Benefits of rhythmic beat training and prosodic training for second language pronunciation instruction

PhD research project

Judith Llanes-Coromina
15/06/2017

Doctorat en Traducció i Ciències del Llenguatge

Supervisor: Dr. Pilar Prieto
# INDEX

0. ABSTRACT .................................................................................................................. 4  
1. INTRODUCTION ......................................................................................................... 7  
   1.1 OBJECT OF ANALYSIS .................................................................................... 7  
   1.2 PRIOR WORK ...................................................................................................... 7  
      1.2.1 The importance of oral skills and pronunciation instruction in the second language classroom .................................................................................................................. 7  
      1.2.2 L2 pronunciation teaching and learning: experimental studies ...................... 10  
      1.2.3 Effects of auditory and articulatory training .................................................. 11  
      1.2.4 Effects of segmental and suprasegmental training ........................................ 13  
      1.2.5 Benefits of gestures in first and second language acquisition ....................... 16  
      1.2.6 Suprasegmental pronunciation instruction and body gestures ...................... 19  
      1.2.7 Why beat gestures can help pronunciation instruction? ................................ 21  
2. GOALS OF THE DISSERTATION ............................................................................. 22  
3. THEORETICAL FRAMEWORK ................................................................................... 23  
   3.1 Gesture and speech as an integrated system ......................................................... 23  
   3.2 Embodied cognition ............................................................................................ 24  
4. HYPOTHESES ........................................................................................................... 25  
5. EXPERIMENTAL STUDIES ....................................................................................... 26  
   5.1 STUDY 1 ................................................................................................................ 26  
      5.1.1 Goals ............................................................................................................. 26  
      5.1.2 Participants .................................................................................................... 27  
      5.1.3 Experimental materials ................................................................................ 27  
      5.1.4 Procedure ....................................................................................................... 28  
      5.1.5 Ratings ........................................................................................................... 29  
      5.1.6 Results ........................................................................................................... 30  
   5.2 STUDY 2 ................................................................................................................ 30  
      5.2.1 Goals ............................................................................................................. 30  
      5.2.2 Participants .................................................................................................... 31  
      5.2.3 Experimental materials ................................................................................ 31  
      5.2.4 Procedure ....................................................................................................... 32  
      5.2.5 Ratings ........................................................................................................... 34  
      5.2.6 Results ........................................................................................................... 34  
   5.3 STUDY 3 ................................................................................................................ 35
5.3.1 Goals ................................................................................................................................. 35
5.3.1 Overall method ..................................................................................................................... 35
5.2.2 Participants ........................................................................................................................... 35
5.2.3 Experimental materials ........................................................................................................ 35
5.4 STUDY 4 .................................................................................................................................. 37
5.4.1 Goals .................................................................................................................................. 37
5.4.2 Participants ............................................................................................................................ 38
5.4.3 Experimental materials and procedure .................................................................................. 38
6. WORKING SCHEDULE ............................................................................................................... 39
7. SELECTED REFERENCES .......................................................................................................... 41
8. GENERAL REFERENCES ............................................................................................................ 44
0. ABSTRACT

In recent years the need for teaching pronunciation and communication skills in a second language (L2) has gained more attention (Bakar & Abdullah, 2015; Offerman & Olson, 2016). Existing pronunciation paradigms are typically based on specific drills that practice with isolated sounds and phrases, and, to our knowledge, there are no integrative methods that take into account the role of the communicative goals of the speaker, together with the rhythmic highlighting of the sound structure of the target language through body gestures and prosody.

Following an embodied cognitive approach, the aim of this project is twofold, namely (a) to investigate the role of encouraging the production of rhythmic gestures (or rhythmic hand/arm movements produced together with prominent prosody) in the acquisition of L2 pronunciation; and (b) to create an educational tool that will allow high-school students to learn pronunciation and communication skills more effectively, specifically when learning English as a L2. The thesis will consist of four interrelated experimental and educational studies which will be run with junior highschool students (e.g., 14- and 15-year-olds). The first study will investigate whether encouraging the production of beat gestures enhances L2 pronunciation reading abilities. The second study will focus on how encouraging beat gestures can enhance L2 pronunciation when producing a spontaneous discourse. The third study will investigate the potential beneficial effects of focusing on pragmatically contextualized prosodic patterns and beat gestures on L2 discourse pronunciation. The fourth study will be a classroom intervention study which will be based on the teacher and student embodied mediation to teach second language pronunciation. Finally, we will create an educational tool based on gestural and prosodic activities to improve English pronunciation. We believe that a paradigm which integrates the pragmatics and rhythmic highlighting will increase the effectiveness of L2 pronunciation training in relation to present educational approaches.
Durante los últimos años, en el ámbito de la enseñanza de segundas lenguas (L2) se ha generado un creciente interés por la necesidad de potenciar la expresión de habilidades comunicativas, especialmente a través del perfeccionamiento de la pronunciación (Bakar & Abdullah, 2015; Offerman & Olson, 2016). Para ello, la mayoría de paradigmas existentes proponen una metodología basada en el diseño de ejercicios destinados a practicar sonidos específicos y frases aisladas. Hasta donde conocemos, actualmente no existen propuestas metodológicas que trabajen específicamente la relación que existe entre los objetivos comunicativos del hablante y la estructura rítmica de los sonidos de la segunda lengua, expresada a través de los gestos y de la prosodia.

El doble objetivo de esta tesis consiste en (a) probar de forma experimental un conjunto de técnicas que investigan el papel de la gestualidad rítmica en el aprendizaje de la pronunciación de una segunda lengua; y (b) diseñar una herramienta educativa que permita a los estudiantes de secundaria desarrollar sus habilidades comunicativas a través del trabajo y del perfeccionamiento de la pronunciación —y de la gestualidad asociada— de una segunda lengua como el inglés. El primer estudio investigará si el impulso a producir gestos rítmicos ayuda a mejorar la pronunciación de una L2 cuando se lee un discurso. El segundo estudio se centrará en fomentar la producción de gestos rítmicos para mejorar la pronunciación L2 cuando se produce un discurso espontáneo. El tercer estudio investigará los posibles efectos de centrarse en los patrones prosódicos y en los gestos rítmicos en la pronunciación de un discurso en inglés como L2. El cuarto estudio consistirá en un entrenamiento —incorporado a la actividad regular del aula— que incluirá diversas actividades basadas en la prosodia y gestualidad, cuyo objetivo será el de comprobar los efectos de la enseñanza explícita de la pronunciación a través del uso de gestos y prosodia. Finalmente, se pretende crear una herramienta educativa basada en las actividades gestuales y prosódicas para mejorar la pronunciación del inglés. Creemos que un paradigma que sea capaz de integrar la pragmática y que otorgue relevancia al componente rítmico de la lengua —expresado a través de los componentes prosódico y gestual del habla—, aumentará la eficacia del entrenamiento de la pronunciación de una L2 en relación con los enfoques educativos actuales.
En els últims anys s'ha fet palesa la necessitat que té el nostre ensenyament reglat d'integrar l'ensenyament de les habilitats orals i de pronúncia en una segona llengua con l'anglès (Bakar & Abdullah, 2015; Offerman & Olson, 2016). La majoria de paradigmes existents en l'ensenyament de la pronúncia d'una segona llengua es basen en la pràctica d'exercicis específics que treballen amb sons i frases aïllades, i s'ha treballat poc en provar de forma experimental una sèrie de mètodes innovadors que potencien tant l'estructura rítmica de la segona llengua a través dels gestos i la prosòdia com la integració pragmàtica d'aquestes activitats.

