



Hand-clapping to the rhythm of newly learned words improves L2 pronunciation: Evidence from Catalan and Chinese learners of French

Florence Baills¹, Yuan Zhang¹, Pilar Prieto^{1,2}

¹Department of Translation and Language Sciences, Universitat Pompeu Fabra, Spain

²Institució Catalana de Recerca i Estudis Avançats (ICREA), Spain

florence.baills@upf.edu, yuan.zhang@upf.edu, pilar.prieto@upf.edu

Abstract

Previous research has shown that rhythmic training enhances phonological speech processing (e.g., [1, 2, 3, 4]). Yet little is known about whether rhythmic training can also help to improve pronunciation in a second language (but see [5, 6, 7]). This study tests the potential benefits of hand-clapping to the rhythm of newly learned French words for the acquisition of pronunciation patterns by Catalan children and Chinese adolescents. In two between-subjects experiments with a pretest and posttest design, participants either repeated words while clapping the rhythmic structure of words or only repeated the words. The French target words were very similar to their lexical counterparts in Catalan, whereas they differed completely from Chinese. Participants' oral production before and after training was rated for accentedness by three French native speakers. Results showed a significant improvement for the clapping group among the Catalan participants, whereas only a near significant tendency appeared for the Chinese adolescents. Individual musical abilities did not interact significantly with the results in either experiments but working memory played a significant role in Chinese participants' pronunciation. The results show that a short phonological training session based on highlighting the rhythmic structure of words seems to help improve pronunciation in a foreign language as long as there is little demand placed on semantic processing.

Index Terms: prosodic acquisition, pronunciation learning, rhythmic training

1. Introduction

There is growing evidence that rhythmic training is beneficial for different aspects of language development (e.g., speech segmentation, phonological skills, memorization, reading skills, etc.) in the native language. Priming studies by Cason and collaborators have shown that the phonological processing of speech by adult participants is enhanced by the temporal expectations generated by a musical rhythmic prime [1, 2]. Regarding production skills, Cason et al. [4] showed that the same priming effect benefits hearing-impaired children with cochlear implants. Similarly, Bhide et al. [8] found that a two-month program of phonological and rhythmic training helped children who were considered poor readers to improve their literacy skills. Rhythmic training coupled with singing has also proved useful for helping children at risk of dyslexia to increase their pre-reading skills [9].

However, studies on rhythm training applied to foreign language acquisition are scarce. To our knowledge, only three studies from very different domains of research have been conducted on the effect of rhythmic training on L2 pronunciation. In a qualitative study involving six advanced

learners of English from various language backgrounds, Fischler [5] explored the effects of a four-week training program on the learning of English sentence and word stress through activities related to rhythm and rap music. She tested participants before and after training on a reading task and a narrative-picture task and counted the errors they made in stress location. Additionally, their oral productions were rated by three English native speakers for intelligibility. Although not statistically backed up, an improvement in both intelligibility and stress placement was reported for the reading task only. Wang, Mok, and Meng [6] tested a computer application that automatically generated a percussive rhythm for any given text in English. 20 Chinese learners of English were asked first to pronounce 15 English sentences naturally, then introduce the same sentences in the interface, listen to the rhythm that was automatically generated and say the sentences out loud immediately afterwards. Sentences pronounced after a rhythmic cue obtained better ratings in terms of native-likeness, but only for participants that performed worse in the natural condition. Finally, Gluhareva and Prieto [7] found positive effects of a short training session that involved observing rhythmic beat gestures—simple up-and-down or back-and-forth hand movements naturally coordinated with the prominent parts of speech—on L2 English pronunciation by Catalan learners. At posttest, participants that had observed the training with beat gestures significantly improved their accentedness ratings on a set of difficult items related to the training content.

The three studies just described employed different types of procedures none of which would be easily applicable to teaching suprasegmental aspects of pronunciation in the language classroom. Perhaps surprisingly, the possible facilitating effect of hand-clapping to the rhythm of the target L2 prosody has not yet been investigated. Whereas in [7], beat gestures were used to highlight higher levels of prosodic structure by marking nuclear pitch accents, hand-clapping lends itself to visually marking prosodic properties at the syllabic level, namely duration and intensity.

