Introduction

In recent years the need for teaching pronunciation in a second language (L2) has gained more attention because having good pronunciation skills in a second language leads people to have effective and efficient communication, giving a positive impression and avoids misunderstandings [1], [2].

There are some aspects of oral communication that are essential for having successful communication in a second language, some of them may affect the part of pronunciation, such as comprehensibility or accentedness, and others the part of oral abilities, such as fluency or expressiveness. We understand comprehensibility as the listener’s judgement of how difficult the L2 speech is to understand (as per [3], [4], [5], [6]). On the other side of the same coin, accentedness is the listener’s perception of how different the accent of a speaker is from that of the L1 community (as per [7], [8]). The term ‘foreign accent’ is, thus, used to describe the traits of the L2 speech that are evidently different from native speakers of a certain community [9]. Higher comprehensibility rates and lower accentedness rates will characterize a more native-like L2 pronunciation.

Another factor that impacts successful communication is fluency. It includes many aspects of a language (e.g., [10]) and it is closely related to oral abilities in the L2. We understand fluency as the overall speaking proficiency [11] and the ‘impression on the listener’s part that the psycholinguistic processes of speech planning and speech production are functioning easily and smoothly’ [12, p. 391]. When the speech is not considered fluent, disfluencies can be taken into account as the total filled pauses (such as um, er, etc.) and the repairs (i.e., all the expressions that have been repeated or reformulated) [13], [14]. Following this line, another factor to account for is expressiveness. It can be viewed as the perceived degree of involvement of the speaker, which can have an acoustic realization through the expansion of features such as loudness, pitch, and tempo [15], [16]. There are some acoustic features which have been mentioned to be useful to determine the expressiveness of L2 speech such as speaking rate, the length of pauses, stress patterns, post-focus compression, or the degree of contrast robustness [17].

Over the last decades many studies have investigated ways of improving second language pronunciation, either from a segmental or a suprasegmental point of view. In general, when suprasegmental and segmental traits are compared in the learning of a language, the latter plays a major role since suprasegmental traits tend to highly contribute to the native listeners’ perception of several features such as accentedness, comprehensibility or intelligibility (e.g., [18], [19], [20], [21]). According to [18] and [6], having an incorrect prosody in the L2 might have consequences in higher ratings of accentedness, and issues of comprehensibility and intelligibility. Furthermore, not only rhythm is important. As [22] states, what differentiates a fluent native speaker from a second language speaker is not only rhythm but also stress and intonation - which are suprasegmental features. Thus, focusing on teaching suprasegmental traits might be an appropriate way of learning the pronunciation of a foreign language.

In the field of L2 pronunciation instruction, the role of gestures and body movement as a pedagogical tool has generally been overlooked. Despite the inherent connection between sound production and body movement identified in gesture research, very few studies in this area report on the teacher’s use of nonverbal resources, including hand gestures. These few studies have shown that gesture use by both teachers and students are beneficial in teaching and learning L2 pronunciation ([23], [24], among others). Specifically, [23] provided an in-depth analysis of how teachers employ body movement in teaching L2 pronunciation and how students respond to such instruction in the flow of naturally occurring classroom interaction.

Abstract

The aim of this study is to assess whether a brief training with rhythmic beat gestures helps L2 pronunciation in a reading aloud task with high school students. In a between-subjects pretest-posttest design, a total of 59 high school students were randomly assigned to one of the following two conditions: the beat gesture group and no-beat gesture group. In the beat gesture condition they were asked to first read two short stories aloud without any gestural instruction (pretest) and in the following two texts they were asked to move their hands (training). Students in the no-beat condition (control condition) were asked to read the four texts aloud (pretest and training) without any gestural instruction. Then, in order to see the benefits of gesture, both groups were asked to read a fifth text aloud (posttest) which was more difficult (more complex syntactic structure and longer) than the ones they read in the pretest or the training. Results showed that speakers who were asked to produce beat gestures during the training had better pronunciation measures (specifically accentedness, comprehensibility, and fluency) in the posttest than the ones that were not asked to produce any specific gesture during the training.