El doble objectiu d'aquesta tesi és: (a) provar experimentalment una sèrie de tècniques que investiguen el rol de la gestualitat rítmica en l'aprenentatge de la pronúncia d'una segona llengua; i (b) crear una eina educativa que permetrà als estudiants de secundària aprendre la pronunciació i habilitats comunicatives en la seva segona llengua. El primer estudi investigarà si la producció de gestos rítmics pot ajudar a millorar la pronunciació L2 quan es llegeix un discurs. El segon estudi investigarà si producció de gestos rítmics pot millorar la pronunciació L2 quan es produeix un discurs espontani. El tercer estudi investigarà els possibles efectes de la producció de gestos rítmics i de la imitació de patrons prosòdics integrats pragmàticament en la pronunciació d'un discurs en anglès. El quart estudi serà un estudi d'intervenció a l'aula que inclourà diversos tallers basats en la prosòdia i gestualitat per a poder valorar els efectes de l'ensenyament de la pronunciació a través de l'ús de gestos i prosòdia. Finalment, crearem una eina educativa que inclourà en les activitats gestuals i prosòdiques per millorar la pronunciació de l'anglès. Creiem que un paradigma que integri tant aspectes de contextualització pragmàtica com la potenciació del ritme de la llengua pot significar un augment de l'eficàcia en l'entrenament de la pronunciació d'una L2 en relació amb tècniques més tradicionals.
1. INTRODUCTION

1.1 OBJECT OF ANALYSIS

The aim of the current PhD proposal is to investigate: (a) the relationship between prosodic and gestural abilities and oral communication skills in high-school students when learning pronunciation in their second language; and (b) the potential benefits of using prosodic and gestural strategies to improve oral communication skills in their second language, and specifically the effects of beat gestures (rhythmic hand/arm movements which are typically produced in association with prominent prosody) in L2 pronunciation improvement. The basic hypothesis of the thesis stems from results of recent studies which have demonstrated the cognitive and developmental benefits of the use of co-speech gesture in different areas, such as in first and second language acquisition.

This PhD project plans to carry out three experimentally designed tasks and one classroom intervention study to test the value of prosodic and gestural training in promoting oral language abilities in junior high school students when learning a second language. The social significance of this research topic is high, as training procedures based on embodied cognitive awareness can be proven valuable for the improvement of oral communication abilities in educational settings.

1.2 PRIOR WORK

1.2.1 The importance of oral skills and pronunciation instruction in the second language classroom

Oral skills are a crucial aspect of communication both in first and second language\(^1\), since a more expressive, fluent and comprehensible discourse may affect on the listener’s perception of such speech (e.g., Morreale, Osborn, & Pearson, 2000, among many others). Morreale, Osborn, & Pearson (2000) reported that oral skills by high school students may contribute to the social adjustment and participation of the students in satisfactory interpersonal relationships. By contrast, teenagers with poor oral communication skills have fewer friendships and they are viewed as less attractive by

\(^1\)In this work, the term ‘second language learning’ is used as a cover term that refers to the process of learning another language after the native or dominant one. This is a common strategy in the field, which uses this term to refer to the learning of a third or a fourth language (Gass, 2013).
the other students. They also stated that being able to communicate in a proper way orally favours a better interaction with the others (Morreale, Osborn, & Pearson, 2000). Thus, effective oral skills are not only crucial for having an intelligible and comprehensible speech (both in a first and in a second language) but also for better interpersonal relationships.

Since English language is spoken all over the world (Jenkins, Cogo, & Dewey, 2011), communicating fluently in English, as well as having good pronunciation skills such as producing intelligible and comprehensible speech, has become more important (Hüttner, 2009). Some authors such as Offerman et al. (2016) reported that having good pronunciation skills in a second language has a positive impact on conveying meaning in an effective and efficient way (Offerman et al., 2016; see also Bakar & Abdullah, 2015). Even though in recent years the need for teaching pronunciation and communication skills in a second language, such as English, has gained more attention (Kissling, 2013), research shows that pronunciation and oral skills in a second language have received little attention compared with areas such as morphology or syntax (e.g., Deng et al, 2009; Offerman et al, 2016). Rajadurai (2001) performed a study with ESL teachers to know the attitude and concern for accurate English pronunciation and results showed that more than 80% of the teachers see pronunciation as an important matter (Rajadurai, 2001; Bakar & Abdullah, 2015). Not only teachers see the importance of pronunciation for spoken English. Students are concerned about the accuracy of their pronunciation (Dalton & Seidlhofer, 2000). According to several studies, some reasons why teachers do not focus enough on pronunciation might be the lack of training in teaching pronunciation, the lack of available resources or the unclear procedures for assessing pronunciation (Baker, 2014; Macdonald, 2002, Smotrova, 2017). Teachers themselves are reported to often feel reluctant or unconfident teaching pronunciation (Baker, 2014; Macdonald, 2002). Darcy, Erwert & Lidster (2012) described some reasons for which ESL (English as a second language) teachers do not focus enough on pronunciation instruction. They performed 14 interviews to IEP (Intensive English Program of Indiana University, US) teachers to know how different elements of pronunciation, such as intonation, rhythm, perception ability, clarity, among others, are approached. Results showed that teachers believe that some specific traits of pronunciation may be taught at different levels of English proficiency, for example, that learners with a higher proficiency level should be taught more suprasegmentals than
segmentals and the learners with a lower level of proficiency should be taught firstly with segmentals. Results also showed that 71% of the teachers that participated in the survey do not teach pronunciation at all and only 14% reported to practise pronunciation with a specific rubric for its evaluation. Moreover, the interviews revealed that teachers find pronunciation instruction difficult to do for several reasons such as lack of time available (43%), lack of training (43%), lack of training (25%) and need for more guidance (25%). Therefore, teachers admit to not having clear guidelines for pronunciation instruction. Moreover, since there are partially contradictory practices and approaches when teaching pronunciation, such as practising production or perception, or practising suprasegmentals or segments (Foote, Holtby & Derwing, 2011), it is unclear for teachers that they should follow a well-established method that can be used to teach ESL pronunciation. The study by Foote et al. (2011) assessed the status of teaching pronunciation in ESL lessons in Canada by comparing the current practises with the ones from 10 years ago. They performed surveys to teachers to know about beliefs, approaches, and resources about teaching pronunciation. Results showed that pronunciation instruction had not changed a lot for 10 years in Canada, only that there are more training opportunities and more pronunciation courses available. They did find differences on how teachers approach segmentals and suprasegmentals in their instruction practices: while ten years ago there was emphasis both in segmentals and suprasegmentals, and now there is greater focus on segmentals.

The experimental research presented in this PhD thesis will be centred on the junior high school student population (ages 14 to 16) at one high school in the north of Catalonia, namely INS Pere Alsius (Banyoles). This high school uses English books from two well-known publishers, Oxford and Burlington, which are two of the most commonly used publishers in high schools all across Catalonia. Mosaic (Oxford) is an elementary level textbook used by second graders (4th ESO) (15-16 years old). Inspection of the book shows that it includes only two pages in the whole book focusing at pronunciation. These pages include many activities based on segmental practice. For example, exercises in which students have to pay attention to the distinction between /s/ and /z/. There are only two exercises in the whole section based on suprasegmentals in which they have to pay attention to sentence stress. In one of the exercises, students are asked to stress the important syllables of a written text. New Action! 3 (Burlington) is an elementary level textbook used by first graders (3rd ESO) (14-15 years old). Inspection
of the book shows that again this book includes only two pages of pronunciation practise presented in an appendix. These pages include exercises to practice the pronunciation of contrasting segments such as /s/, /z/ and /iː/, /I/. For suprasegmentals we can find activities based on rhythm and intonation of the sentences. In general, in both textbooks the principal focus of this short pronunciation section is on segmentals while suprasegmentals play a secondary role.

