust. Soc. Am*. *27*(4), 769-773.

[29] Cassidy, S. Harrington, J. M. 1995. The place of articulation distinction in voiced oral stops: evidence from burst spectra and formant transitions. *Phonetica*. 52, 263-284.

[30] Yang, J., Xu, L. 2021. Vowel production in prelingually-deafened Mandarin-speaking children with cochlear implants. *J. Speech Lang. Hear. Res*. 64(2), 664-682.