PRODUCTION OF MANDARIN STOP CONSONANTS IN PRELINGUALLY DEAF CHILDREN WITH COCHLEAR IMPLANTS *

Jue YU Xinyu XIA

School of Foreign Languages, Tongji University, Shanghai, China
erinyu@126.com, xinyuxia@yeah.net

ABSTRACT

The present study examined the production performance of word-initial stops by prelingually deaf Mandarin-speaking children with cochlear implants (CIs). The main purpose was to obtain a comprehensive understanding about how CI children acquired Mandarin stops. The results showed that, compared with normal-hearing (NH) children, CI children followed the same development path but applied some non-developmental phonological rules; made more articulatory errors, especially in the aspirated-unaspirated bilabial cognate pair if given no visual information; presented significantly different VOT patterns, particularly in aspirated stops which reflected their inadequacy in VOT manipulation; and were affected by the articulatory constraints in a similar way to NH children across different vowel contexts. Only the length of CI device use was found to be a significant contributor for accuracy performance of CI children but no systematic correlation was found between VOT production skills and the use of cochlear implant devices.

Keywords: Cochlear implants, Mandarin stops, articulatory, accuracy, VOT pattern

1. OBJECTIVES AND BACKGROUND

Recently, many researchers have focused on the consonant development in cochlear implanted (CI) children. Among all the consonants, stops are of particular interest because they form the most popular sound group in all world languages and cross-linguistically, are one of the earliest sound classes to be normally acquired [4], [10]. While for CI children, they frequently demonstrate more difficulties in perceiving and discriminating stop consonants for their short duration, ambiguous spectral pattern and coarticulation with vowels. There are substantial studies on CI children's development and performance of stop consonants from English and other language backgrounds. Yet few published studies have comprehensively documented the production performance of stop consonants in Mandarin-speaking children with CIs.

Compared with English stops, Mandarin Chinese also contains six stops produced in bilabial, alveolar and velar position. But all Mandarin stops are voiceless and further contrasted by the presence or absence of aspiration. Besides, Mandarin stops only occur in word-initial position, not final. Acquisition of stops is really important for Mandarin-speaking children with CIs, for it can also facilitate the development of affricates. Moreover, failure in contrasting aspiration feature in Mandarin leads to confusion in lexical meaning.

This paper will examine the production performance of word-initial stops by prelingually deaf Mandarin-speaking children with CIs. The main purpose is to obtain a comprehensive understanding about how CI children acquire Mandarin stops. Three research questions are proposed:

1. To what extent the children with CIs can accurately produce Mandarin stops, in terms of accuracy rates and error patterns?
2. What are the differences in VOT patterns of Mandarin stops produced by CI and NH children, in terms of aspirated-unaspirated distinction and place distinction across different vowel contexts?
3. How are the Mandarin-speaking children's post-implant stop production skills associated with the factors of CI devices?

2. METHOD

2.1. Stimuli and procedure

Each participant produced a list of 23 Mandarin CV monosyllables that contained all the target stop consonants /p, t, k, p^h, t^h, k^h/ in the word-initial position. Considering the variation of VOT values related to the following vowel environment, all the possible combinations of a given stop and vowels in CV syllable structure are elicited [5], [8]. The effect of lexical tone on the word-initial voiceless stops was not of concern in the present study [1], [12].

All the target words were elicited through a modified imitative task and randomised in order not to be predictable. For each target word, the participants first heard an audio prompt (naturally

* This paper is sponsored by the Fundamental Research Funds for the Central Universities, China (Grant No. 22120170497)

produced by a female native Mandarin speaker), and were then asked to repeat the word once immediately after the audio prime (repetition was allowed). The reasons for using audio prompt instead of an adult clinician's utterance are first to avoid the possible exaggeration by the clinician to facilitate the child's best production, and second, to maintain the homogeneity for each target stimuli [15].

