Training the pronunciation of L2 novel phonetic features: a comparison of observing versus producing hand gestures

Xiaotong Xi[1], *Peng Li*[1], *Florence Baills*[1], *Pilar Prieto*[2, 1]

[1] Universitat Pompeu Fabra, Barcelona, Catalonia, Spain
[2] Institució Catalana de Recerca i Estudis Avançats, Barcelona, Catalonia, Spain
xiaotong.xi@upf.edu, peng.li@upf.edu, florence.baills@upf.edu, pilar.prieto@upf.edu

Abstract

This paper compares the results of two complementary experiments exploring the effects of training with gesture observation (Experiment 1) and gesture production (Experiment 2) on L2 pronunciation learning. Importantly, this study controls for the appropriateness of the gestures performed by learners during the production training and its effect on learning outcomes. A total of 106 Catalan naïve learners were trained in the learning of Mandarin aspirated plosives by being asked to observe (Experiment 1) or to produce (Experiment 2) a fist-to-open-hand gesture encoding the aspiration feature. Results revealed that participants who observed the hand gestures during phonetic training improved their pronunciation of the aspiration feature significantly more than those who produced the hand gesture during training. However, when the accuracy of gestures performed by participants during training was taken into account, the pronunciation improvement was similar between those who performed gestures well and those who were trained by merely observing hand gestures. These results suggest that observing hand gestures facilitates the pronunciation of L2 novel segmental features and that training with producing hand gestures also can be effective, but only for participants who are able to perform the hand gestures appropriately.

Index Terms: gesture observation, gesture production, speech production, aspiration feature

1. Introduction

The potential effects of training by observing and producing hand gestures on L2 pronunciation learning have been examined in many empirical studies. Observing several types of hand gestures showed beneficial effects on the learning of several L2 phonological features, such as beat gestures (up-and-down hand gestures which highlight prosodic prominence) for the pronunciation of rhythmic patterns [1], or pitch gestures (hand gestures mimicking pitch counters in space) for the production of intonational patterns [2] and perception of lexical tones [3]. Gesture production has also been shown to play a beneficial role in L2 pronunciation learning, like beat gestures in the pronunciation of rhythmic patterns [4] and pitch gestures for the perception of lexical tones (e.g., [3]). Research has also compared the effects of observing versus producing gestures on pronunciation learning, with mixed results. One study showed stronger effects of gesture production than gesture observation on the production of rhythmic patterns [4]. Other studies reported that observing vs. producing yielded either equally beneficial effects on the learning of lexical tones [3] or an equal absence of effect on the learning of vowel length contrasts (e.g., [5]). With regard to L2 segmental features, despite the positive effects of observing hand gestures mimicking articulatory information on the learning of such features (e.g., [6]), little is known about the benefits of gesture production as compared to gesture observation. Crucially, to our knowledge, none of the previous studies has assessed the potential effects of the accuracy of gesture performance on L2 pronunciation learning. This study thus compares the effects of observing (Experiment 1) versus producing (Experiment 2) a fist-to-open-hand gesture which encodes the aspiration feature of Mandarin aspirated plosives on the pronunciation of this feature by Catalan naïve learners.

Mandarin Chinese has three pairs of plosives (/p/-/p^h/, /t/-/t^h/, and /k/-/k^h/) contrasting in aspiration. This aspiration feature does not exist in Catalan phonology, and the acquisition of the aspirated plosives is considered to be one of the main difficulties experienced by Romance language speakers when learning Mandarin [7]. Thus, we intended to train Catalan speakers to produce aspirated plosives with the help of a fist-to-open-hand gesture adopted from an observation study of L2 Mandarin classroom in which this gesture was used as a teaching tool to facilitate the learning of the aspiration feature [8].

2. Method

In Experiment 1, 50 Catalan native speakers without prior knowledge of Mandarin Chinese were trained in one of two conditions: (a) in the Speech Observing (SO) condition, they watched two instructors producing words in Mandarin containing the three pairs of plosives, and (b) in the Gesture Observing (GO) condition, they watched the instructors producing the same words while performing a fist-to-open-hand gesture when they pronounced the aspirated consonant. In Experiment 2, a different set of 58 Catalan naïve learners were trained in one of two conditions: (a) the Speech Producing (SP) condition, in which participants watched and imitated the same instructors producing the same words, and (b) the Gesture Producing (GP) condition, where participants imitated both words and hand gestures. In both experiments, participants' pronunciation was tested before and after training through an imitation task in which they listened to and imitated Mandarin words containing the target sounds. Five Mandarin native speakers assessed their production of the aspiration feature on a 9-point Likert scale from 9 'definitely accurate' to 1 'not accurate at all'. The independent analysis of these two experiments showed that training that involved observing hand gestures yielded better pronunciation scores of the aspiration feature compared to training that involved observing speech (Exp. 1), and similarly, training that involved producing hand

gestures outperformed training that involved producing speech (Exp. 2).

Based on these two experiments, we sought to further explore the effects of observing versus producing hand gestures, with a focus on the accuracy of the gestures performed by participants during training. With this goal in mind, we performed two analyses of variance (ANOVA). The first ANOVA included the training conditions of the two experiments (four levels: SO, GO, SP, GP) as the fixed factor. A mean score for the production of the aspiration feature was calculated for each participant at pre- and posttest by adding up the scores for the six items rated by five raters and then dividing by the total number of scores. The participant's mean improvement score was calculated by subtracting the pretest and posttest mean scores, and this was set as the dependent variable in both ANOVAs.