Generally, there is a lack of pronunciation instruction in the ESL language manuals. Nevertheless, nowadays in the market we can find many pronunciation books that focus exclusively on English pronunciation instruction. Some of them are specific for learning pronunciation of English as a second Language (Baker, 1977; Hewings, 2007, among others), others are to reduce the accent (Chwat, 1994) or to train a specific accent (Cook, 1991), and others that specifically compare Spanish pronunciation and English pronunciation, so that Spanish learners of English can improve their pronunciation (Poms & Dale, 1985; Estebas 2009, among others). A good proposal for Spanish learners is the one made by Estebas (2009) which is a monograph for Spanish speakers to learn English pronunciation and which includes a wide set of segmental exercises combined with suprasegmental traits, such as stress, rhythm, intonation, among others.

All in all, it is clear that pronunciation instruction remains a challenge in L2 teaching and learning (Darcy et al. 2012). In general, there is a lack of empirical research on pedagogical methods for teaching L2 pronunciation, and conversely, there is a need for more research into how ESL teachers can enhance oral and pronunciation skills.

1.2.2 L2 pronunciation teaching and learning: experimental studies

There are some aspects of oral communication that are essential to having successful communication, both in a first language and in a second language, such as fluency, comprehensibility, accentedness and expressiveness. These aspects have all been taken as elements to assess oral competence in a second language. Fluency includes many aspects of a language (e.g., Derwing, Rossiter, Munro, & Thomson, 2004). Fluency has many different definitions. We can understand fluency as the overall speaking proficiency (Chambers, 1997) and the ‘impression on the listener’s part that the psycholinguistic processes of speech planning and speech production are functioning easily and smoothly’(Lennon, 1990, p. 391). When the speech is difficult to rate,
disfluencies can be taken into account as the total filled pauses (such as *um, er*, etc.) and the repairs (all the expressions that have been repeated or reformulated) (Wright, 2013; Wright & Zhang, 2014). Comprehensibility can be understood as the listener's perception on how difficult to understand is the speech produced (Baker & Burri, 2016; Derwing et al., 1998; Derwing & Munro, 1997). Expressiveness can be measured as the perceived degree of involvement of the speaker (Swerts & Krahmer, 2010). Accentedness in a second language can be understood as the degree to which the listener thinks an utterance is different phonetically from native speaker utterances (Munro & Derwing, 2001).

Research over the years has provided many different methods to teach L2 pronunciation in an effective way. Pennington & Richards (1986) stated several recommendations for teaching pronunciation in second language teaching, such as that pronunciation must be settled as a long-term goal, that the teaching of pronunciation should be gradually reducing the influence of native sound patterns on segmental, voice-setting and prosodic features, among others. In this subsection we review this work on instructional practices.

### 1.2.3 Effects of auditory and articulatory training

Over the last decades many studies have investigated ways of improving second language pronunciation, either from a segmental or a suprasegmental point of view. From a segmental point of view, it has been shown that phonetic training can improve speech perception and production abilities in the second language, be it using one training session (Carney et al. 1977, Pisoni et al. 1982) or longer trainings (Logan & Pruitt, 1995). Nevertheless, the positive effects of phonetic training do not depend on the amount of time that a learner is exposed to L2 sounds, but to the quality of the training, i.e. whether the training succeeds in drawing the attention of the learners, and the ability to process L2 sound phonetic cues produced by native speakers (Aliaga-García & Mora, 2009). According to Kartushina et al. (2015), the main goal of L2 perception training studies is usually to enhance the perception of L2 contrasts, such as consonantal contrasts, tone contrasts, and typically the ones that are more difficult because they have similarities to one L1 category. These trainings include tasks such as discrimination or identification exercises involving minimal pairs of words, etc.
(Kartushina, Hervais-Adelman, Frauenfelder & Golestani, 2015). Aliaga-Garcia & Mora (2009) investigated the effects of phonetic training on perceptual and productive competence of learners of English as an L2 focusing on four L2 sound contrasts that Catalan and Spanish bilinguals have reported to be difficult (Cebrian, 2002, 2006; Mora, 2007; Mora & Fullana, 2007). Results showed that learners produced and perceived the target pairs of sounds significantly more accurately after the training, i.e., the training sessions were successful for the experimental group in the improvement of pronunciation accuracy, but there were no gains in the overall production and perception of the sounds.

The High Variability Phonetic Training approach (Logan, Lively & Pisoni, 1991) has been shown to be effective in lab conditions. This approach places a strong emphasis on the auditory exposure of the trainees to normal phonetic variability in a phonetic category. In other words, HVPT is a perception-only approach that focuses the use of several speakers in several phonetic contexts so that the stimulus variability is increased (Wong, 2015; Carlet & Cebrian 2014). Iverson, Pinet and Evans (2012) performed an auditory training for experienced and inexperienced second-language learners of English (French native speakers) using HVPT. Eight sessions of high-variability phonetic training for English vowels were performed. To evaluate their improvement participants were asked to take part into several perception and production tests. Results showed that HVTP training is different from the one produced in more natural situations because both groups of participants showed similar degrees of learning and that it facilitates production and perception. Therefore training perception can benefit the production of several sounds when there is explicit focus on the phonetic form L2 pronunciation is effective (Mora & Levkina, 2017).

Moreover, HVTP training does not only focus on the ability to learn the items that are trained but also to assess the “generalization of learning” (e.g., Logan and Pruitt, 1995; Carlet & Cebrian 2014). This generalization is understood as the ability to transfer the knowledge acquired to several dimensions; if this generalization is visible it means that robust learning is present (Logan and Pruitt, 1995; Carlet & Cebrian, 2014; Mora & Levkina 2017). In the study by Carlet & Cebrian (2014), they investigated the effects of a short-term high variability phonetic training on the perception of English consonant and vowel contrasts. Some of the tasks of the training were discrimination and identification tasks, and results from the post training on identification showed that
there was a positive effect of the training for some vowels and consonants, but not all of the ones tested. Moreover, they tested generalization of the improvement on novel words pronunciation and they showed that robust learning might occur even after a short period of training. In the study by Mora & Levkina (2017), they investigated whether a treatment in improving L2 perception of target vowels contrast would be effective, whether this treatment would be effective for production of the target vowels contrast and whether the improvement can generalize to new items or speakers. Results after the treatment show that there was an improvement in perceptual sensibility to the contrast /i:/-/ɪ/, that there was a generalization to new nonwords.

Some studies have assessed an approach which combines L2 perception with L2 production training, with positive effects (Aliaga-Garcia and Mora, 2009; Delvaux et al., 2013; Massaro et al., 2008) Some of these studies have also used several production tasks with computer-based visual feedback, showing that they are especially successful for improving L2 production accuracy (see Kartushina et al. for a review, 2015).

Nevertheless, auditory cues are not the only important element in communication, visual cues play also an important role (Cebrian & Carlet, 2012). Cebrian & Carlet (2012) tested the role of audiovisual cues in the perception of several consonants of English and they assessed the effect of cue salience and differences in L1-L2 in phonemic status on non-native perception. Participants were divided in several groups: three groups of listeners, a group of native speakers of English, and two groups of EFL learners and stimuli were presented in three different conditions: audiovisual stimuli, auditory-only and visual-only. Results showed that the salience of visual cues plays an important role in the perception of native speakers. However, in non-native speakers production is more influenced by the status of the native and the L2 sounds. There was also showed a positive effect of L2 experience when perceiving a more native-like manner in both auditory and visual conditions.

