An ultrasound study of English vowel productions by newly-arriving Mandarin-speaking young children

Boram Kim [1,2] *Megan Keough* [2] *Yung-hsiang Shawn Chang* [3]

[1] CUNY Graduate Center (US) [2] Haskins laboratories (US) [3] National Taipei University of Technology (Taiwan)
bkim@graduatecenter.cuny.edu, megan.keough@yale.edu, shawnchang@mail.ntut.edu.tw

Background: Previously studies (e.g., Baker & Trofimovich, 2005; Guion, 2003; Oh et al., 2011) have shown that early bilinguals are better at separating the second language (L2) categories from the first language (L1) than late bilinguals given the same amount of experience. However, these studies have all been acoustic investigations and it remains unknown whether a similar account from articulatory measurements can be observed. The current study is based on an ongoing project of constructing a longitudinal corpus of English vowels produced by L1 and L2 children aged 4 to 6 and adults. In particular, the L2 speakers are Mandarin-speaking children and adults who have just arrived in the US within the previous four months. The same participants will be recorded again after one year of residence in the US. Here we present the productions of two tense-lax vowel pairs /i-ɪ/ and /u-ʊ/ in English as well as the /i/ and /u/ in Mandarin. Mandarin has no tense-lax vowel distinction. Therefore, following Best's (e.g., Best & Tyler, 2007) Perceptual Assimilation Model (PAM), we expect our Mandarin children participants, who have limited prior experience with English, to assimilate /i-ɪ/ and /u-ʊ/ into /i/ and /u/ respectively (i.e., single category assimilation) both in the articulatory and acoustic domains. On the other hand, we expect our adult participants, who have varied experience with English as Foreign Language (EFL) learners, to distinguish the /i-ɪ/ and /u-ʊ/ to various degrees.

Methods: The participants were native Mandarin children (1F 5;5, 1M: 4;5) and adults (2F, 1M) and age-matched native English child (1M 6;3) and adult (1F). An ultrasound system of a Philips EPIQ7G was used to collect real-time images of the tongue surface. A 2D transducer (C8-5) was positioned under the participant's chin with a plastic transducer holder (Derrick et al., 2018). The transducer was adjusted to view the entire tongue surface between the shadows of the jaw and the hyoid bone. The time synchronous audio signal was digitally recorded at 44,100 Hz, 16-bit resolution, using a directional Sennheiser microphone. Stimuli consisted of English and Mandarin CV(C) words that occur frequently and were likely to be known by children participants. The pictures of the English words were presented three times. Using Praat software (Boersma & Weenink, 2017), the start and the end of target vowels were demarcated based on their waveform and the spectrogram. The first two formants (F1 and F2) were extracted from the mid-point of each vowel and normalized using z-score transformation. Using the vowel mid-point time information from the acoustic data as a reference, the tongue image was selected and contoured using GetContours (Tiede, 2016).

Results and discussion: As expected, we found that the tense-lax vowel pairs were merged for our L2 Mandarin children speakers in the acoustic domain. For the front vowel pair /i-ɪ/, the L1 English children showed distinct categories for the tense and lax vowels, while the L2 children showed considerable overlap among the English vowel pair and the Mandarin /i/. In addition, the Mandarin children did not differentiate the English back vowel pair clearly. Interestingly, however, the native English children *also* showed a great deal of overlap for the back vowels. This contrasts with the clear separation of those vowels for the L1 adult speakers. As predicted, the L1 adult speakers showed two clear categories for both tense and lax pairs. On the other hand, the L2 adult participants showed a distinction between the front tense and lax vowels, but not the back vowel pair (i.e., single category assimilation). When we compared the tongue gestures, we found that the L2 child participants appeared to use tongue height but not frontedness to distinguish /i/ and /ɪ/ (Fig. 1c), while the L1 child manipulated both characteristics (Fig. 1a). We also found that both L1 and L2 children produced /u/ vowels that were fronted compared to their /ʊ/ vowel, both in the acoustic and the articulatory domains (Fig.1b and 1d). This is in some sense expected for the L1 child because he lived in Western Canada until age five; /u/-fronting is a known characteristic of Vancouver English (Boberg, 2011). However, it is somewhat surprising that the L2 child participant—whose exposure to English has been concentrated in New England—showed similar acoustic and articulatory results for /u/. One possibility is that, because of their relatively small vocal tract size, the children rely more on lip protrusion when producing /u/. In future work, we could further investigate this finding by incorporating lip gesture tracking into our analysis.

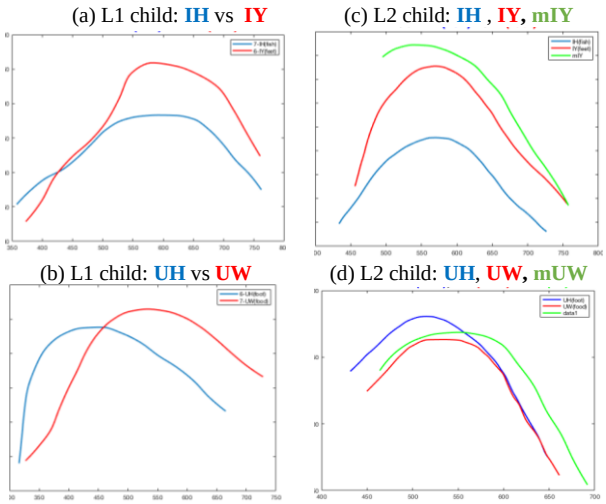

Fig 1. Tongue contours from children participants.
*English /i/(IY) and /u/ (UW) in blue, English /ɪ/ (IH) and /ʊ/ (UH) in red and Mandarin /i/ (mIY) and /u/ (mUW) in green:

References

Baker, W., & Trofimovich, P. (2005). Interaction of native-and second-language vowel system (s) in early and late bilinguals. *Language and speech*,*48*(1), 1-27. **Best, C. T., & Tyler, M. D. (2007).** Nonnative and second-language speech perception: Commonalities and complementarities. Language experience in second language speech learning: In honor of James Emil Flege, 17, 13-34.; **Boberg, C. (2011).** Reshaping the vowel system: An index of phonetic innovation in Canadian English. University of Pennsylvania Working Papers in Linguistics, 17(2), 4. **Boersma, P., & Weenink, D. (2017).** Praat: Doing phonetics by computer [Computer software]. Version 6.0. 33. **Derrick, D., Carignan, C., Chen, W., Shujau, M., & Best, C. T. (2018).** Three-dimensional printable ultrasound transducer stabilization system. The *Journal of the Acoustical Society of America*, 144(5); **Guion, S. G. (2003).** The vowel systems of Quichua-Spanish bilinguals. *Phonetica*,60(2), 98-128. **Oh, G. E., Guion-Anderson, S., Aoyama, K., Flege, J. E., Akahane-Yamada, R., & Yamada, T. (2011).** A one-year longitudinal study of English and Japanese vowel production by Japanese adults and children in an English-speaking setting. *Journal of phonetics*, *39*(2), 156-167.;
Tiede, M. (2016). GetContours.https://github.com/mktiede/GetContours.

Keywords: Speech & language acquisition, L2 Speech production, Longitudinal study