Phonetic Interference of L3 Japanese on L2 English Word-Initial Stop Production in Mandarin Trilinguals

Min Zeng
Waseda University, Japan

Keywords — composite L2-L3 category, L2 production, multilingualism, voice onset time

I. Introduction

While word-initial stops in Mandarin, English and Japanese are primarily cued by voice onset time (VOT) [1, p. 60], [2], [3, p. EL96], word-initial voiceless stops in these languages differ in lag patterns. Word-initial stops can be classified into three categories: voicing lead (−125 to −75 ms), short lag (0 to 25 ms), and long lag (60 to 100 ms) [4, p. 403]. Mandarin [p^h, t^h, k^h] and English /p, t, k/ fall under the long lag category, and are produced by monolinguals with VOT values centering at 100 ms in Mandarin and 80 ms in English [5, p. 20]. In contrast, Japanese /p, t, k/, weakly aspirated, are produced by monolinguals with VOT values that fall between those of prototypical short lag and long lag stops [6, p. 75]. We observe a hierarchy of these voiceless stops: Mandarin [p^h, t^h, k^h] exhibit longer VOT values than those of English /p, t, k/, which in turn exhibit longer VOT values than those of Japanese /p, t, k/.

Previous studies on the production of word-initial stops by native Mandarin-speaking bilinguals and trilinguals have revealed two important findings. First, despite experiencing native language (L1) transfer, Mandarin bilinguals, even those with low proficiency in their second language (L2) English, appear to produce English voiceless stops with VOT values akin to those of native English speakers [7, p. 560]. Second, Mandarin trilinguals experience cross-linguistic phonetic interactions [8, p. 101]. To the best of our knowledge, no study has focused on the acquisition of L2 stops by Mandarin trilinguals who are advanced in both their L2 and L3. This study aims to fill this gap. Our research question is, how do Mandarin trilinguals, who are advanced in both their L2 English and L3 Japanese, produce English word-initial voiceless stops? We hypothesize that the performance of the Mandarin trilinguals will reflect the phonetic interference from both their L1 Mandarin and the later acquired L3 Japanese.

II. Method

Thirty-one Mandarin trilinguals, 34 Mandarin bilinguals and 34 Japanese bilinguals formed the MT, MB and JB groups, respectively, participating in the production experiment. The MT group, consisting of international students at a Japanese university, had an average 3.69-year residency in Japan. Their average age of acquisition was 6.9 years for English and 18.39 years for Japanese. Regarding language proficiency, they scored higher than 85 on a TOEFL iBT test (CEFR B2-C1 level) and passed the Japanese Language Proficiency Test N1 level. The MB and JB groups, university students from Shanghai and Tokyo, respectively, both had L2 English levels ranging from beginner to intermediate (CEFR A1-B2 level). The production material included nine English words (i.e., *panda*, *Paris*, *parrot*, *taxi*, *tablet*, *tango*, *candy*, *camel*, and *carrot*), where the target stops /p, t, k/ occur at the onsets of their first syllables. The participants produced these words using a carrier sentence "The target word is ___."

In total, we collected 2673 valid productions (99 participants * 9 words * 3 repetitions). Using Praat to segment the VOT values of the target stops, we selected the zero-crossing points and relied primarily on the waveforms and secondarily on the vowel formants to pinpoint the boundary between the stop and its following vowel. One linear mixed model (LMM) was applied to these VOT values (the dependent variable) in R. The fixed factor of the LMM was Group (three levels: MT, MB, and JB). The random intercepts were Participant and Stimulus. We assessed the main effects of the fixed factor and performed the post-hoc comparison of contrasts.

III. Results

The descriptive statistics are illustrated with boxplots (Fig. 1). The average VOT values of each English stop produced by the MT, JB, and MB groups are as follows: /p/-64.0, /t/-65.3, /k/-74.2 ms by the MT group; /p/-49.5, /t/-52.3, /k/-68.1 ms by the JB group; /p/-75.0, /t/-75.2, /k/-83.4 ms by the MB group. Regarding statistical analysis, the LMM results indicated significant main effects on Group (χ^2 (2) = 35.7, $p < 0.001$), with a large effect size ($\eta^2 = 0.17$). The post-hoc analysis, as detailed in Table 1, revealed a nuanced hierarchy: the Mandarin bilinguals produced L2 English word-initial voiceless stops with significantly longer VOT values than those of the Mandarin trilinguals, which in turn were longer than those of the Japanese bilinguals.

IV. Discussion and conclusion

First, the VOT values of the English /p, t, k/ produced by the Mandarin trilinguals were significantly longer than those of the Japanese bilinguals, suggesting that the trilinguals were influenced by the phonetic system of their L1 Mandarin. Native Mandarin speakers produce the Mandarin [p^h, t^h, k^h] with significantly longer VOT values ([p^h]-104.9, [t^h]-104.4, [k^h]-103.1 ms) [9, p. 772] than those of the Japanese /p, t, k/ produced by native Japanese speakers (/p/-22, /t/-28, /k/-47 ms) [10, p. 70]. Carrying over the Mandarin production patterns into the production of L2 English stops, the Mandarin trilinguals exhibited significantly longer VOT values compared to those of the Japanese bilinguals.