1.2.4. Effects of segmental and suprasegmental training

Several classroom intervention studies have focused their attention on the effects of learning pronunciation through segmental and/or suprasegmental trainings. Derwing & Rossiter (2003), performed an experiment to compare three types of methodology based
on a segmental, global (mostly suprasegmental features) and no specific pronunciation instruction. Results showed that the global one, that included suprasegmental abilities, improved more in comprehensibility and fluency. Derwing, Munro & Wiebe (1998) explored in a classroom-based experiment three types of pronunciation teaching instruction on fluency, comprehensibility and accentedness ratings. They tested three methods: segmental, global (i.e., suprasegmentals) and no pronunciation-specific instruction. Results show that while segmental and global training triggered improvements in comprehensibility and accentedness in a sentence-pronunciation exercise, only the global group improved with spontaneous speech. The classroom-based study by Gordon, Darcy & Erwert (2013) compared the benefits of a segmental training and a suprasegmental training. Participants were divided into three groups and they were under pronunciation instruction for 3 weeks (25 min per day, 3 days per week). Results showed that comprehensibility was improved only in the suprasegmental training group. Finally, the study by Gordon & Darcy (2016), a short classroom intervention study was performed to compare the effects of explicit vs. non-explicit pronunciation instruction in three different groups of participants, as follows: pronunciation instruction on suprasegmental features; pronunciation instruction on four vowel sounds (segmental instruction); and no explicit instruction. The three groups received four hours of pronunciation instruction during three weeks following the same teaching sequence. Results showed that the group that was instructed on suprasegmental features obtained higher comprehensibility ratings.

In general, the abovementioned studies provide evidence for the importance of a suprasegmental approach when learning a second language. In general, when suprasegmental and segmental traits are compared in the learning of a language there is a major role of suprasegmental training, since suprasegmental traits tend to highly contribute to the native listeners' perception of several features, such as accentedness, comprehensibility or intelligibility (Anderson-Hsieh, Johnson, & Koehler, 1992; Edmunds, 2010; Field, 2005; Ulbrich, 2013). According to Anderson-Hsieh, Johnson & Koehler (1992) and Derwing & Munro (2009), having an incorrect prosody in the L2 might have consequences in higher ratings of accentedness, and issues of comprehensibility and intelligibility. Moreover, learning the rhythm of a language, which is part of prosody, is essential for improving L2 pronunciation (Bekius et al., 2016). According to Bekius et al. (2016), people that have a better rhythmic regularity
perception have better reading skills in L2. Furthermore, not only rhythm is important. As Adams-Goertel (2013) states, what differentiates a fluent speaker from a second language is rhythm but also stress and intonation, which are suprasegmental features. Thus, focusing on teaching suprasegmentals traits might be an appropriate way of learning the pronunciation of a foreign language.

Sardegna (2009) performed a longitudinal study on the long-term effectiveness of pronunciation instruction, specifically on suprasegmentals. In her study, she assessed the progress of the students in English primary phrase stress, constructions stress and word stress has during a semester of a pronunciation course. The study assessed the pronunciation improvement of the students in reading aloud in English words and dialogs. Results showed that the instruction of pronunciation learning strategies was effective for the improvement of the read discourses in primary stress and word stress, among others. Therefore, a guided practice in suprasegmentals for pronunciation improvement extended to a classroom-based training might be useful for improving pronunciation. A method based in self-assessment of pronunciation as the centre of pronunciation teaching was proposed by Sardegna & McGregor (2013). In this method, based on suprasegmentals, the pronunciation goals were based on the student needs. Students were encouraged by explicit instruction, guided practice and learning strategies, i.e., through teaching strategies, and opportunities were given for students to pay attention to their performance during the pronunciation exercises and see their own outcomes. With this methodology students were able to recognize, comprehend, correct and improve the pronunciation challenges on their own (Sardegna & McGregor, 2012). Therefore, a program that enables self-reflective practice for learners might be very productive.

Derwing, Munro, Foote, Waugh, & Fleming (2014) performed a pronunciation training program with ESL speakers who had lived in an English-speaking context for an average of 19 years. The training was based on how learner's perception of certain segments of prosody makes them improve in terms of comprehensibility, accentedness and fluency. Some of the activities of the training were to record themselves after completing a perception component of such assignment, dictation tasks, making stress

---

2When a word in the sentence is more stressed than the others: What KIND of tipes? (Sardegna, 2009).
3When a multi-word, like compound nouns, is more stressed than the others: When I lived on RANDOLPH Street, I had NINETY-five neighbors. (Sardegna, 2009)
4 When a syllable of the word is stressed more than the others: deISions, acadEmic (Sardegna, 2009)
on both words and sentences, among others (Derwing, Munro, Foote, Waugh, & Fleming, 2014). All these training activities were used to focus on the prosody of the English language and they were effective in enhancing comprehensibility, fluency and accentedness of the speech produced.

While suprasegmental training has proven to be successful in improving second language learners’ overall fluency and comprehensibility, almost no work has tested the efficacy of specific techniques and paradigms. Despite the apparent importance of suprasegmental instruction, there exists little concrete evidence showing the superiority of one suprasegmental training method over another, as the previous studies have trained participants in a mixture of features using various teaching methods and techniques. The main objective of this PhD thesis is to contribute to assessing the efficiency of training methods that highlight rhythmic structure through body and hand gestures in order to inform L2 pronunciation instruction practices. In the next section we review the work that has been carried out on the positive effects of using gestures for first and second language acquisition.

Apart from pronunciation instruction, increasing the frequency of exposure to the second language can help reducing the foreign accent. In a recent study, Llanes, Mora & Serrano (2016) showed that studying abroad in an English language speaking country can also help reducing the foreign accent. In their study, they tested two different groups of students, the first group was instructed with a study abroad course in the UK and the second group with an intensive course at home in Spain. Results showed that VOT and foreign accent reduction were significantly correlated, i.e., the higher the VOT values the less foreign accent.

1.2.5 Benefits of gestures in first and second language acquisition

Research in the last few decades has provided ample evidence that gestures are an effective tool for language learning (see Gullberg, 2006 for a review). For example, it has been shown that iconic (or representational) gestures have a positive effect on the recall of information in a first language by adults (So, Sim Chen-Hui, & Low Wei-Shan, 2012; Thompson, 1995) and children (Cook, Mitchell, & Goldin-Meadow, 2008; Goldin-Meadow, Kim, & Singer, 1999; So et al., 2012). With a recall task using videos in which verbal communication was depicted either with gestures or without them, Riseborough (1981) showed that participants remembered information presented with
gestures better than information presented without. Similarly, Thompson (1995) performed an experiment in which adult participants were asked to recall words after the information had been conveyed to them by either an invisible speaker, a visible speaker, or a visible speaker also using iconic gestures. The results showed that participants remembered words better when they were accompanied with iconic gestures. Cook et al. (2012) carried out an experiment in which adult participants had to remember some letters while explaining the solution to maths problems and, at the same time, producing either iconic meaningful and non-meaningful gestures. The results of the experiment demonstrated that participants remembered more items when meaningful iconic gestures were produced than when either no gesture at all or non-meaningful gestures were produced. Similarly, studies carried out with children have shown that they tend to remember target words better when they are accompanied by representational gestures than when they are not (Cook et al., 2008; Goldin-Meadow et al., 1999; Tellier, 2008).

In the realm of L2 acquisition, several studies have shown that using iconic gestures improves the memorization of novel words in adults (Kelly, Barr, Church, & Lynch, 1999; Kelly, McDevitt, & Esch, 2009; Macedonia & Kriegstein, 2012; Macedonia, Müller, & Friederici, 2011) and children (Macedonia, Bergmann, & Roithmayr, 2014; Porter, 2012; Tellier, 2008). In the study by Tellier (2008), for example, French children were instructed to learn several novel words in English, one group with pictures and the other group with iconic gestures. Results indicated that children who learnt the words with iconic gestures performed the recall task better.