2.2. Participants

Altogether, 20 prelingually deaf Mandarin-speaking children with unilateral multi-channel CIs (11 male and 9 female) were recruited and 10 NH children (5 male and 5 female) were used as a baseline. The CI participants were aged between 4.33 to 12.67 years (mean: 9.12 years) at the time of recording and their average length of CI experience was 6.28 years. All of them were non-verbal prior to implantation and were reported to have no visual, developmental or cognitive problems except for a hearing impairment. They received their implants at an average age of 2.84 years, a relatively early implantation. After the surgery, all of the children received intensive speech and language training at professional rehabilitation centres in Shanghai and mainly used oral communication. Detailed demographic information for CI children is omitted due to the length limitation.

The average age of the NH participants were 6.97 years (range from 3.3 to 10.5), chronologically slightly younger than CI group, but well matched the average hearing age. All had been reported without language or speech impairments. All participants use spoken Mandarin at home and in their daily life.

2.3. Data analysis

Both NH and CI children's stop productions were first transcribed by two native Mandarin speakers trained in phonetics and if there was inconsistency between the two transcribers, the first author double-checked it and made a final decision. The transcribers were instructed to code the production of initial stops as accurate (including acceptable but distorted) or mispronounced. The mispronounced sounds were then analysed for the accuracy rates and error patterns.

After transcription, all speech samples (except those mispronounced ones), were annotated for further acoustic analysis in Praat [2]. Acoustic analysis mainly focused on VOT, the dominant measure for discriminating stop categories. VOT was measured as the time interval from the onset of the first release burst to the onset of voicing [11]. In order to exclude the influence of speaker variability and speech rate, normalization was conducted to all the duration parameters by calculating VOT ratio in the whole syllable. The general VOT patterns of Mandarin word-initial stops produced by CI and NH children were first compared. Then the aspirated-unaspirated VOT distinction in CI and NH children were revealed. Finally, vowel context was examined to determine whether VOT pattern in CI children was affected by the articulatory constraints in a similar way to NH children.

A mixed factor repeated measures ANOVA was conducted to examine the effect of group (between-subject factor) and place of articulation, manner of articulation (aspiration), and vowel context (within-subject factor) on VOT values. Bonferroni correction was applied for the pair-wise comparisons of the within-subject factor. Finally, Pearson correlation analysis was performed, so the interrelation between the use of CI devices and Mandarin stop production in prelingually deaf children can be observed.

3. RESULTS AND DISCUSSION

3.1. Accuracy metrics and error patterns

In order to gain a better understanding of the differences in accuracy and error pattern between CI and NH children, mean percentages of different error types for each stop in both groups were presented in Table 2. It turned out that most CI and NH children had a complete phonetic inventory of initial stop consonants before 4 years of age (for CI children, 4 years of hearing age) and could averagely produce most initial stops with the mastery level accuracy, (above 75% accuracy) [14], [15], which is consistent with the results of previous studies (such as [7], [12]).

Not surprisingly, the CI group produced six Mandarin stops with significantly lower accuracy than the NH group. Moreover, both groups produced the aspirated-unaspirated bilabial cognate pair with the lowest accuracy, followed by velar stops and the alveolar ones. This meant that CI children followed the same development path for initial stops with NH children, though somehow delayed in acquisition. What was noteworthy in the present study, was that the production accuracy of initial stops decreased with the visibility of the places of articulation in both groups (especially in CI participants). It seemed that the labial stops (especially the unaspirated one, with 43.9% error ratio in CI group and 11.5% in NH group), became the most challenging consonants for all Mandarin-speaking children. The possible explanation for this controversial issue lied on the difference in the elicitation and recording procedures. Audio prompt was used to directly elicit speech production rather than a traditional imitation task which involved an adult clinician's utterance. Therefore, visual information was not provided. Generally speaking, aspirated stops were better mastered than unaspirated ones in both groups. This may be due to the perceptual saliency of aspirated stops. For the voiceless unaspirated stops, the duration is shorter and intensity is lower. This makes the recognition task more difficult, which will definitely distort the stop production.

