Do 3- to 4-year-old Mandarin-English bilinguals separate long-lag VOTs in these two languages?

Jing Yang[1], Li Xu[2]

[1]Communication Sciences and Disorders, University of Wisconsin-Milwaukee, USA
[2]Communication Sciences and Disorders, Ohio University, USA
jyang888@uwm.edu, xul@ohio.edu

ABSTRACT

Previous research has reported that Mandarin and English both have short-lag and long-lag VOTs in similar regions along the VOT continuum. However, Mandarin long-lag VOTs are longer than English long-lag VOTs in adults. The present study examined whether 3- to 4-year-old bilingual Mandarin-English children and corresponding monolingual children demonstrate the subtle but noticeable language difference that adult speakers do. The participants included 10 bilingual Mandarin-English children, 20 corresponding monolingual children, and 22 corresponding monolingual adults. The speech materials included 9 Mandarin words containing three Mandarin aspirated stops and/or 9 English words containing three English voiceless stops. The VOT values were measured using Adobe Audition 1.0. The results showed that unlike adults, the monolingual children and bilingual children at this young age did not show the language difference on long-lag VOTs. However, the VOT patterns were different in the bilingual children who differed in the amount of experience in English.

Keywords: Bilingual Mandarin-English children, monolingual children, VOT

1. INTRODUCTION

Voice Onset Time (VOT) refers to the time interval between the release of stop closure and the onset of glottal pulsing. This durational measurement contains information indexing the coordination of oral-laryngeal musculature for stop production and has been widely used as an acoustic feature to signify the voicing and aspiration status of stop consonants. According to the classical categorization proposed by Lisker and Abramson in 1960s, VOT values can be divided into three broad categories: voicing-lead (<-75 ms), short-lag (0 – 25 ms), and long-lag (> 60 ms) [1]. Mandarin and English both have stops produced at bilabial, alveolar and velar regions. However, these two languages have distinct phonological contrasts. English is characterized by a voiced vs voiceless distinction and Mandarin is characterized by an unaspirated vs aspirated distinction. Regardless of the phonological dichotomy, for word-initial stops in isolation, the two languages are realized with similar phonetic representation of short-lag vs. long-lag VOT distinction. Nonetheless, previous cross-linguistic comparisons of stop consonants by adult speakers revealed that Mandarin long-lag VOTs were slightly yet significantly longer than English long-lag VOTs but the short-lag VOTs in these two languages were quite similar [2, 3]. These findings indicate that adults still show language-specific VOT features for the phonological contrast. The longer long-lag VOTs in Mandarin aspirated stops than English voiced stops were not just reflected in adult speakers but also manifested in 5- to 6-year-old monolingual children of each language [4]. By contrast, this language difference was only present in the bilingual Mandarin-English children who were less experienced in English but not manifested in the bilingual children who were highly experienced in L2 (English).

So far, many bilingual studies have documented the development of VOT separation for aspirating vs voicing language contrast in bilingual children [5-9]. These studies reported that bilingual children successfully acquired distinct VOT patterns for both languages eventually, but not in a monolingual-like manner. Meanwhile, a substantial amount of previous research on the acquisition of voicing contrast in monolingual children has shown that long-lag VOTs are acquired later than short-lag VOTs [10-18]. These studies reported that while children produced voiced stops with adult-like short-lag VOTs before 3 years of age, the long-lag VOT for the voiceless stops were not as long as the adult targets although the long-lag VOTs demonstrated a gradual increase to approximate the adult values. Additionally, the long-lag VOTs of voiceless stops showed greater variability than the short-lag VOTs of voiced stops. Given the later acquisition and greater variability of long-lag VOTs in children, we wonder whether the language difference in long-lag VOTs shown in adults and older children is also present in monolingual children at a younger age; Further, whether and to what extent the language difference in long-lag VOTs are manifested in bilingual children at a younger age?