The aim of the present study is to assess whether a short training session using clapping to highlight the structure of words can improve pronunciation in a second language. There is growing evidence that human perception is based on cross-modal interactions and that sensory modalities influence each other, especially in speech [10]. Our hypothesis is that clapping, an activity that is strongly linked to rhythm, can facilitate the auditory perception of new words, increase phonological awareness, and ultimately lead to better pronunciation. Just as beat gestures that accompany L2 speech help speakers externalize the prosodic features of a foreign language [11], we hypothesize that hand-clapping will make more salient the metrical structure of words by acoustically highlighting it with different intensity and duration patterns. In

the current study, we tested the effect of rhythmic hand movements on the pronunciation of French words by Catalan children (Experiment 1) and Chinese adolescents (Experiment 2). Our choice of L1 populations was based on the respective degree of difficulty they presumably faced in learning the prosodic and lexical phonological target language. First, while Catalan is a stress-based language like French (though in French the position of stress is generally associated with the final syllable of the word), Chinese is a tonal language. Thus French stress placement is presumably more difficult to learn for Chinese than Catalan speakers. Second, while the meaning of the new French words taught in the experiment was transparent for the Catalan-speaking participants because the words were ‘cognates’ to words in Catalan, this was not the case for the Chinese-speaking participants.

In Experiment 1, Catalan children learned twenty new cognate French words under one of two audiovisual conditions: 1) training which involved observing and replicating the behavior of a native speaker simultaneously saying a word and clapping to highlight prosody, or 2) training which involved observing and replicating a native speaker who merely spoke the word without clapping. In Experiment 2, Chinese adolescents learned the same twenty French words under the same two audiovisual conditions. We hypothesized that observing and performing hand-clapping would lead to improved pronunciation of the French words by both groups.

2. Experiment 1

2.1 Methods

2.1.1 Participants

Thirty 7- and 8-year-old children ($M_{\text{age}} = 7.36$; $SD = 0.479$, 14 females) from the city of Girona, Catalonia, took part in the experiment at their school premises after their parents signed a written consent. They were all Catalan-Spanish bilinguals with Catalan as their dominant language. None of them had any prior knowledge of French. They were informed that they would learn words in French and were randomly divided into two groups of 15 participants each.

2.1.2 Control measures

In order to take into account individual differences, prior to the experiment, a series of control tests were carried out for each participant during a separate session that lasted around 40 minutes. In this session measures were taken of a) short-term memory through a memory span task with lists of Catalan words [12], b) narrative ability through a story retelling task involving a short cartoon [13], c) imitation skills through an elicited imitation task involving 12 words in six different languages, d) perceptual abilities through a phonological discrimination task with nonwords [14], e) perceptual rhythmic abilities through a standard musical rhythm and accent perception test [15], and f) a rhythm hand-clapping replication test with six rhythmic samples (4/4 meter, 2 measures).

2.1.3 Materials

Materials for the training session consisted of 20 videos prepared at the professional broadcasting studio of Universitat Pompeu Fabra. Each video was designed to teach one French word. The rationale for the selection of words was first that

their meaning should be transparent to Catalan speakers (e.g., Cat. *avió*/ Fr. *avion* ‘plane’, Cat. *ordinador*/ Fr. *ordinateur* ‘computer’). Second, the selected words should also include a variety of consonant and vowel sounds in the target language that are not part of the Catalan inventory of sounds. Finally, they needed to have a straightforward visual representation. The training session was preceded by a short video in Catalan explaining the task and a pretest (see below).

Each training video displayed the following sequence: First, a black and white line drawing illustrating the French word appeared on the screen. Three seconds later, a video clip showed one of two female native French speakers saying the word either (1) clapping her hands while saying the word (*clapping condition*), or (2) remaining completely still while saying the word (*non-clapping condition*) (see Figure 1). Finally, a 5-second black screen was left for participants to (1) repeat the word and clap their hands, or (2) just repeat the word, depending on the experimental group to which they had been assigned. The 20 training videos were randomized into six different orders.