Meaningful gestures not only improve recall, they also boost comprehension processes in listeners (Cocks, Morgan, & Kita, 2011; Goldin-Meadow et al., 1999; Hostetter, 2011; Kelly et al., 1999; McNeil, Alibali, & Evans, 2000). Several studies have shown how incongruent combinations of iconic gestures with speech negatively affect comprehension (Gunter, Weinbrenner, & Holle, 2015; Kelly, Ozyürek, & Maris, 2010). Kelly et al. (2010) tested participants with iconic gestures that referred to action primes (e.g., chopping vegetables) and combined them with target words which were either congruent or incongruent with those gestures. The results showed that listeners were quicker and more accurate when relating primes to targets when the gesture-speech information was congruent than when it was incongruent. In a more recent study by Gunter et al. (2015), participants were asked to watch a video where a person performed
pointing gestures that were consistent or inconsistent with a previously denoted location. The results proved that the inconsistent pointing gestures impaired participants’ language comprehension.

In the area of child development, many studies have shown that representational and deictic gestures guide children towards the semantic content of a message, that is, they help children comprehend the discourse they are listening to (Clark, Hutcheson, & van Buren, 1974; Goldin-Meadow & Wagner, 2005; Kelly et al., 1999; McNeil et al., 2000). McNeil et al. (2000) pointed out that “reinforcing gestures are an effective scaffold for children’s comprehension of complex spoken messages because they guide comprehension toward the meaning of the spoken language” (McNeil et al., 2000, p. 114). Hostetter (2011) undertook a quantitative meta-analysis of 63 studies on the overall communicative role of co-speech gestures involving adults and/or children and found strong evidence that gestures foster comprehension in listeners. While the size of the beneficial effect was moderate overall, it varied depending on factors such as the types of meanings encoded in gesture or the semantic overlap of the gestures with speech. However, all the gestures involved in the above-mentioned studies were concrete representational or deictic gestures, that is, gestures with a substantial semantic component. Our own scrutiny of the studies included in Hostetter’s (2011) meta-analysis revealed that all the gestures used in those studies were either iconic, metaphoric, or deictic gestures—none of them included beat gestures.

Importantly, the majority of studies on the benefits of gestures have dealt with representational/iconic or metaphoric gestures, that is, gestures which either express specific semantic information in their own right or more abstract ideas. However, little is known about the potential effects of another type of gesture found in adult repertoires, namely beat gestures, which are non-referential hand gestures that co-occur with prosodically prominent positions in speech (McNeill, 1992; see also section 3.1 for a review). Igualada et al. (2017) and Llanes-Coromina et al. (under review) have recently shown that beat gestures help children to better recall and comprehend information from a narrative discourse. In a foreign language, the work by Kushch, Igualada & Prieto (2016) has shown that beat gestures can have a positive effect in novel word learning.
However, little is known about how beat gestures can help learning pronunciation aspects of language. In the section below, we review the work on pronunciation instruction which is focused on enhancing the prosodic features of the target language.

1.2.6 Suprasegmental pronunciation instruction and body gestures

In 1965, the teaching method called "Méthode Structuro-Globale Audio-Visuelle or "SGAV" (Gluberina, 1965) defended that in foreign language learning, phonetics should be taught before any other aspect of the language. With this method, oral skills are taught before all written elements. On the basis of this method, the so-called Verbo-Tonal Method (VTM) (Gluberina & Asp, 1981) was implemented. VTM focuses on the rhythmic traits of the foreign language to make them more visible and to emphasize the phonemic traits of such language. Alazard et al. (2010) compared the effects of an oral skills' training using the VTM (experimental group) with activities based on prosody and for the control group articulatory exercises, such as activities based on reading, text comprehension, reading aloud or creative writing. Results of an oral reading task showed that the VTM learning method improves reading fluency. Therefore, we believe that a prosodic training will not only help improving fluency when reading but also other traits of the oral 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, both by teachers and students has been shown to be beneficial in teaching and learning L2 pronunciation (Smotrova, 2017; Baker 2014, among others). Specifically, a recent study by Smotrova (2017) 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. In the study, some video-recorded classroom interactions from a reading class were analysed and results indicated that teacher's gestures facilitate the production of some suprasegmental features such as word stress and speech rhythm.

Some teaching proposals have emphasized the relevance of body language and gestures for learning pronunciation, suggesting that the combination of the non-verbal resources with the learning/teaching of the pronunciation of a second language can be very
effective. (see Smotrova, 2017 for a review). Hudson (2011) or Rosborough (2011) propose some instructional strategies based on gestures to teach segmental traits of pronunciation. Baker (2014) and Gilbert (2008) suggest that practices related with rhythm should be accompanied by a gesture, for example moving a foot, tapping the desk, or clapping to mark the rhythm. Brown (2014) recommends the use of beat gestures to reinforce tonic syllables, word stress and rhythm of stressed syllables.

However, despite this work, to our knowledge, there are very few experimental studies implementing a classroom training assessing the role of gestures in L2 pronunciation. To our knowledge, only two studies have been carried out with this idea in mind. Gluhareva and Prieto (in press) have proposed that beat gestures, that is, hand gestures that mark the rhythm of speech, are an effective aid for L2 acquisition of pronunciation. In a brief within-subjects training study, participants were asked to watch an English instructor producing a set of target sentences in English within a discourse situation either accompanied by rhythmic beat gestures or not. Twenty Catalan participants improved their accentedness significantly on the most difficult trained items only in the beat perception condition. A follow-up study is currently being carried out (Kushch, Gluhareva and Prieto 2017) and aims to extend the findings of gesture production by investigating whether participants show higher gains in accent improvement if they are instructed to imitate the experimenter and produce beat gestures themselves, rather than only observe them. In the experiment, 30 participants were randomly assigned to two groups. In the first group participants watched the videos in which an instructor gave spontaneous responses to English discourse prompts with accompanying beat gestures and had to orally repeat the utterances after the instructor (no-gesture production group). In the second group, participants watched the same videos and had to repeat the utterances, accompanying them with beat gestures. The results of this experiment showed that producing beat gestures had a significant beneficial effect in comparison to gesture observation alone.

Thus, more research on how gestures, specifically rhythmic beat gestures, might help to improve pronunciation skills, such as fluency, comprehensibility, expressiveness or accentedness, needs to be done. Our hypothesis is that beat gestures may help highlighting suprasegmental aspects of language and at its turn help learning some pronunciation aspects of the language. The property of beat gestures to make prosodic
structure more visible can play a facilitative role in the phonological awareness of suprasegmental features of pronunciation in the target language.

1.2.7 Why beat gestures can help pronunciation instruction?

The main aim of this thesis is to experimentally test the value of using beat gestures in the second language pronunciation classroom. There are two main features of beat gestures that strongly relate them with the phonological structure of the target language: (a) beat gestures act as visual highlighters of the prosodic structure of the language; and (b) beat gestures can have a positive effect on speech fluency, both in first and second languages.

On the one hand, with respect to the function of beat gestures as prosodic highlighters, beat gestures have been shown to be intrinsically coordinated with rhythmic and prosodic structure in speech (e.g., Loehr, 2007; Esteve-Gibert & Prieto 2014). There is a strong temporal connection between the presence of prosodic prominence (or pitch accentuation) and beat gestures. Typically, the most prominent parts of co-speech gestures (the gesture stroke or apex) are temporally aligned with prominent parts of speech (i.e., accented syllables) (e.g. McNeill, 1992; Kendon, 1980). Yasinnik et al. (2004) showed that during a narration more than 90% of instances of the gesture apexes occurred together with a pitch-accented syllable (see also Jannedy & Mendoza-Denton, 2005 for a review).