2. METHODS

2.1. Participants

The present study included a total of 52 participants: 10 bilingual Mandarin-English children at 3 to 4 years old, 14 monolingual Mandarin peers, 6 monolingual English peers, 12 Mandarin-speaking adults aged between 23 and 58 years old, and 10 English-speaking adults aged between 20 and 44 years old. The bilingual children were recruited from central Ohio region and were divided into two subgroups based on the amount of experience in L2 reported by their parents. The Bi-low children started to learn English at a relatively older age (M=37.4 mo) and had a shorter period of English learning experience (M=6 mo) with relatively low percentage of English use (< 50% daily use). The Bi-high children started to learn English at a relatively younger age (M=18.5 mo) and had a longer period of English learning experience (M=30.3 mo) with relatively high percentage of English use (> 50% daily use). All bilingual children were raised by Mandarin-speaking parents and had limited amount of exposure to English before they enrolled in English daycare or kindergarten. The monolingual Mandarin speakers were recruited from Beijing, China and the monolingual English speakers were recruited from central Ohio region. None of the children was reported as having speech, language, or hearing problems.

2.2. Materials

The recording materials included two word lists: 9 Mandarin disyllabic words and 9 English monosyllabic/disyllabic words. The Mandarin words contained three aspirated stops /p^h, t^h, k^h/ at the word-initial position each followed by three vowels /a, i, u/, respectively. Due to the phonotactic constraints in Mandarin, the vowel context /i/ was substituted with /ɤ/ for the velar stop /k^h/. The English words contained three voiceless stops /p, t, k/ in isolation at the word-initial position each followed by three vowels /ɑ, i, u/. For certain words, the vowels /oʊ, ɪ, ɔ, ʊ/ were used as the alternative vowel environments to ensure the familiarity of the target words to the young age children.

2.3. Recording and data analysis

The recordings were conducted in a quiet room. Each monolingual speaker was asked to produce one list of words in their native language and each bilingual speaker was required to produce the two lists of words in both Mandarin and English. To ensure a better control of stimulus presentation and to elicit the target words as expected [19], a visual-auditory word repetition task conducted through custom MATLAB program was used for recording. Individual pictures presenting the target words were displayed on a laptop computer screen, which were followed by audio prompts produced by native adult speakers of each language. The participants were asked to repeat the target word immediately after the audio primes. Each word list was repeated twice for each participant. The speech samples were recorded through a Shure SM10A head-mounted microphone connected to an amplifier with the computer with a 16-bit quantization rate and 44.1 kHz sampling rate.

The landmarks of the onset and offset of stops and following vowels were marked for individual token based on the waveform with a visual check of the spectrogram using Adobe Audition 1.0. VOT was measured from the beginning of stop burst to the beginning of vowel periodicity which was set as the zero-crossing point of the first glottal pulse for vowel production. Subject mean VOT values were calculated for further statistical analyses.

The VOT data were subject to statistical tests. First, a two-sample Kolmogorov–Smirnov (K-S) test was used to compare the distribution of all long-lag VOT data between Mandarin and English for each group of speakers. Then, a Linear Mixed-Effects model was applied to compare the language difference within each group of speakers and group difference within each language. For language comparisons, the language mode, place of articulation, and vowel context were defined as the fixed effects and the subject effect was set as a random effect. For group comparisons, the four groups of speakers, place of articulation, and vowel context were set as the fixed effects while the subject effect was defined as a random effect. As the effects of place and vowel on VOT have been well examined in previous studies [3, 18, 20], these two fixed effects were not reported and discussed here.

3. RESULTS

Figure 1 displays the comparison of the distribution of all long-lag VOT data between Mandarin and English for each group of speakers. Compared to monolingual and bilingual children, the adult speakers for both languages showed concentrated long-lag VOT distributions. English long-lag VOTs mostly occurred around 80 ms while the Mandarin long-lag VOTs mostly occurred around 100 ms. The K-S test revealed significantly different long-lag VOT distribution between the two languages in adult speakers (p=0.001). As for the monolingual children, although a certain portion of Mandarin long-lag

VOTs were located at a higher value region than English long-lag VOTs, the overall distributions of the long-lag VOTs in Mandarin and English did not differ as much as those in the monolingual adult speakers. The K-S test yielded no significant difference between the two languages in monolingual children (p>0.05). Similar to monolingual children, neither Bi-low nor Bi-high children showed evident separation between Mandarin and English long-lag VOTs. The K-S tests confirmed this observation (all p>0.05). However, the long-lag VOTs for both languages in the Bi-low children were distributed in a wider range than the Bi-high children and monolingual children.