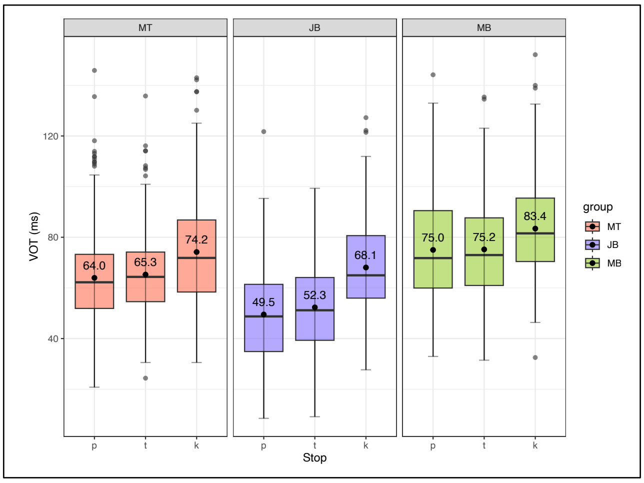

Fig. 1. VOT distributions of the English stops produced by the MT, JB, and MB groups.

TABLE I. SUMMARY OF THE RESULTS OF THE POST-HOC TEST COMPARING THE MT, JB, AND MB GROUPS

Group	Estimate	*SE*	*df*	*t* ratio	*p* value
MT vs. JB	0.53	0.16	96	3.14	0.0067*
MT vs. MB	−0.45	0.16	96	−2.69	0.0249*
JB vs. MB	−0.98	0.16	96	−5.97	< 0.001*

Notably, the Mandarin trilinguals produced the English /p, t, k/ with significantly shorter VOT values than the Mandarin bilinguals, indicating that the trilinguals experienced phonetic interference from their L3 Japanese. Based on the category formation process proposed by the SLM-r [11, pp. 40-41], we speculate that the trilinguals, with extensive exposure to Japanese, might have discovered the phonetic differences between the voiceless stops in L3 Japanese and L1 Mandarin. Specifically, as the trilinguals modified their realization rules to attune to the new L3 stops they had heard and seen in meaningful conversations, they might have noticed that these stops carry phonetic features, such as the articulatory features of moderately aspirated voiceless stops, resembling their corresponding stops in L2 English, which are markedly different from those in L1 Mandarin. Consequently, the trilinguals experienced the L2 English voiceless stops shifting toward the corresponding stops in L3 Japanese, establishing a composite L2-L3 articulatory category that maintained phonetic contrast with the articulatory category for stops in L1 Mandarin. Therefore, the VOT values of the English voiceless stops produced by the Mandarin trilinguals, 'compromise' VOT values carrying the articulatory features of stops in Japanese, were significantly shorter than those produced by the Mandarin bilinguals.

In summary, the results supported our hypothesis: the performance of the Mandarin trilinguals in L2 English revealed that they experienced phonetic interference from their L1 Mandarin and, more importantly, from their later-acquired L3 Japanese.

REFERENCES

[1] H.-G. Byun, "Acoustic characteristics for Japanese stops in word-initial position: VOT and post-stop fo," *Journal of the Phonetic Society of Japan*, vol. 25, pp. 41–63, 2021.

[2] R. Y.-H. Lo, "The dual role of post-stop fundamental frequency in the production and perception of stops in Mandarin-English bilinguals," *Frontiers in Communication*, vol. 7, p. 864 127, 2022.

[3] A. A. Shultz, A. L. Francis, and F. Llanos, "Differential cue weighting in perception and production of consonant voicing," *The Journal of the Acoustical Society of America*, vol. 132, no. 2, EL95–EL101, 2012.

[4] L. Lisker and A. S. Abramson, "A cross-language study of voicing in initial stops: Acoustical measurements," *Word*, vol. 20, no. 3, pp. 384–422, 1964.

[5] J. Yang, "Comparison of VOTs in Mandarin–English bilingual children and corresponding monolingual children and adults," *Second Language Research*, vol. 37, no. 1, pp. 3–26, 2021.

[6] T. J. Vance, *The sounds of Japanese*. Cambridge University Press, 2008.

[7] L.-m. Chen, K.-Y. Chao, J.-F. Peng, and J.-C. Yang, "A cross-language study of stop aspiration: English and Mandarin Chinese," in *2008 Tenth IEEE International Symposium on Multimedia*, 2008, pp. 556–561.

[8] Y. Sun and S. Profita, "Cross-linguistic study on VOT of Chinese trilingual speakers," *Studies in Literature and Language*, vol. 20, no. 1, pp. 98–102, 2020.

[9] Q. Feng and M. G. Busà, "Acquiring Italian stop consonants: A challenge for Mandarin Chinese-speaking learners," *Second Language Research*, vol. 39, no. 3, pp. 759–783, 2023.

[10] K. Shimizu, " 閉鎖子音の有声性・無声性の音声的特徴に関する考察 [A study on phonetic characteristics of voicing contrasts of stop consonants]," *Journal of the Phonetic Society of Japan*, vol. 22, no. 2, pp. 69–80, 2018.

[11] J. E. Flege and O.-S. Bohn, "The revised speech learning model (SLM-r)," in *Second language speech learning: Theoretical and empirical progress*, R. Wayland, Ed., Cambridge University Press, 2021, ch. 1, pp. 3–83.