Index Terms: beat gestures, pronunciation instruction, reading task, ESL.
Beat gestures are hand movements in which the hand makes a two-phase movement (in/out, up/down), and tend to coincide with rhythmically important parts of the speech [25]. The rhythmic features of beat gestures strongly relate them with the phonological structure of the target language. First, rhythmic beat gestures act as visual highlighters of the prosodic structure of the language; in pronunciation instruction, beat gestures can help marking the rhythm and the prominent syllables in a foreign language [26]. For example, [27] showed that beat gesture observation represents an effective aid for L2 acquisition of pronunciation. Second, emphasizing the rhythm of a second language might enhance reading fluency skills, as [28] has shown that better rhythmic regularity perception directly relates to better reading skills.

According to the authors, therefore, we hypothesize that a brief training with the production of beat gestures might be an effective aid for pronunciation instruction, specifically for pronunciation and fluency indicators, as they are markers of rhythm. The aim of the current study is to investigate whether a brief training with rhythmic beat gestures can help L2 pronunciation in a reading aloud task. In the current study, participants took part in a between-subjects experiment with two different conditions, namely beat gesture and no gesture, and they were randomly assigned to one of them. Following the study by [27], which showed that observing beat gestures have beneficial effects in accentedness and comprehensibility ratings, we hypothesise that speakers asked to produce beat gestures while reading aloud in the training will obtain better pronunciation and oral ability measures in the more difficult posttest than the speakers that were not asked to produce gestures during the training.

2. Method

2.1. Participants

Fifty-nine native Catalan speakers \((M = 95.159; SD = 10.127)\) from Institut Pere Alsius (Banyoles, Girona) participated in the study. They were 14 years old \((M = 14.08; SD = 0.281)\) and there were 29 female and 30 male participants. At the moment of the experiment, they were all in their 3rd year of high school and their English level as a second language was approximately elementary or lower intermediate. Parental consent was obtained before the experiment to process their data. Parents also stated that their children showed a normal development and that they do not have any language disorder.

2.2. Materials

Participants were given five different short texts and they were asked to read them aloud imagining that they were talking to someone. Both the task instructions and the five experimental items were presented in a power point presentation (each item in a power point slide). In this experiment, we used a total of five short texts; four about different cities or countries, and one about a festive event. The texts were double-checked by the students’ teachers to assure that they would be understood by the students. These five texts were then adapted in accordance with the English teachers’ comments, so that they were not too difficult to be read by students at the elementary level. The four texts about cities/countries (see example 1 below) were used for the pretest and training; the mean length of words was 22.5. They were made of sentences composed of coordinate lists and coordinate structures linked by “and”. These four texts were considered easy to read. The text about the festive event (see example 2 below) was used in the posttest and the mean length of words was 29. Sentences composed of main verbs and nested complement structures were used. As this text was considered difficult to read (as opposed to the pretest and training materials), participants were placed in a discourse-demanding situation (see [27]). While the order of the pretest and training texts was randomized, the posttest was always the same.

(1) Example of a pretest text: Portugal is a country with excellent sandy beaches, a very pretty coast, cosmopolitan cities, and incredible weather.

(2) More complex posttest target text: Hyde Park Wonderland is a spectacular festive event in the heart of the capital. For six weeks, Hyde Park is transformed into a magical wonderland of winter festivities.

2.3. Procedure

Participants were tested in a quiet room at the Institut Pere Alsius. Even though the activity was part of the English as a foreign language lesson, participants were tested individually in a different room. Each student was asked to sit in front of a computer so that they could follow the instructions and go through the experimental items. The experimenter did not look at the participant while s/he was reading the texts. The entire experimental task was video recorded with a Nikon d7000, located 5 meters in front of the participant. Before starting the task, participants were asked to first read the text silently to themselves one time in order to ensure that there were no vocabulary or pronunciation issues. Then, participants were instructed to imagine that they were talking to someone who wanted to travel to the indicated place. They needed to read the text to this person in a spontaneous and expressive way to give information about the specific country or city.