On the neural level, it has also been shown that different areas of the brain are activated depending on whether the speech was synchronised with beat gestures or with other stimuli that are not gestures, e.g. a dot moving on the screen (Biau, Fernandez, Holle, Ávila, and Soto-Faraco, 2016). The areas activated when a beat gesture is produced are the language areas of the brain, while when a dot in the screen appears only visual-perception areas are activated (Biau, Fernandez, Holle, Ávila, and Soto-Faraco, 2016). Therefore, beat gestures are directly related with the language-related areas of the brain. Moreover, the positive effects of beat gestures can be due to their direct role on attention, as shown in the ERP study done by Biau and Soto-Faraco (2013) where beat gestures were shown to help listeners to regulate the parsing of a stream of the speech and to focus their attention on the information conveyed in the gesture. Therefore, beat gestures function as attention getters related to language-related areas of the brain, in contrast with the visual non-gestural stimuli that activate visual-perception areas.
On the other hand, with respect to the positive effects of beat gestures on speech production and speech fluency, a study by Lucero, Zaharchuk & Casasanto (2014) investigated the effects of beat gestures on word production. They carried out two experiments. In the first one, they tested whether producing beats or iconic gestures can help speakers to produce words. Results showed that there was a beneficial effect from beat gestures when producing words when compared to the iconic and the no gesture conditions. In the second experiment, they tested 4 groups instead of three: No Gesture Instruction, Bimanual Beat, Right Hand Beat and Left Hand Beat, and participants were asked to produce words in one of these conditions. Results showed that when participants were instructed to produce beat gestures bimanually or with their left hand, the production of the word was quicker. Therefore, these results provide evidence that beat gesture can facilitate word production and thus we hypothesize that these gestures can help make a discourse more fluent.

2. GOALS OF THE DISSERTATION

Following an embodied cognitive approach, the aim of the current thesis is twofold, (a) to investigate how effective it is to highlight rhythmic features of the target language through beat gestures and prosody in the context of L2 pronunciation learning; and (b) to investigate the effects of a gesture-based classroom intervention study in pronunciation skills in the learning of English as a second language in high school students. One of the practical goals of the thesis is also to create an educational tool to teach pronunciation and oral communication skills more effectively, specifically when learning English as a second language, and which can be integrated into the high-school curriculum.

The target population for all the experiments in this thesis will be junior high school students (ESO 3 and ESO 4 courses, that is 14- to 16-year-old adolescents). They typically have A2 level English students (elementary level).
3. THEORETICAL FRAMEWORK

The present PhD thesis is couched within two theoretical frameworks, namely, gesture and speech as an integrated system and embodied cognition.

3.1 Gesture and speech as an integrated system

Speakers express their thoughts in two dimensions, speech and gesture. Research from the previous decades has shown that these two dimensions constitute a single communicative system that is tightly integrated semantically, pragmatically, and temporally (e.g., Kendon, 1980; Levinson & Holler, 2014; McNeill, 1992). There is clear evidence that gestures help people to clarify unclear, ambiguous, and/or difficult speech because they offer a second channel of communication (Goldin-Meadow & Alibali, 2013; Holler & Wilkin, 2009; Hostetter, 2011; Kelly & Church, 1997; see Goldin-Meadow & Alibali, 2013, for a review). In general, across studies, there is strong evidence that gestures boost listeners’ recall and comprehension of speech in both adult and children populations.

According to McNeill (1992), gestures can be classified into five main categories, namely iconic (or representational), deictic (or pointing, typically produced with the index finger extended), metaphoric (i.e., gestures representing speech content), emblems, and beat gestures (i.e., hand movements which are rhythmically integrated with speech but contain no referential content).

At the phonological level, the highest physical effort of gesture (the gesture apex or the gesture stroke) typically coincides with the most prominent part of speech. A particularly important role in marking the rhythm of speech is played by beats—“flicks of the hand(s) up and down or back and forth that seem to beat time along with the rhythm of speech” (McNeill, 2005, p. 40). Due to this function, beat gestures co-occur with stressed syllables and, apart from hands, can be produced with the head, eyelids, and other body parts (Efron, 1972; Loehr, 2007).

From an articulatory point of view, there are three models of gesture and speech production in adults that are important in the context of this thesis). Firstly, the
Growthpoint Theory (McNeill, 2005) which is presented as a unit that is formed by two simultaneous types of structuring meaning: “(1) spatial, analog, holistic, imagistic and (2) sequential digital, combinatorial, linguistic.” (Iverson, 2006, p.241). Therefore, this theory states that gesture and speech need to be considered in a jointly way. Secondly, the Interface Hypothesis by Kita & Özyürek (2003) is based on the fact that gestures originate from a spatio-motoric interface representation that is organised to enable speaking. According to this theory, it is speaking what imposes constraints on how the information should be organised. This approach claims that gestures are produced during the process that organises the spacio-motoric imagery to an optimal form of speaking (Kita & Özyürek, 2003; Kita & Özyürek, 2007). Thirdly, the sketch model (de Ruiter, 2000) assumes that gesture and speech have a shared communicative function and that, therefore, there is a common communicative intention (De Ruiter & De Beer, 2013). In contrast to the other models presented, this model not only takes into account iconic/metaphoric gestures, but it also includes deictics, emblems & pantomims, but no beat gestures. The three models presented focus on the representational aspects of gesture and do not consider gesture within a discourse and pragmatic dimension, and do not include beat gestures.

3.2 Embodied cognition

The grounded or embodied cognition theory (Barsalou, 2008) states that cognition is based in the sensory-motor processes and in the morphology and internal states of our body (Barsalou, 2008; Ionescu & Vasc, 2014). According to grounded cognition, modal simulations, bodily states and situation action are elements that underlie cognition. Some studies about grounded cognition highlight the big role played by the body in cognition, since states of the body can produce cognitive states and also can be the results of them (e.g., Barsalou et al., 2003; Lakoff & Johnson, 1980; Barsalou, 2008). Embodied cognition mixes mechanisms of perception and action and then perception, action and cognition are unified in the brain (Barsalou, 2010). While it is hardly surprising that exposure to gestures which include a semantic component can enhance memory and comprehension (as many studies have corroborated), less is known about the effects of gestures which do not carry referential meaning, such as beat gestures. In accordance with the grounded or embodied cognition paradigm (Barsalou,
and since gestures represent an embodied way to represent and highlight some aspects of language, we expect that actively involving sensory-motor perceptive and production processes through the use of beat gestures will trigger an increase in attentional networks and thus benefit the coding of linguistic information. Therefore, encouraging children to move their hands can only be beneficial for their learning. (Cook et al., 2008).

4. HYPOTHESES

The studies discussed in the preceding section suggest that the connection between the rhythm of speech and body movement may provide a theoretical basis for creating gesture-based pedagogies for teaching pronunciation. However, the findings in gesture studies relevant to pronunciation instruction have not yet been applied in L2 pedagogy, with a few exceptions as discussed above.

The present PhD thesis will be composed of four independent studies which will experimentally test the value of using beat gestures (rhythmic hand/arm movements produced in alignment with prominent prosody) and prosodic prominence in the second language pronunciation classroom. For each study, the specific research questions are the following:

STUDY 1: Does encouraging the production of beat gestures during reading enhance pronunciation skills of Catalan teenage learners of English?

STUDY 2: Does encouraging the production of beat gestures during the production of descriptions enhance pronunciation skills of Catalan teenage learners of English?

STUDY 3: Does encouraging the production of beat gestures and the delayed repetition of prosodic patterns in relevant pragmatic contexts improve the intonational skills of Catalan teenage learners of English?

STUDY 4: The idea for this study is to prepare a set of gesture-based classroom instructional tools for pronunciation and test their efficacy through a classroom intervention study. The main research question is whether a prosodic and gestural
classroom training with high-school students will enhance pronunciation skills in Catalan teenage learners of English.