Table 1: The mean percentages of different error types for each stop produced by CI and NH children

Group	Error type	/p/	/t/	/k/	/pʰ/	/tʰ/	/kʰ/
NH	Manner	5.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	POA	6.5%	2.0%	2.0%	2.5%	0.0%	2.0%
	Deletion	0.0%	0.0%	5.0%	5.0%	0.0%	0.0%
	Total	11.5%	2.0%	7.0%	7.5%	0.0%	2.0%
CI	Manner	13.5%	0.0%	1.4%	0.6%	0.0%	2.5%
	POA	24.0%	8.2%	10.4%	15.6%	4.6%	10.3%
	Deletion	6.4%	1.0%	3.2%	0.0%	0.0%	0.0%
	Total	43.9%	9.2%	15.0%	16.2%	4.6%	12.8%

Regarding the effects of vowel context on stop production, CI children obviously produced stops in each vowel context with significant lower accuracy than NH children, some of which even four times lower. For example, the CI participants produced stops followed by central vowel with 26.9% error ratio and the NH group with 7.0%. Moreover, it showed more accuracy of stop production when followed by front vowel in both CI and NH group (error ratio: CI group 8.7%; NH group 2.0%), i.e. the more front places of articulation, the more accuracy of initial stop production. This finding is controversial with previous studies [5], [6] and still needs further examination due to the insufficient data in front vowel context caused by the phonotactic constraint on velar stops. But anyway, it can be concluded that vowel context did exert an impact on the articulation accuracy of Mandarin stops.

Finally, further qualitative examination of the error patterns indicated that besides the general developmental phonological rules applied by most NH children such as deletion, affrication and deaspiration, CI children also applied the non-developmental phonological rules, such as /h/ replacement for aspirated initial consonant (/pʰ/--[h]), labilization (/p/--[w]) and nasalization (/p/--/m/), which led to more diverse substitutions.

3.2. Acoustic analysis: VOT

3.2.1. Overall VOT means and distribution

Undoubtedly, the VOT duration of Mandarin aspirated stops were longer than that of unaspirated ones in both NH and CI children. But, compared with [3], the present study demonstrated much longer VOT duration (or higher VOT ratio) for aspirated voiceless stops, mostly over 100ms; while slightly shorter VOT duration for unaspirated voiceless stops (except for /k/) in all participants (See Table 2). It indicated that all aspirated stops in Mandarin were located at the highly aspirated region, while unaspirated ones (except for /k/) in unaspirated region. Compared with English voiceless stops, it is more salient in aspirated-unaspirated distinction. Both NH and CI children were able to make aspirated-unaspirated phonemic distinction, but the velar stop /k/ seemed somehow controversial in both groups. It belonged to the slightly aspirated stops in NH group; while approached the unaspirated category in CI group according to [3].

Table 2: Overall VOT means and normalized ratio produced by CI and NH children

Group	VOT	/p/	/t/	/k/	/pʰ/	/tʰ/	/kʰ/
CI	VOT mean raw	24.3	23.4	38.3	120.7	104.2	112.4
	VOT ratio	0.06	0.06	0.08	0.23	0.21	0.21
NH	VOT mean raw	23.5	18.5	51.6	158.1	141.0	154.7
	VOT ratio	0.05	0.04	0.1	0.31	0.27	0.29

In addition, the CI group generally produced aspirated stops with shorter VOTs than the NH group, which reflected the immature respiration control and higher subglottal pressure in CI children. With regard to the influence of place of articulation on VOT, the VOTs of alveolar stops were the shortest among all Mandarin stops in both groups. Velar stops showed different results under different articulatory conditions. It only owned the longest VOT under unaspirated condition. This finding is partially consistent with [13], [9].

Statistical analysis revealed a main effect for group differences between CI and NH children (($F(1, 562) = 29.32$, $p<0.001$). Subsequent post hoc Bonferroni adjusted *t* tests showed significant differences in aspirated stops between CI and NH children. Interaction effect of group by stop was found significant ($F(5, 562) = 11.84$, $p < 0.001$). These findings indicated that there were significant VOT differences in all Mandarin stops produced by CI and NH children, especially the aspirated ones. Moreover, for each group, VOT can be a reliable acoustic measure to systematically distinguish Mandarin stops.

