

Seeing written forms influences lab-based phonetic convergence

Daiki Hashimoto¹, Reiko Asada¹, Keigo Tatusya¹

¹Joetsu University of Education (Japan)

Many experimental studies have explored vocal imitation in a laboratory setting. These studies demonstrate that vocal imitation occurs even though an interlocutor is physically absent in speech. Goldinger [1] is known as a seminal work in this line of research. In his experiment, participants first read aloud a list of words in the beginning, the recordings of which are treated as baseline tokens. Then, the participants shadowed the words quickly after presentation of model speech stimuli. The baseline and shadowed tokens are compared with the model speech stimuli. His finding is that shadowed tokens are more similar to model speech stimuli than baseline tokens are. This finding suggests that the participants imitate the auditory speech stimuli while shadowing the speech stimuli without interacting with an interlocutor. His experimental paradigm has been adopted by a large number of studies so far, and his finding has been successfully replicated.

Although many previous studies employ a shadowing experiment to explore vocal imitation, the methodologies slightly differ from researcher to researcher. We would like to attract attention to the fact that shadowers saw orthographic forms in some experiments [2, 3, 4], whereas shadowers did not see any written forms in the other experiments [5, 6, 7, 8]. The aim of this study is to explore to what extent seeing written forms influences L2 learners' shadowing of model speech. Addressing this question provides an insight to a methodology that can be used to study vocal imitation or accommodation, and also allows us to increase our knowledge of how pronunciation and orthography are memorized in the cognitive system of a L2 learner.

In order to address this question, we explored to what extent Japanese English learners align their formant values with those of a native American English speaker, when they shadow the model speech stimuli with /æ/ or /ɑ/ while seeing written forms and without seeing written forms. Our prediction is that seeing written forms may interfere in L2 learners' imitation of model speech. This is because presentation of orthographic forms may activate the knowledge about grapheme-to-phoneme correspondences in L1 [9, 10]. Note that the two vowels are likely to be spelled as <a> and <o> respectively in English, and the graphemes are associated with the phonemes /a/ and /o/ in Japanese.

Twenty-two Japanese English learners participated in the current study. They first read aloud 42 target words (21 words with /æ/ and 21 words with /ɑ/) while seeing written forms on the computer monitor. They were also asked to read aloud nine keywords, the recordings of which were used to standardize the vowel space. Then, half the participants shadowed model speech stimuli of the target words while seeing written forms in the first shadowing session, and shadowed the model speech stimuli without seeing written forms in the second shadowing session. The other participants took the two shadowing sessions in the reverse order. The recordings were manually annotated, and acoustically analysed. F1 and F2 values were extracted at mid-point using a Praat script [11], and they were Labanov-transformed.

The results are shown in Figure 1. The figure shows the mean formant values of TRAP /æ/ and LOT /ɑ/ produced by a model talker (red), and by a participant in the pre-shadowing session (purple), the shadowing session with written forms (blue), and the shadowing session without written forms (green). The mixed-effects regression analyses were performed on our dataset. We constructed four mixed-effects linear regression models that estimate a F1 value of TRAP, a F2 value of TRAP, a F1 value of LOT, and a F2 value of LOT respectively, using the *lmer* function in the *lme4* library [12]. These models include the binary variable *task* (non-orthography shadowing session vs. orthography shadowing session), the binary variable *group* (orthography-shadowing-first group vs. non-orthography-shadowing-first group), the interaction term of the two key variables, and by-speaker and by-word random intercepts. The statistical results indicate that the formant values are significantly different between the orthography shadowing session and the non-orthography shadowing session. For TRAP vowels, F1 values are significantly lower in an orthography shadowing session than in a non-orthography shadowing session. As for LOT vowels, F1 and F2 values are significantly lower in an orthography shadowing session than in a non-orthography shadowing session.