Figure 1. Histograms with kernel density estimation showing the distribution of long-lag VOTs for Mandarin aspirated stops and English voiceless stops in each group of speakers.

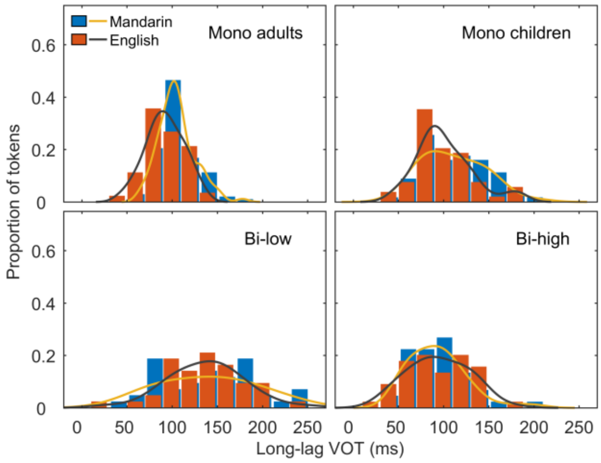

Figure 2 presents the mean and standard error of the long-lag VOTs collapsed across the three places and vowel contexts for each language in each group of speakers. Although the long-lag VOTs in both monolingual Mandarin adults and children were longer than those in monolingual English adults and children, a significant language difference was only yielded between Mandarin adults and English adults (F=4.673, p=0.043). With regard to the two groups of bilingual children, neither Bi-low nor Bi-high children demonstrated evident difference of long-lag VOTs between Mandarin and English. No significant language difference was found for them. Interestingly, although none of the bilingual groups showed significant language difference, the Bi-low children produced both Mandarin and English long-lag VOTs longer than the monolingual speakers while the Bi-high children produced both Mandarin and English long-lag VOTs shorter than the monolingual speakers.

Group difference of long-lag VOTs for each language was also examined. Averaged across the places of articulation and vowel contexts, the average VOT value for Mandarin aspirated stops was 140 ms in the Bi-low children, 93 ms in the Bi-high children, 110 ms in the monolingual Mandarin children and 105 ms in the monolingual Mandarin adults. The average VOT value for English voiceless stops was 136 ms in the Bi-low children, 94 ms in the Bi-high children, 101 ms in the monolingual English children and 92 ms in the monolingual English adults. The Linear Mixed-Effects modelling yielded a significant group difference for both Mandarin aspirated long-lag VOTs (F=3.277, p=0.034) and English voiceless long-lag VOTs (F=8.717, p=0.001). The results of the estimates for fixed effects with the group of monolingual adults as the reference level revealed a significant difference between the Bi-low children and monolingual Mandarin adults as well as between the Bi-low children and monolingual English adults. No significant difference was found between monolingual children or Bi-high children and the reference group of monolingual adults for Mandarin or English.

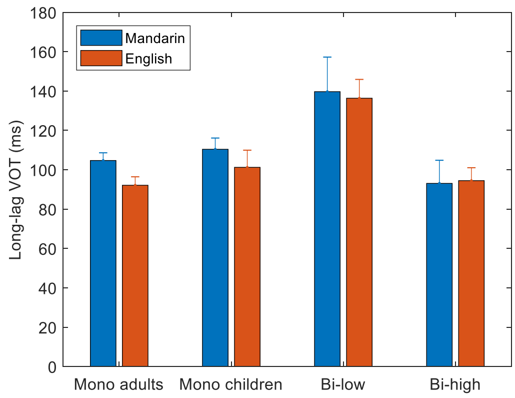

Figure 2. Comparison of the means and standard errors of long-lag VOTs collapsed across place and vowel between Mandarin and English for each group of speakers.

4. DISCUSSION

The present study extended our earlier research of cross-language comparison between Mandarin and English VOTs to a younger age group of 3- to 4-year olds. Given that the VOT difference between Mandarin and English is mainly reflected in long-lag VOTs in adult speakers and the long-lag VOTs show later acquisition and continuing development in children after 3 years of age in comparison to the short-lag VOTs, the present study focused on the long-lag VOTs of Mandarin aspirated stops and English voiceless stops.