5. EXPERIMENTAL STUDIES

This PhD project will consist of three experimental studies and one classroom-based intervention study which will assess the value of examines the teacher and student gesture employed in teaching and learning suprasegmental features of second language (L2) pronunciation. The first two experiments will provide evidence on the effects of encouraging students to produce beat gestures both in reading and in speaking tasks. The third study will investigate whether a delayed repetition of prosodic patterns in pragmatically relevant contexts will enhance the use of intonation in English as a L2. Finally, taking into account the results of the first three experiments, a gesture and prosodic intervention study will be performed with junior high school Catalan learners of English. Finally, an educational tool based on this proposed training will be put together. As mentioned before, all the experiments will be performed with junior high-school students in a high-school in Banyoles (Girona).

5.1 STUDY 1

5.1.1 Goals

The first experiment investigates whether encouraging the production of beat gestures enhances L2 pronunciation in a reading task. Our hypothesis is that speakers in the gesture-encouraged condition will have a better pronunciation when using their hands while reading than when they are not asked to produce any gesture. Thus, encouraging gestures may have a positive effect on reading a discourse in a more fluent and comprehensible way.

Participants took part in a between subjects experiment with two different conditions, namely gesture encouraged and no-gesture encouraged conditions, and they were randomly assigned to one of them.

The data for this study was collected during the months of February and March, 2017. Yet the pronunciation ratings have not yet been analysed.
5.1.2 Participants

A total of 59 teenage participants (29 females and 30 males) were recruited in a high school in Banyoles, Girona (INS Pere Alsius). Their mean age was fourteen-years-old ($M=14.08; SD=2.81$) and they were Catalan-dominant speakers ($M=95.159; SD=10.1270$). The parents of the participants filled in a prior consent to the participation in the study and they also stated that their children had a normal development and that they do not have any language disorder. Participants also filled in a language questionnaire and an anxiety questionnaire (Horwitz et al. 1986). According to the teachers, the general level of the students was an A2 (elementary level) according to the European Framework, which corresponds at the level taught at 3rd of ESO.

5.1.3 Experimental materials

Participants were given four different short fragments about several countries and cities and they were asked to read them aloud imagining that they were talking to someone. Both the task instructions and the four experimental items were presented in a power point slide (see Figure 1 for a sample). All fragments consisted of one long sentence with approximately 17 to 25 words. Target words in the fragments were double-checked with the teacher of the students participating to check that they could be comprehended by the students. The order of the texts was randomized so that participants read the fragments in different orders.

![Portugal](image)

Portugal is a country with excellent sandy beaches, a very pretty coast, cosmopolitan cities, and incredible weather.

Figure 1. Example of a sample slide with the text that participants needed to read.

---

5 We would like to thank the EFL teacher Glòria Coromina from INS Pere Alsius for the assessment of the materials for this test.
The instructions of the encouraged-condition were presented in a separate slide with the following text: "use your hands while speaking”.

5.1.4 Procedure

Participants were tested in a quiet room at the INS Pere Alsius. Even though the activity was part of the English as a foreign language lesson, participants were tested individually in a different room. The 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 was not looking at the participant while s/he was reading the texts. All of the experimental task was videorecorded with a Nikon d7000, located at 5 meters in front of the participant.

Before starting with the reading task, participants were told that they need to imagine that they were talking to someone that wanted to travel to the indicated place and they needed to read the text to this person, in a spontaneous and expressive way, to give information about the specific country or city. The specific instruction was as follows: Read the following texts silently for yourself. Imagine that you want to describe a set of countries and cities. Then read the following texts aloud as if you were reading them to him/her. Firstly, they were asked to read the text silently, for themselves once, and then to read it aloud.

In a between-subjects design, participants were randomly distributed into two conditions: gesture encouraged and no-gesture encouraged conditions (they were asked to participate in a reading task involving four short texts. The first two short texts were read aloud with no specific instructions, but the latter two were different depending on the condition. While in the no-gesture encouraged condition (control condition) they were asked to read the four texts without any gestural instruction, in the gesture-encouraged condition they were asked to read the latter two stories while moving their hands. Therefore, all participants read the first two texts without gesture encouragement and then only the gesture encouraged group was asked to use their hands while reading the last two fragments Figure 1 shows a summary of the experimental procedure.
Once the experiment was complete, participants filled out both the language questionnaire and the anxiety test, via Google forms.

A total of 236 oral fragments were obtained (59 students x 4 texts).

5.1.5 Ratings

All the participants’ audio recordings were rated by five native speakers of American English as this is the variety of English to which students are most exposed to, according to their ESL instructors. The native speakers are living in Catalonia but they are born in the USA. All of them reported normal hearing.

All the audios were uploaded to Survey Gizmo (https://www.surveygizmo.com) and a survey was created for the raters so that they could assess all the recordings in a random order. The raters analysed a total of 236 speech fragments and were asked to rate the following aspects of speech from 1 to 5: fluency, comprehensibility, expressiveness and accentedness. There was a thirty-minutes training session with the experimenter and the 4 native speakers so that they could practice the rating procedures for these aspects of speech. A working definition of those variables can be found in section 1.2.2.
5.1.6 Results

Perceptual assessment by the author of this study seems to point to the result that participants in the encouraging condition seem to improve their pronunciation (specifically their fluency and comprehensibility). Thus, we have initial indication that beat gestures might help enhance foreign language L2 pronunciation. Figure 2 shows still images (top) and a fragment of speech (bottom) of a participant uttering an item in the non-encouraged condition (left panel) and an item in the encouraged condition (right panel). At the prosodic level, we can observe that there is a flat intonation in the non-encouraged condition and that there is a pitch accent in the word weather and an expanded pitch range in the encouraged-condition.

Figure 3. Still images (top) and a fragment of speech (bottom) of a participant uttering an item in the non-encouraged condition (left panel) and an item in the encouraged condition (right panel).

5.2 STUDY 2

5.2.1 Goals

The aim of the second experiment is to assess whether encouraging Catalan learners of English to use gestures when producing short oral discourses in English enhances L2 pronunciation. As in the previous experiment, participants took part in a between
subjects experiment with two different conditions, namely gesture encouraged and no- 
gesture encouraged conditions, and they were randomly assigned to one of them. 
The data for this study was collected during the months of February and March, 2017. 
Yet the pronunciation ratings have not yet been analysed.

5.2.2 Participants
A total of 48 teenage participants (31 females and 17 males) were recruited in a high 
school in Banyoles, Girona (INS Pere Alsius). Their mean age was 15.2 (\(M= 
15.2; SD=38\)) and they were Catalan-dominant speakers (\(M=94.68; SD=12.60\)). Parents 
of the participants filled in a prior consent to the participation in the study and they also 
stated that their children had a normal development and that they do not have any 
language disorder. Participants also filled in a language questionnaire and an anxiety 
questionnaire (Horwitz et al. 1986). The activity was part of the English as a foreign 
language lesson but participants were tested alone in a different room. The level of the 
students was an A2 (elementary level) according to the European Framework which 
corresponds at the level taught at 4th of ESO.

5.2.3 Experimental materials
The materials consisted of 5 different PowerPoint slides, which were created on the 
basis of five different cities and countries. The five items were divided into one 
familiarization trial followed by four experimental trials. Each slide provided students 
with two concepts about a city/country, the name of the city/country and a picture of it. 
First, an introductory slide (Figure 3) explained to them how the task was going to be 
and that they needed to use the two provided concepts presented in the slide while 
describing the city/country (see Figure 4 for a sample slide). They were asked to 
imagine that they were describing these places to a person that was interested in 
travelling. Before starting with the experimental trials (4 slides) they were asked to 
familiarize themselves with the task with the first trial.
5.2.4 Procedure

Participants were tested in a quiet room at the INS Pere Alsius. Even though the activity was part of the English as a foreign language lesson, participants were tested individually in a different room. The 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 was not looking at the participant while s/he was reading the texts. All the experimental task was videorecorded with a Nikon d7000, located at 5 meters in front of the participant.