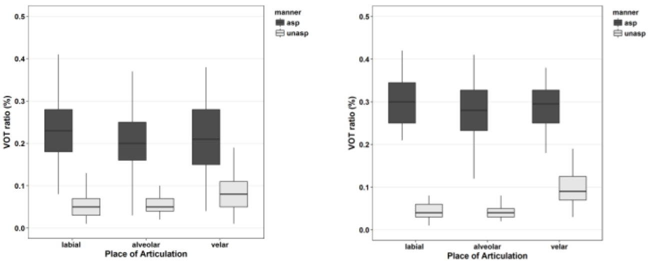
3.2.2 Aspirated-unaspirated VOT distinctions

Figure 1 shows the aspirated-unaspirated VOT distinctions in the CI group and the NH group respectively. VOT mean ratio is used instead of VOT raw mean, in order to diminish the influence of speaker variability and speech rate. The results showed that although both NH and CI children were able to well distinguish aspirated-unaspirated cognate pairs, there was a great discrepancy in VOT patterns between the two groups. In the NH group, the VOT duration of aspirated stops is about 3-6 times longer than that of the unaspirated cognate pair, while no such significant differences were observed in the CI group. This meant that compared with NH children, CI children were still incapable of VOT manipulation.

Moreover, the aspirated-unaspirated VOT pattern for the velar cognate pairs was very different from the other pairs, which showed the least salient distinction between aspirated and unaspirated in both groups (especially in CI children). This meant that velar stops were the most difficult to be acquired by all children.

The ANOVA results yield a significant VOT difference between aspirated and unaspirated stops (F_{CI} (1, 346) = 549.676, $p < 0.001$; F_{NH} (1, 216) = 821.27, $p < 0.001$), as well as between different places of articulation (F_{CI} (2, 346) = 5.052, $p < 0.01$; F_{NH} (2, 216) = 13.29, $p < 0.001$) in both CI and NH children. Interaction effect of manner by place was also found significant (F_{CI} (2, 346) = 3.650, $p < 0.05$; F_{NH} (2, 216) = 5.77, $p < 0.01$). These statistical results indicated that within either group, Mandarin stops demonstrated different aspirated-unaspirated VOT patterns across the different places of articulation, which was exactly shown in Figure 1.

Figure 1: Box plot showing aspirated-unaspirated VOT distinctions across different places of articulation. (Left: CI group; Right: NH group)

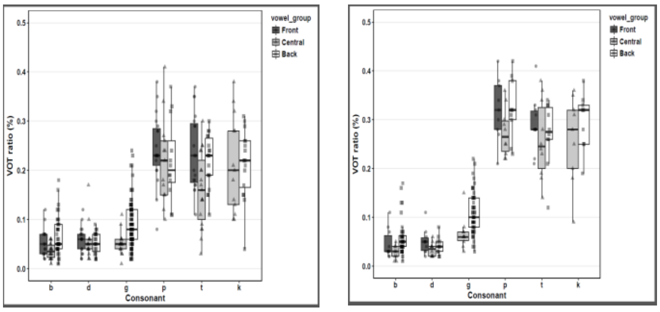


3.2.3. Vowel context

Between-group differences in VOT patterns of Mandarin stops across different vowel context were examined in Figure 2. Due to the phonotactic limitation, velar stops were not able to be followed by front vowel.

In either CI group or NH group, all Mandarin stops showed shorter VOT duration when followed by central vowel than front or back vowel, which is consistent with previous results [13], [9]. The ANOVA results revealed a significant main effect of the vowel context (F_{CI} (2, 336) = 9.223, $p < 0.001$; F_{NH} (2, 206) = 5.882, $p < 0.01$) but the interaction effect of stop by vowel was not found significant (F_{CI} (2, 336) = 1.787, $p > 0.05$; F_{NH} (2, 206) = 0.708, $p > 0.05$). Therefore, the VOT patterns of Mandarin stops were affected when followed by different vowels.