In a between-subjects design, participants were randomly distributed into two conditions: beat gesture and no-beat gesture condition. In the pretest, both groups were asked to read two different texts aloud, without any specific instruction. After reading the two texts, the beat gesture group was explicitly asked to read two different texts aloud while using their hands, contrary to the no-beat gesture group (control condition) which was asked to continue reading the texts aloud as the first two (therefore, in the no-beat gesture group there was no difference between the pretest and training). Thus, all participants read the first two texts without specific instruction about their hands and then only the beat gesture group was asked to use their hands while reading the two training texts. This training lasted approximately five minutes. After reading these four texts, all participants were asked to participate in a ten minute distraction activity. This allowed for the de-emphasis of gesture use in the beat gesture group.

Afterwards, all participants took part in the posttest, which consisted of reading a fifth text aloud without any specific gesture instruction. In other words, participants’ gesture performance was not controlled in this part of the task. As the posttest was more difficult than the pretest, we expect the scores to be generally lower in the posttest. If gesturing functions as an aide for the learners, then we expect the experimental group to maintain their scores, or to be somewhat shielded from the effects of difficulty in the posttest.
2.4. Pronunciation ratings

The participants’ audio recordings from the pre-test and posttest were rated by five native speakers of American English, as this is the variety of English to which students are most exposed according to their ESL instructors. The native speakers were living in Catalonia but were born in the USA. All of them reported normal hearing. The raters underwent a fifteen-minute training session with the experimenter where each aspect of speech was explained and the raters were given audio samples to rate. Raters then compared and discussed their ratings. This ensured that raters understood the difference between each of the aspects to be judged, and in the end there were no apparent discrepancies between the raters.

The experimental audio files were uploaded to Survey Gizmo (https://www.surveygizmo.com) and a survey was created for the raters so that they could assess the recordings in a random order. Each rater analyzed all of the 177 speech fragments on a scale of 1 (least native-like) to 5 (most native-like) for the following aspects of speech: accentedness, comprehensibility, fluency, and expressiveness.

2.5. Statistical analyses

Four GLMMs were run using SPSS Statistics 23.0 with four dependent variables: accentedness, comprehensibility, fluency, and expressiveness. Condition (Beat Group/No-Beat Group) and Test (Pretest/Posttest) were set as fixed factors. Participants and raters were set as random intercept.

3. Results

3.1. Perceptual review of the data

A perceptual review of the video recordings showed that when participants were instructed to move their hands during the training session, all participants used gestures, with beat gestures being the most common gesture type. Perceptual analyses also showed that in both the pre-test and posttest, there was very little gesture use, and no major differences between the two tests were observed.

3.2. Accentedness

The results of the GLMM analysis showed a significant effect of Condition on accentedness (F (1,881) = 9.487, β = 0.404, p = 0.002), no statistical significance for Test (F (1,881) = 3.670, β = 0.190, p = 0.056), and there is a nearly significant interaction between Condition and Test (F (1,881) = 3.416, p = 0.065). Pairwise contrasts show that Beat Condition received a better accentedness score than the No-Beat Condition in the posttest (β = 0.586, p = 0.000), but not in the pretest (β = 0.221, p = 0.172); also, that in the no-beat group there was a significant decrease from pretest to posttest (β = 0.372, p = 0.017), but no significant effect in the beat group (β = 0.007, p = 0.957). A graph of the results can be seen in Figure 1.

3.3. Comprehensibility

The results of the GLMM analysis showed a significant effect of Condition on comprehensibility (F (1,881) = 6.147, β = 0.326, p = 0.013), no statistical significance for Test (F (1,881) = 3.245, β = 0.184, p = 0.072), and a significant interaction between Condition and Test (F (1,881) = 7.016, p = 0.008). Pairwise contrasts show that the Beat Condition received higher comprehensibility ratings than the No-Beat Condition in the posttest (β = 0.597, p = 0.001), but not in the pretest (β = 0.055, p = 0.732); also, that in the no-beat group there was a significant decrease from the pretest to the posttest (β = 0.455, p = 0.005), but no significant effect was found in the beat group (β = 0.087, p = 0.491). Figure 2 shows the mean accentedness ratings as a function of Condition (Beat vs. No-Beat).