Unlike monolingual adults and the older monolingual children reported in our previous study, the 3- to 4-year-old monolingual children did not show significant difference in the long-lag VOTs between Mandarin and English. However, the trend of longer long-lag VOTs in monolingual Mandarin children than in monolingual English children can be observed. Previous research showed that long-lag VOTs develop at a later age than short-lag VOTs. The long-lag VOTs develop from a widespread distributional pattern to a gradually concentrated pattern to approximate the adult targets [14, 15, 17]. Although 5- and 6-year-olds demonstrated a less concentrated VOT distribution for the long-lag VOTs than the adults, they had approximated the adults' targets and showed significant language difference between Mandarin and English long-lag VOTs. By contrast, for the younger children, they were still refining the production of long-lag VOTs and had not acquired the adult-like pattern. Therefore, the subtle difference between Mandarin and English long-lag VOTs found in adults might not be well shown due to the continuing development of long-lag VOT in young children.

Unlike monolingual children who possess only one phonetic system, bilingual children need to manipulate two phonetic systems. The young bilingual children tested in the present study, regardless of the amount of experience in English, could hardly separate the long-lag VOTs between these two languages. These findings were inconsistent with previous research on the development of language-specific VOT patterns in bilingual children [5-9, 21, 22]. These studies reported that bilingual children, at a young age, successfully acquired distinct VOT patterns for each language, even though not always in an adult-like manner. It is noteworthy that the target languages investigated in these studies occupy different VOT ranges along the continuum. For example, in Kehoe et al., [7] study of German-Spanish bilinguals, German stops are characterized by short-lag vs. long-lag contrast while Spanish stops exhibit voicing-lead vs. short-lag distinction. In Simon [8] study of Dutch-English bilingual child, Dutch has a voicing lead vs. short-lag VOT distinction but English has a short-lag vs. long-lag distinction. The two target languages in the present study are both realized as short-lag vs. long-lag distinction with slight difference shown on the long-lag VOTs by adult speakers. As shown in the present study, the younger bilingual children including both experienced and inexperienced bilingual children demonstrated almost completely overlapped Mandarin and English long-lag VOTs. However, the two groups of bilingual children still differed from each other in that the Bi-low children demonstrated widespread distributions of long-lag VOTs for both L1 and L2 than the Bi-high children and monolingual speakers. In addition, the Bi-low children produced both L1 and L2 long-lag VOTs longer than the Bi-high children and other groups of monolingual speakers did.

The comparison of the four groups of speakers for each language revealed that the Bi-low children tended to produce English long-lag VOTs longer than monolingual English children and adults. Note that English long-lag VOTs in adult norm are normally shorter than Mandarin long-lag VOTs. The longer English long-lag VOTs in the Bi-low children than the monolingual speakers and Bi-high children likely showed the assimilation of English VOTs toward Mandarin representations. Interestingly, the Bi-low children not just differed from other three groups on the English long-lag VOTs, they also produced Mandarin long-lag VOTs longer than the other groups, similar to the Bi-low children at a relatively older age reported in our former study [4]. Additionally, the modification of Mandarin long-lag VOTs seemed to occur with the same direction of increased VOT values for the English long-lag VOTs and deviated from the target Mandarin VOT values produced by monolingual Mandarin speakers. These results suggested that young bilingual children experienced great change in their phonetic systems during the early stage of L2 learning.

Different from the Bi-low children, the Bi-high children had a longer period of English learning and a greater amount of exposure to English. They have acquired English long-lag VOTs as monolingual English speakers and showed no significant difference from the monolingual speakers. Meanwhile, the Bi-high children's Mandarin long-lag VOTs were shorter than the monolingual Mandarin children and adults although the difference between the Bi-high children and monolingual Mandarin speakers did not reach the statistical significance. This result suggested the tendency of approximation of Mandarin long-lag VOTs towards the English long-lag VOTs in the Bi-high children.