Figure 2: Box plot showing VOT patterns under different vowel context. (Left: CI group; Right: NH group. Stops presented in Pinyin form)



3.3. Interactions between CI device use and Mandarin stop production

The interrelation between stop production performance (in terms of articulatory accuracy and VOT pattern) and the use of CI devices (mainly in terms of age at implantation and the length of CI device use) were examined by Pearson correlation analysis in this part. The length of CI device use was found to be the only significant predictor for overall stop articulatory accuracy (r^2=0.514, $p<0.05$). No significant correlation was found between the use of cochlear devices and VOT measure (VOT ratio) of Mandarin stops. The acoustic discrepancy may be relevant to other factors, such as the intrinsic temporal resolution of CI devices or individual unbalanced performance, which needs to be further studied.

4. CONCLUSION

The present study examined the production performance of word-initial stops by prelingually deaf Mandarin-speaking children with cochlear implants (CIs). Some interesting results were illustrated here: 1) compared with NH children, CI children followed the same development path but applied some non-developmental phonological rules, such as /h/ replacement, labilization and nasalization; 2) comparatively speaking, CI children made more articulatory errors than NH children, especially in the aspirated-unaspirated bilabial cognate pair if given no visual information; 3) CI children presented significantly different VOT patterns with NH children, particularly in aspirated stops which reflected their incapability of VOT manipulation; 4) vowel context did exert an impact on the production performance of Mandarin stops; 5) Only the length of CI device use was found to be a significant contributor for accuracy performance of CI children, but no systematic correlation was found between VOT production skills and the use of cochlear implant devices.

5. REFERENCES

[1] Bradshaw, M. 1999. A Crosslinguistic Study of Consonant-Tone Interaction. (Ph.D. dissertation, The Ohio State University). Retrieved from ProQuest Dissertations and Theses. Publication No. AAT 9941291.

[2] Boersma, P., & Weenink, D. 2010. Praat, a system for doing phonetics by computer. (Version 5.1.45). Available at: www.praat.org. Accessed December 1, 2013. Google Scholar.

[3] Cho, T., & Ladefoged, P. 1999. Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, 27(2), 207–29.

[4] Dodd, B., Holm, A., Hua, Z., & Crosbie, S. 2003. Phonological development: a normative study of British English-speaking children. *Clinical Linguistics & Phonetics*, 17(8), 617 – 43.

[5] Donaldson, G.S., & Kreft, H.A. 2006. Effects of vowel context on the recognition of initial and medial consonants by cochlear implant users. Ear and Hearing, 27 (6) ,658-677.

[6] Dubno J R, Dirks D D, Schaefer A B. 1987. Effects of hearing loss on utilization of short-duration spectral cues in stop consonant recognition. [J]. *Acoustical Society of America. Journal*, 81(6).

[7] Hua, Z., & Dodd, B. 2000. The phonological acquisition of Putonghua. *Journal of Child Language*, 27, 3–42.

[8] Klatt, D. H. 1975. Voice Onset Time, frication, and aspiration in word-initial consonant clusters. Journal of Speech and Hearing Research, 18, 686–706.

[9] K.-Y. Chao, G. Khattab, and L.- M. Chen. 2006. Comparison of VOT Patterns in Mandarin Chinese and in English, *Proc. 4th Annual Hawaii International Conference on Arts and Humanities* Honolulu, pp. 840–859.

[10] Ladefoged. P. 1992. Another View of Endangered Languages. *Language*, 68(4), 809-811.

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

[12] 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 and Neck Surgery*, 130, 592–697.

[13] Rochet, B. L., & Fei, Y. 1991. Effect of consonant and vowel context on Mandarin Chinese VOT: production and perception. *Canadian Acoustics*, 19(4), 105–6.

[14] Shriberg, L. 1993. Four new speech and voice-prosody measures for genetics research and other studies in developmental phonological disorders. Journal of *Speech, Language, and Hearing Research*, 36, 105–140.

[15] Tye-Murray, N., & Kirk, K.I. 1993. Vowel and Diphthong Production by Young Users of Cochlear Implants and the Relationship Between the Phonetic Level Evaluation and Spontaneous Speech. *Journal of speech and hearing research*. 36. 488-502.