For the clapping condition, the two instructors naturally clapped their hands to each word’s syllables to highlight the rhythmic structure. Thus, the claps which occurred on the stressed syllables presented higher intensity and longer duration. At the same time, it was ensured that the sound produced by the hand-clapping did not cover the voices of the instructors.

The pre- and posttests were elicited imitation tasks made up of identical materials. They consisted of a sequence of audio recordings of 14 French words embedded in a PowerPoint file. Ten of these words were taken from the training materials and four were completely new. The ten words from the training materials consisted of audio files extracted from the video recordings. The four new items were recordings featuring one of the same two native speakers who had appeared in the training videos. The sequence of the 14 items was randomized to produce several different presentation orders. Each sequence was preceded by instructions telling the participant to repeat each word that s/he heard and then touch the screen to proceed to the next word.

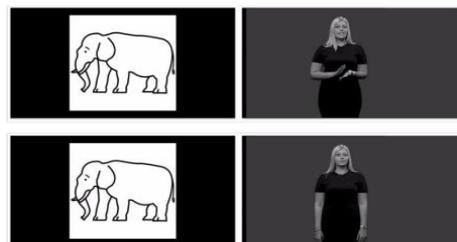


Figure 1: Stills from the training video in the clapping condition (top panel) and in the non-clapping condition (bottom panel).

2.1.4 Procedure

Participating children were tested and trained individually in a quiet room at their school, accompanied by one of the researchers. They were seated in front of a tablet computer and wore comfortable high quality headsets equipped with a high quality microphone.

After listening to the recorded instructions, participants listened to and repeated the 14 French words of the pretest at their own pace by activating each sound file. Then, depending on which experimental group the child had been assigned to,

the researcher (the first author of the paper) launched one or the other training video. The 15 children in the clapping condition group saw the drawing depicting the word, watched the video clip of the speaker producing the word while clapping her hands, and consecutively repeated the word while also clapping their hands (see Figure 1). The 15 children in the non-clapping condition group saw the drawing, watched the speaker producing the word without clapping, and simply repeated the word. Each participant saw two different video sequences, for a total of 40 videos. When the training session finished, the children performed the posttest, which was identical to the version of the pretest that they had done. For pretest, training session, and posttest, the order of presentation of the items was assigned randomly.

The full experimental procedure lasted approximately 25 minutes. The audio recordings of the participants repeating words, with or without clapping constituted the dataset for subsequent analysis. A total of 840 such recordings were obtained, 240 from the pretest (30 participants \times 14 items) and an equal number from the posttest.

2.1.5 Pronunciation assessment

The audio files of the children's oral productions were rated by three non-linguist French native speakers, who were unaware of the purpose of the experiment. The productions were presented in random pretest/posttest pairs together with the native model word. The raters were asked to compare the participants' production with the original auditory stimuli and evaluate the general accentedness of the words (see [16], for a validation of accentedness as a valid evaluation of pronunciation) on a scale from 1 - 'not accented' - to 7 - 'extremely accented'.

2.1.6 Statistical analysis

A high degree of reliability was found between the three raters' judgments (Interclass Correlation Coefficient ICC = .834 with a 95% confidence interval from .813 to .852, $F(837,1674) = 6.018$, $p < .001$). The average of the three raters' scores for each item was used as an index of the variable ACCENTEDNESS for subsequent statistical analysis. To test the effect of the type of training (clapping vs. non-clapping) on participants' accentedness, a General Linear Mixed Model (henceforth GLMM) using IBM SPSS Statistics 23 was run with ACCENTEDNESS as the dependent variable. GROUP (two levels: non-clapping vs. clapping), SESSION (two levels: pretest and posttest), GROUP*SESSION, RELATEDNESS (two levels: trained word vs. new word) and GROUP*SESSION*RELATEDNESS were set as fixed factors. All the control measures described in 2.1.2 were also set as fixed factors. One random effects block was specified, with PARTICIPANT and ITEM intercepts.