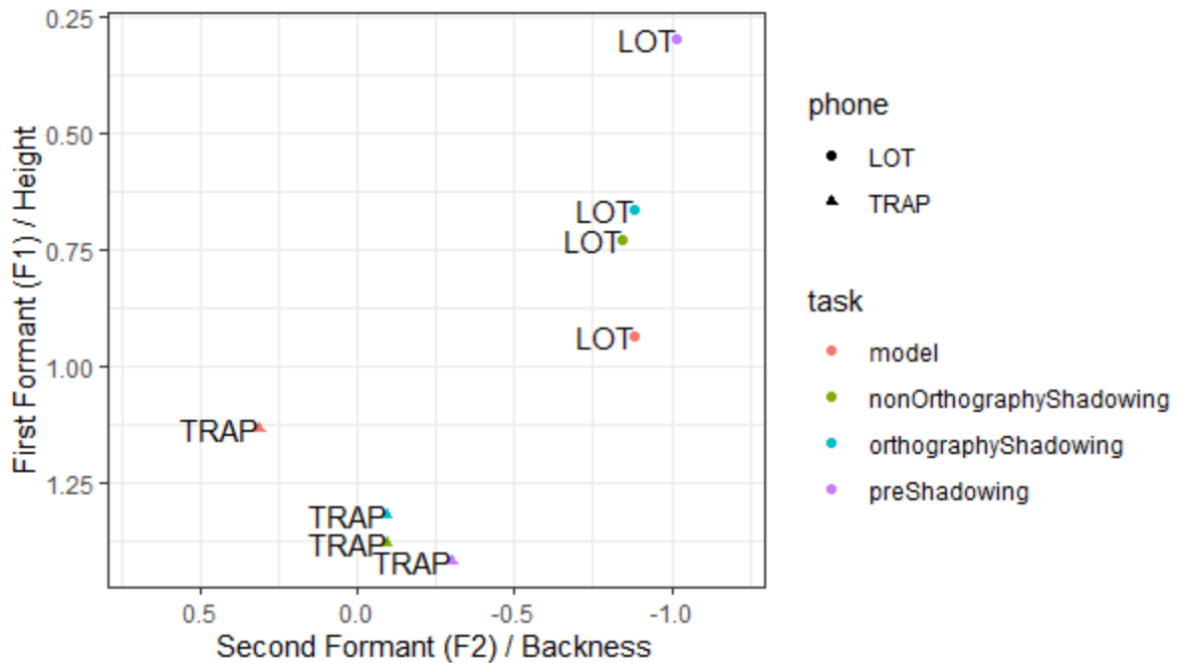


Fig. 1. Mean Labanov-transformed formant values of LOT /a/ and TRAP /æ/ of model speech stimulus (red), pre-shadowing token (purple), non-orthography-shadowing token (green), and orthography-shadowing token (blue)

References

- [1] Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105(2), 251–279.
- [2] Walker, A., & Campbell-Kibler, K. (2015). Repeat what after whom? Exploring variable selectivity in a cross-dialectal shadowing task. *Frontiers in Psychology*, 6: 546.
- [3] Gnevshva, K., Szakay, A., & Jansen, S. (2021). Phonetic convergence across dialect boundaries in first and second language speakers. *Journal of Phonetics*, 89: 101110.
- [4] Hashimoto, D., Gomi, K. & Shiraishi, R. (2022). Updating L2 phonetic memories: Shadowing training vs. listening training. *Journal of Phonetic Society of Japan* 26. 13–26.
- [5] Shockley, K., Sabadini, L. & Fowler, C. A. (2004). Imitation in shadowing words. *Perception and Psychophysics*, 66(3), 422–429.
- [6] Babel, M. (2009). *Phonetic and social selectivity in speech accommodation*. Berkeley: University of California, Berkeley dissertation.
- [7] Dufour, S., & Nguyen, N. (2013). How much imitation is there in a shadowing task? *Frontiers in Psychology*, 4: 346.
- [8] Pardo, J. S., Urmanche, A., Wilman, S., & Wiener, J. (2017). Phonetic convergence across multiple measures and model talkers. *Attention, Perception, & Psychophysics*, 79, 637–659.
- [9] Rafat, Y. (2015). The interaction of acoustic and orthographic input in the acquisition of Spanish assibilated/fricative rhotics. *Applied Psycholinguistics*, 36(1), 43–66.
- [10] Bürki, A., Welby, P., Clément, M., & Spinelli, E. (2019). Orthography and second language word learning: Moving beyond “friend or foe?”. *The Journal of the Acoustical Society of America*, 145(4): EL265-EL271.
- [11] Boersma, P., & Weenink, D. (2020). Praat version 6.1.14. <http://www.fon.hum.uva.nl/praat/> (accessed May 28, 2020)
- [12] Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–38.