3.4. Fluency

The results of the GLMM analysis showed a significant main effect of Condition on fluency (F (1,881) = 5.911, β = 0.364, p = 0.015), no significant effect for Test (F (1,881) = 1.431, β = 0.119, p = 0.232), and a significant interaction between Condition and Test (F (1,881) = 9.670, p = 0.002). Crucially, pairwise contrasts show that the beat group received better fluency ratings than the no-beat group in the posttest (β = 0.673, p < 0.001), but not in the pretest (β = 0.056, p = 0.746); also, that in the no-beat group there was a significant decrease from pretest to posttest (β = 0.428, p = 0.005), but no
significant effect in the beat group ($\beta = 0.190$, $p = 0.140$). A graph of the results can be seen in Figure 3.

### 3.5. Expressiveness

The results of the GLMM analysis showed a significant main effect of Condition ($F(1, 881) = 8.704$, $\beta = 0.514$, $p = 0.003$), no statistical significance for Test ($F(1.881) = 0.852$, $\beta = 0.104$, $p = 0.356$), and no significant interaction between Condition and Test ($F(1, 881) = 1.147$, $p = 0.284$). However, if we explore the pairwise contrasts of the non-significant interaction between Condition and Test, we see that the Beat Condition received higher expressiveness ratings than the No-Beat Condition in the posttest ($\beta = 0.634$, $p=0.002$), but not in the pretest ($\beta = 0.394$, $p = 0.066$). And no significant differences are found between pretest and posttest for any of the two conditions (for the No-Beat Group, $\beta = 0.224$, $p = 0.183$; for the Beat Group, $\beta = 0.017$, $p = 0.911$).

![Mean fluency ratings](image1)

**Figure 3:** Mean fluency ratings (from 0 = less fluent to 5 = more fluent), broken down by training Condition (Beat vs. No-Beat) and test (pretest vs. posttest). Error bars represent 95% confidence intervals of the means.

![Mean expressiveness ratings](image2)

**Figure 4:** Mean expressiveness ratings (from 0 = less expressive to 5 = more expressive), broken down by training Condition (Beat vs. No-Beat) and test (pretest vs. posttest). Error bars represent 95% confidence intervals of the means.

### 4. Discussion

Our findings suggest that the use of beat gestures by students help boost various aspects of speech, particularly in regards to accentedness, comprehensibility, and fluency.

Regarding accentedness, the students who were not trained with gesture showed a significant decrease in their posttest scores with a harder text. This was not the case for the students who were trained with gesture. There was no significant difference in their pretest and posttest scores. This finding is reinforced by the fact that the posttest scores of the no-gesture group were significantly lower than that of the gesture-trained group. Similar results were found for comprehensibility and fluency ratings. Moreover, in all three cases there was an improvement in the posttest for these aspects of speech, though neither of which was significant.

The only aspect that did not show such robust results was that of expressiveness (even though non-significant tendencies are documented too). Expressiveness can be regarded as a more inherent trait of the individual performance than fluency, accentedness or comprehensibility and thus it might be more difficult to modify it than the other speech dimensions presented in this study.

### 5. Conclusions

Results showed that speakers asked to produce beat gestures while reading aloud in a short training session obtain better pronunciation measures (specifically accentedness, comprehensibility, and fluency) in the posttest than the ones that were not asked to produce any specific gesture during the training session. Although the same tendency was found, expressiveness measures were not significantly affected, as they were more stable across pretest and posttest readings. The results also show that using beat gestures to highlight prosodic structure in speech can be profitably used not only for oral description tasks (e.g., [30]) but also for reading aloud tasks in the second language classroom.

These results add to the growing body of literature that support beat gestures as tools in the second language classroom. The novel aspect of this article is that students who are encouraged to gesture while reading aloud have been shown to maintain accentedness, comprehensibility, and fluency ratings in discourse-demanding situations. Future work may take advantage of more precise acoustic analyses of L2 speech. For example, [29] found that VarcoV, especially when coupled with speech rate, is an effective predictor of accentedness ratings. Finally, the long term effects of beat gesture training have not been explored. Previous studies (as well as the current study) use experimental designs with very short training periods. It would certainly be interesting to assess potential long-term effects, as it would add reinforcement to the argument that beat gestures have an important place in the second language classroom.

### 6. References