2.2. Results

Results showed a significant main effect of SESSION ($F(1, 747) = 28.014$, $p < .001$), GROUP*SESSION ($F(1,747) = 4.184$, $p < .005$), and RELATEDNESS ($F(1,747) = 20.250$, $p < .001$) on accentedness scores. No significant effects were found for any of the control variables. Post-hoc pairwise comparisons revealed that there was a significant difference between the pretest and the posttest both in the non-clapping group ($F(1,747) = 5.238$, $p < .05$) and in the clapping group

($F(1,747) = 27.097$, $p < .001$). Effect sizes calculated from the means and SDs following Cohen's (1988) convention revealed that the non-clapping group showed a small effect size ($d = .20$) whereas the clapping group showed a moderate and larger effect size ($d = .62$). Additionally, the triple interaction GROUP*SESSION*RELATEDNESS indicated that only the related items improved significantly between pre- and posttests for both groups (clapping: $F(1,747) = 84.706$, $p < .001$; non-clapping: $F(1, 747) = 6.888$, $p < .005$).

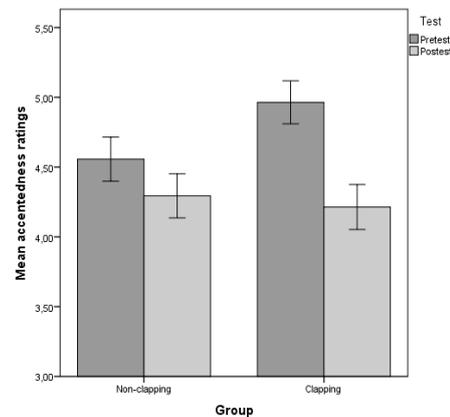


Figure 2: Mean accentedness ratings across the two conditions in Experiment 1, GROUP (non-clapping and clapping) and SESSION (pretest and posttest).

2.3 Discussion

Our results show that a short audiovisual training session in which Catalan children observed and imitated clapping to the rhythm of newly learned cognate French words helped to improve their pronunciation of these words by Catalan children more than training with word repetition only. Clapping to the rhythm of words may have helped to mark syllable structure and the saliency of the final stressed syllables so that participants could better perceive and produce the prosodic features of the target words.

3. Experiment 2

3.1 Methods

3.1.1 Participants

Fifty 13- to 15-year-old Chinese adolescents ($M_{age} = 13.6$; $SD = 0.535$, 16 females) from Shandong Province, China, took part in the experiment at their school premises after their parents signed a written consent. They were all Mandarin Chinese native speakers without any knowledge of French. They were informed that they would learn words in French and were randomly divided into two groups of 25 participants.

3.1.2 Materials, procedure, and statistical analysis.

Materials for the training session were exactly the same materials as in Exp. 1 (see section 2.1.3), the only difference being that the audio files were uploaded onto an online survey platform that automatically randomized the items. The experiment lasted approximately 25 minutes. A total of 1400 experimental responses were obtained (50 participants \times 14 items \times 2, i.e., pre- and posttest). The same pronunciation

assessment as in Exp. 1 was applied to evaluate the oral production of participants in their imitations of French words at pretest and at posttest and the statistical analysis was exactly the same as in Exp. 1 (see sections 2.1.5 and 2.1.6).

3.2. Results

A good degree of reliability was found between the two raters' judgments (ICC = .721 with a 95% confidence interval from .690 to .749, $F(1395,1395) = 3.580, p < .001$). Results showed a significant main effect of SESSION ($F(1, 2484) = 24.920, p < .0001$) and WORKING MEMORY ($F(3, 2484) = 2.976, p < 0.05$) as well as a near significant effect of GROUP*SESSION ($F(1, 2484) = 2.879, p = .090$). Effect sizes for the clapping group was higher ($d = 0.362$) than for the non-clapping group ($d = 0.178$).

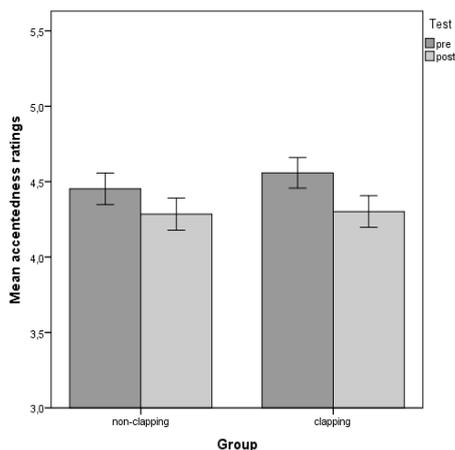


Figure 3: Mean accentedness ratings across the two conditions in Experiment 2, GROUP (non-clapping and clapping) and SESSION (pretest and posttest).