5. CONCLUSION

In sum, the present study revealed that the subtle difference between Mandarin and English long-lag VOTs was not present in young bilingual Mandarin-English children, regardless of the amount of experience in English. This might be accounted for by the combined effect of the continuing development of long-lag VOTs in children at relatively young age and L1-L2 interactions. It is noteworthy that the small number of participants in the present study compromises the statistical power of the outcomes and generalization of the findings. A future study with more participants is warranted.

7. REFERENCES

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

[2] Chao, K. Y., Chen, L. M. 2008. A cross-linguistic study of voice onset time in stop consonant productions. *Computational Linguistics and Chinese Language Processing 13*(2), 215-232.

[3] Chen, L. M., Chao, K. Y., Peng, J. F. 2007. VOT productions of word-initial stops in Mandarin and English: A cross-language study. *In Proceedings of the 19th Conference on Computational Linguistics and Speech Processing (*ROCLING, 2007).

[4] Yang, J. (under review). Comparison of VOTs in Mandarin-English bilingual children and corresponding monolingual children and adults.

[5] Deuchar, M., Clark, A, 1996. Early bilingual acquisition of the voicing contrast in English and Spanish. *Journal of Phonetics* 24(3), 351-365.

[6] Khattab, G. 2002. VOT production in English and Arabic bilingual and monolingual children. *Amsterdam Studies in the Theory and History of Linguistic Science Series 4*, 1-38.

[7] Kehoe, M.M., Lleó, C., Rakow, M. 2004. Voice onset time in bilingual German-Spanish children. *Bilingualism: Language and Cognition 7*(1), 71-88.

[8]. Simon, E. 2010. Child L2 development: A longitudinal case study on Voice Onset Times in word-initial stops. *Journal of Child Language* 37(1), 159-173.

[9] Fabiano-Smith, L., Bunta, F. 2012. Voice onset time of voiceless bilabial and velar stops in 3-year-old bilingual children and their age-matched monolingual peers. *Clinical Linguistics & Phonetics* 26(2), 148-163.

[10] Bond, Z. S., Wilson, H. F. 1980. Acquisition of the voicing contrast by language-delayed and normal-speaking children. *Journal of Speech and Hearing Research* 23, 152-161.

[11] Engstrand, O., Williams, K. 1996. VOT in stop inventories and in young children's vocalizations: preliminary analyses. *Speech, Music, and Hearing Quarterly Progress and Status Report* 37(2), 97-99.

[12] Gilbert, J. H. V. 1977. A voice onset time analysis of apical stop production in 3-year-olds. *Journal of Child Language 4*, 103–113.

[13] Hitchcock, E. R., Koenig, L. L. 2015. Longitudinal observations of typical English voicing acquisition in a 2-year-old child: Stability of the contrast and considerations for clinical assessment. *Clinical Linguistics & Phonetics* 29(12), 955-976.

[14] Lowenstein, J. H., Nittrouer, S. 2008. Patterns of acquisition of native voice onset time in English-learning children. *The Journal of the Acoustical Society of America* 124(2), 1180-1191.

[15] Macken, M. A., Barton, D. 1980. The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language* 7, 41–74.

[16] Whiteside, S. P., Marshall, J. 2001. Developmental trends in voice onset time: Some evidence for sex differences. *Phonetica* 58(3), 196-210.

[17] Zlatin, M. A., Koenigsknecht, R. A. 1976. Development of the voicing contrast: A comparison of voice onset time in stop perception and production. *Journal of Speech and Hearing Research* 19, 93–111.

[18] Yang, J. 2018. Development of stop consonants in 3-to-6-year-old Mandarin-speaking children. *Journal of Child Language.* doi:10.1017/S0305000918000090.

[19] Edwards, J., Beckman, M. E. 2008. Methodological questions in studying consonant acquisition. *Clinical Linguistics & Phonetics* 22(12), 939-958.

[20] Cho, T., Ladefoged, P. 1999. Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics* 27(2), 207-229.

[21] Johnson, C., Lancaster, P. 1998 The development of more than one phonology: A case study of a Norwegian-English bilingual child. *International Journal of Bilingualism 2*, 265– 300.

[22] Lee, S. A. S., Iverson, G. K. 2012. Stop consonant productions of Korean–English bilingual children. *Bilingualism: Language and Cognition* 15(2), 275-287.