Next, for the 29 participants in the GP condition, accuracy in the performance of hand gestures during training was assessed. To do this, three Mandarin native speakers rated the GP participants' gesture performance on a 9-point Likert scale from 9 'very good' to 1 'very bad' by focusing on the accuracy of gesture form and whether it was performed simultaneously with the target sound. Inter-rater reliability was checked using Cronbach's alpha. The results revealed an excellent level of agreement for gesture accuracy scores ($\alpha = .93$). Based on their gesture accuracy scores, participants in the GP group were then divided by means of a TwoStep Cluster analysis into two subgroups, a Well-Performed Gesture (WPG) group ($n = 14$) and a Poorly-Performed Gesture (PPG) group ($n = 15$). A second ANOVA analysis was then performed with condition (five levels: SO, GO, SP, WPG, PPG) as the fixed factor and improvement score as the dependent variable.

3. Results

Results of the first ANOVA showed a significant effect of condition (F (3,104) = 3.934; $p = .011$) and the post-hoc tests revealed that observing hand gestures improved participants' aspiration pronunciation significantly more than producing hand gestures (Mean Difference = 0.960, $p = .027$), a result which is inconsistent with previous comparison studies (e.g., [3], [4]).

The second ANOVA analysis revealed a significant effect of condition (F (4,103) = 3.148; $p = .017$) and the post-hoc results showed that GO outperformed PGP (Mean Difference = 1.160, $p = .039$), whereas there was no significant difference between GO and WGP (Mean Difference = 0.746, $p = .375$). Figure 1 shows the mean improvement score across the five conditions.

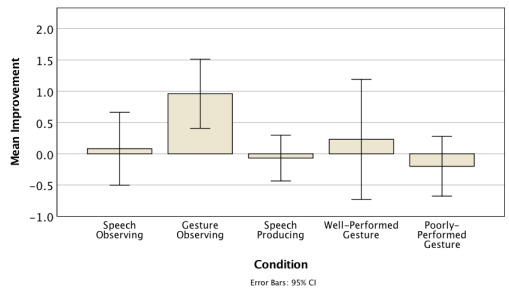

Figure 1: *Mean improvement of aspiration production across condition (five levels: SO, GO, SP, WGP and PGP). Error bars indicate 95% CI.*

4. Discussion and conclusions

The results of this study reveal that, at an early stage of L2 learning, observing gestures mimicking articulatory information has a beneficial effect on the pronunciation of L2 segmental features, which is in line with previous studies (e.g., [6]). As for producing hand gestures, it is only effective when learners are capable of doing so appropriately. Previous study has shown that the efficacy of observing hand gestures on learning L2 novel segments depends on the appropriateness of the gesture form [9]. This study reveals the importance of the appropriateness of the gestures performed by learners, which suggests that both the appropriateness of gesture form and the assessment of gesture performance should be taken into consideration for future studies assessing the effect of producing gesture. From a pedagogical standpoint, teachers who wish to have their students use gestures to support pronunciation learning should assess the potential difficulties of performing gestures, especially for L2 beginners.

5. Acknowledgements

This study was funded by the Spanish Ministry of Science, Innovation and Universities (PGC2018-097007-B-I00) and a Generalitat de Catalunya AGAUR project (2017 SGR-971). The authors sincerely thank Yuan Zhang (Universitat Pompeu Fabra) for her participation in the rating session.

6. References

[1] D. Gluhareva and P. Prieto, "Training with rhythmic beat gestures benefits L2 pronunciation in discourse-demanding situations," *Language Teaching Research*, vol. 21, no, 5, pp. 609–631, 2017.

[2] C. Yuan, S. González-Fuente, F. Baills, and P. Prieto, "Observing pitch gestures favors the learning of Spanish intonation by Mandarin speakers," *Studies in Second Language Acquisition*, vol. 41, no. 1, pp. 5–32, 2019.

[3] F. Baills, N. Suárez-González, S. González-Fuente, and P. Prieto, "Observing and producing pitch gestures facilitates the learning of Mandarin Chinese tones and words," *Studies in Second Language Acquisition*, vol. 41, no. 1, pp. 33–58, 2019.

[4] O. Kushch, "Beat gestures and prosodic prominence: impact on learning," Ph.D. dissertation, Dept. Transl. Lang. Sci., Univ. Pompeu Fabra, Barcelona, Catalonia, Spain, 2018.

[5] Y. Hirata, S. D. Kelly, J. Huang, and M. Manansala, "Effects of hand gestures on auditory learning of second-language vowel length contrasts," *Journal of Speech, Language, and Hearing Research*, vol. 57, no. 6, pp. 2090–2101, 2014.

[6] M. Amand and Z. Touhami, "Teaching the pronunciation of sentence final and word boundary stops to French learners of English: Distracted imitation versus audio-visual explanations," *Research in Language*, vol. 14, no. 4, pp. 377–388, 2016.

[7] N. F. Chen, V. Shivakumar, M. Harikumar, B. Ma, and H. Li, "Large-scale characterization of Mandarin pronunciation errors made by native speakers of European languages," in *INTERSPEECH 2013 – 14th Annual Conference of the International Speech Communication Association, Lyon, France, Proceedings*, 2013, pp. 2370 –2374.

[8] Y. Zhang, "The importance of using gestures in pronunciation teaching," *Language Teaching and Linguistics Studies*, vol. 6, pp. 51–56, 2002.

[9] X. Xi, P. Li, F. Baills, and P. Prieto, "Hand gestures facilitate the acquisition of novel phonemic contrasts when they appropriately mimic target phonetic features," *Journal of Speech, Language, and Hearing Research*, to be published.