3.3. Discussion

Although the interaction GROUP*SESSION was only near-significant, an analysis of effect sizes revealed that accentedness ratings in the clapping group improved more in the posttest than did those of the non-clapping group.

4. General discussion and conclusions

Rhythmic training sessions have proved useful to enhance the phonological perception of words and sentences [1, 2, 3], to help primary school children acquire reading skills [8], to improve oral production by children with cochlear implants [4], and to improve stress placement [5] and general pronunciation in learners of English [6,7].

In this study, we extend previous findings by showing in Exp. 1 that Catalan children improved their pronunciation of French words significantly more after a short audiovisual training session based on repeating and hand-clapping the prosody of the words than after a training session that involved merely repeating the words. Hand-clapping may have lightened the cognitive demand involved in learning the rhythmic properties of the words, thus allowing participants to focus on other aspects such as difficult phonemes or phonemic combinations. However, this hypothesis warrants more fine-grained exploration through instrumented acoustic analysis..

These results were not replicated by the Chinese adolescents in Exp. 2, where an improvement in pronunciation occurred but did not quite reach significance. A detailed acoustic analysis of the data was carried on the relative duration of the word-final syllable and of the word-final rhyme (Zhang, Baills & Prieto, submitted) and showed that the Chinese adolescents in the clapping condition improved significantly more than participants in the non-clapping condition with respect to the realization of the last stressed syllable of the word (e.g., by lengthening more significantly duration, see [17]).

The pronunciation improvements seen in Exp. 1 were perceptually stronger than those seen in Exp. 2. There are several possible explanations for this difference. First, the age of the participants may have played a role. Since the age of onset of L2 learning is particularly important for the acquisition of L2 phonology (for a review, see [18]), which may not be possible after a sensitive period for neural plasticity has ended (e.g., [19]), adolescent language learners in Exp. 2 may have more difficulty replicating novel phonological targets than younger children. Second, the distance between participants' native languages and the target language, French, must also be taken into account. Whereas the words used for the training session would be semantically transparent for Catalan speakers given that French and Catalan are closely related Romance languages (and indeed the words were selected because of that similarity), Chinese speakers had no such advantage and were required to cope simultaneously with meaning as well as phonetic realization. Their attention during training may thus have been directed more toward learning the meaning of the new vocabulary items rather than learning how to pronounce them. We hypothesize that the cognitive demands placed on semantic processing added a supplementary difficulty to the task, diminishing the impact of training. Interestingly, the main effect of Working Memory found in Exp. 2 GLMM analysis indicated that participants who had better memorization skills were more able to focus on pronunciation whereas participants who had a harder time learning the new words obtained lower pronunciation ratings. It would therefore be of interest to test the effect of increasing the duration of the hand-clapping training sessions for the latter group of Chinese speakers and also the impact of clapping-based training on more advanced learners.

All in all, hand-clapping to the rhythm of new words in a second language was shown to be an effective tool to improve second language pronunciation under certain conditions. We suggest that is well worth further study given that this technique would be easy to implement in the language classroom and has great potential to help learners grasp second language rhythm and improve pronunciation.

5. Acknowledgements

We thank the pupils at the Col·legi Dr. Masmitjà in Girona and Zhang Qiu Experimental Middle School. We are grateful to A. Massanas, J. Li, C. Monnier and J. Borràs-Comes for their assistance. This research was supported by funding from the Spanish Ministry of Economy and Competitiveness (FFI2015-66533-P) and the Generalitat de Catalunya (2014 SGR-925). The first author acknowledges a grant awarded by the Dept. of Translation and Language Sciences at UPF.

6. References

- [1] N. Cason and D. Schön, “Rhythmic priming enhances the phonological processing of speech”, *Neuropsychologia*, vol. 50, no. 11, pp. 2652-2658, 2012.
- [2] N. Cason, C. Astésano, and D. Schön, “Bridging music and speech rhythm: Rhythmic priming and audio-motor training affect speech perception”, *Acta Psychologica*, vol. 155, pp. 43-50, 2015.
- [3] C. François, J. Grau-Sánchez, E. Duarte, and A. Rodríguez-Fornells, “Musical training as an alternative and effective method for neuro- education and neuro-rehabilitation”, *Frontiers in Psychology*, vol. 6, no. April, 2015.
- [4] N. Cason, C. Hidalgo, F. Isoard, R. Stéphane, and D. Schön, “Rhythmic priming enhances speech production abilities: Evidence from prelingually deaf children”, *Neuropsychology*, vol. 29, no. 1, pp. 102-107, 2015.
- [5] J. Fischler, “The rap on stress: teaching stress patterns to English language learners through rap music”, *MInneTESOL Journal*, vol. 26, 2009.
- [6] H. Wang, P. Mok, and H. Meng, “Capitalizing on musical rhythm for prosodic training in computer-aided language learning”, *Computer Speech and Language*, vol. 37, pp. 67-81, 2016.
- [7] D. Gluhareva and P. Prieto, “Brief training with beat gestures favors L2 pronunciation in discourse-demanding situations”, *Language Teaching Research*, vol. 21, no. 5, 2017.
- [8] A. Bhide, A. Power, and U. Goswami, “A rhythmic musical intervention for poor readers: A comparison of efficacy with a letter-based intervention”, *Mind, Brain, and Education*, vol. 7, no. 2, pp. 113-123, 2013.
- [9] M. Habib, C. Lardy, T. Desiles, C. Commeiras, J. Chobert, and M. Besson, “Music and dyslexia: A new musical training method to improve reading and related disorders”, *Frontiers in Psychology*, vol.7, no. January, 2016.
- [10] G. Calvert, C. Spence, & B.E. Stein (Eds.). “The handbook of multisensory processes”. Cambridge: MIT press, 2004.
- [11] S. McCafferty, “Gesture and the materialization of second language prosody”, *International Review of Applied Linguistics in Language Teaching*, vol. 44, pp. 197-209, 2006.
- [12] M.F. Bunting, N. Cowan, J.S. Saults, “How does running memory span work?”, *Quarterly Journal of Experimental Psychology*, vol. 59. pp. 1691–1700, 2006.
- [13] I. Vilà-Giménez, A. Igualada, & P. Prieto, “Observing storytellers who use rhythmic beat gestures improves children’s narratives discourse performance.”, *Developmental Psychology*, under review.
- [14] L. Macchi, C. Vansteene, N. Timmermans, and F. Boidien, “Épreuve Lilloise de Discrimination Phonologique (ELDP) : Présentation et illustration par deux études de cas cliniques”, *Les Cahiers de l’ASELF*, Vol. 10, no. 3, 2013.
- [15] M. Zentner and H. Strauss, “Assessing musical ability quickly and objectively: Development and validation of the Short-Proms and the Mini-Proms”, *Annals of the New York Academy of Sciences*, Vol. 1400, pp. 33–45, 2017.
- [16] K. Saito, P. Trofimovich, & T. Isaacs, “Using Listener Judgments to Investigate Linguistic Influences on L2 Comprehensibility and Accentedness: A Validation and Generalization Study”, *Applied Linguistics*, Vol. 38, no. 4, pp. 439–462, 2017.
- [17] J. Vaisière, “Rhythm, accentuation and final lengthening”. In: J. Sundberg, L. Nord, and R. Carlson (eds.), *French in music, language, speech and brain*. Wenner-Gren International Symposium (Series Macmillan Press), vol. 59, pp. 108-120, 1991.
- [18] T. Piske, I.R.A. Mackay, and J.E. Flege, “Factors affecting degree of foreign accent in an L2: A review”, *Journal of Phonetics*, vol. 29, pp. 191–215, 2001.
- [19] M. Long, “Maturational constraints on language development”, *Studies in Second Language Acquisition*, vol. 12, pp. 251–285, 1990.