First, participants were asked to carefully read the instructions presented in the first slide. The specific instruction was as follows: *You need to briefly describe a set of countries and cities so that a person that wants to travel there know some traits of these places. Do it very briefly and use the provided concepts.*
countries and cities so that a person that wants to travel there can know some traits of these places. They were told that they needed to briefly describe a set of countries and cities to an interested party. They were also told that two important concepts about these cities/countries will be provided so that they can easily talk about the target places.

Participants were randomly distributed in the two between-subject groups, namely gesture-encouraged group and no-gesture encouraged group, corresponding to the two conditions of the experiment. The participants in the no-gesture encouraged one were asked to produce the four stories without any gestural instruction. By contrast, participants in the gesture encouraged group were asked to produce the first two discourses with no gestural instructions and then, in the following two, they were asked to use their hands before producing the brief discourse. The gesture-encouraging instructions for the encouraged-condition were presented in an isolated slide with the following text: "use your hands while speaking". Figure 6 shows a summary of the experimental procedure.

![Figure 6. Summary of the experimental procedure](image)

Once the experiment was complete, participants filled out both the language questionnaire and the anxiety test, via Google forms.

The typical length of the speech fragments produced by students was of a total of 2-3 sentences. A total of 192 oral fragments were obtained (48 students x 4 texts).
5.2.5 Ratings

All the participants' audio recordings will be rated by five native speakers of American English because this variety of English is the one that are most exposed to the students, according to their ESL teachers. The native speakers are living in Catalonia but they were born in the USA.

All the audios will be uploaded to Survey Gizmo (https://www.surveygizmo.com) and a survey was created for the native speakers so that they could rate all the recordings in a random order. The native speakers will analyse a total of 192 speech fragments and they will be asked to rate the following aspects of speech from 1 to 5: fluency, comprehensibility, expressiveness and accentedness. A working definition of those variables can be found in section 1.2.2. There was a thirty minute training session with the experimenter and the 4 native speakers so that they could practice the rating procedures for these aspects of speech.

5.2.6 Results

Perceptual pilot results from 10 participants show that when participants are asked to use their hands they produce a more fluent and comprehensible discourse, thus their pronunciation seems to be better. Moreover, as in experiment one, these preliminary results also show that when participants are asked to "use their hands" while producing the discourse, they tend to produce more beat gestures (rhythmic hand/arm movements produced in alignment with prominent prosody). Thus, we have more evidence to believe that beat gestures enhance English pronunciation both when reading and producing a discourse. See Figure 7 for a still image of a participant uttering an item in the non-encouraged condition (left panel) and an item in the encouraged condition (right panel).

![Figure 7. Still images of a participant uttering an item in the non-encouraged condition (left panel) and an item in the encouraged condition (right panel).](image-url)
5.3 STUDY 3

5.3.1 Goals

The aim of this study is to investigate whether encouraging the production of beat gestures in a delayed discourse imitation task within a pragmatically relevant context can improve L2 intonational patterns.

To our knowledge, there are no integrative methods that take into account the role of the communicative goals of the speaker, together with the rhythmic highlighting of the sound structure of the target language through body gestures and prosody. Studies by Morley (1994) stated that, for example, communicative approaches should be also applied to pronunciation teaching, and that the focus on suprasegmentals and how they affect in communicating the meaning of the discourse should be more important.

A between-subjects experiment will be conducted with a delayed imitation in four different conditions, namely beat encouragement and non-beat encouragement in context discourse, and beat-encouragement with isolated sentences and non-beat encouragement with isolated sentences. Participants will be randomly assigned to one of the four conditions.

5.3.1 Overall method

A discourse-based delayed imitation task targeting a variety of L2 intonation patterns will be designed.

5.2.2 Participants

Fifty teenage students from an ESO-3 class (14 and 15-year-olds) will take part in the experiment. They will be Catalan native speakers and their parents will fill in a consent form of participation.

5.2.3 Experimental materials

The materials for the delayed imitation task targeting a variety of intonation contours will be created on the basis of the target videorecorded discourses provided in the Speaking Skills Practise of the learnenglish webpage of the British Council for
teenagers\textsuperscript{6} (see Appendix 1 for an example of the whole dialogue). After watching these short videorecorded dialogues, participants will be prompted to imitate the target intonation phrases by one of the characters in the role play. A sample dialogue is reproduced in figure 8:

![Image](image_url)

Figure 8 Example of the dialog “Buying new shoes”.

In order to assess the importance of the contextual discourse, some isolated sentences from the same fragments will be used in the control group (based on Llanes, Mora & Serrano, 2016).

The pre-test and post-test materials will consist of the reading of the training dialogues of the Speaking Skills Practise of the learnenglish webpage of the British Council for teenagers (http://learnenglishteens.britishcouncil.org/skills/speaking-skills-practice).

5.2.4 Procedure

\textsuperscript{6}http://learnenglishteens.britishcouncil.org/skills/speaking-skills-practice
Participants will be asked to perform a pre-test, a training and a post-test. In the pre-test and post-test phases, students will be asked to read a set of target phrases of dialogues that will be trained in the training phase. These pre-test and post-test oral fragments will be later assessed by external assessors (native speakers of English). In the training phase, students will be randomly assigned to one of the four different groups: beat-encouragement and non-beat encouragement group, either within a pragmatic context or without it. In the gesture encouraged condition, participants will be asked to use their hands when reading the dialog; in the no-gesture encouraged condition, participants will have no instruction (see figure 9 for an example of the materials they will have to read).

We expect that the gesture encouraging group participating in a delayed imitation task with complete dialogues will obtain better pronunciation accuracy scores than the other three groups.

![Figure 9. Example of the task](image)

### 5.4 STUDY 4

#### 5.4.1 Goals

The fourth experiment will be the basis for creating a tool for pronunciation instruction based on the encouragement in the production of beat gestures and prosody. This classroom intervention will consist of a set of embodied pronunciation training activities which will be experimentally tested. For example, one of the training tasks will consist of exercises based on practicing mainly the rhythm of the L2 with beat gestures and prosody.
5.4.2 Participants

Sixty students of 16-years old (fourth of ESO) will take part in the experiment. They will be Catalan learners of English in the INS Pere Alsius high-school in Banyoles. The parents of the participants will fill in a consent form of participation.

5.4.3 Experimental materials and procedure

The intervention study will have a pre-test-intervention-post-test design. The pre-test and the post-test will consist of a short oral production elicitation task the topic of which will be to explain their favourite book as the pre-test and to explain their favourite series/film as the post-test. These oral fragments will be later assessed by external raters (native speakers of English).

For the intervention phase, groups of students will be randomly assigned to two different groups: either a beat-encouragement or non-beat encouragement group. For each of the two groups, the intervention will consist of 15-minute sessions during 5 weeks in the ESL classroom at the high-school. The intervention materials will be based on the results of the previous studies carried out in this thesis, as well as on several proposals of pronunciation learning books and on several studies produced on pronunciation instruction. A good study to take into account for the activities of the training is the paper by Kendrick (1997). In their study, they performed several activities for improving students' L2 pronunciation such as exercises in which they were asked to discriminate and produce specific segments to be aware of the sounds and to be able to self-correct; exercises to develop awareness about weak syllables; activities to make them aware of speech rhythm; exercises on stress of prominent words; activities in drama and role-play intonation, etc. The book by Estebas (2009), which has many activities on both segmentals and suprasegmentals, will also be useful as a baseline. The study of Derwing, et al. (2014) also includes many activities focusing on the prosody of English language and that were effective in enhancing comprehensibility, fluency and accentedness. All these materials will constitute the basis to start preparing a five-week embodied training to enhance pronunciation and oral-skills of the students.
### 6. WORKING SCHEDULE

| March-June 2017 | - Data collection for Studies 1 and 2  
- Workshops for 3rd and 4th ESO students and 1st of BATXILLERAT. INS Pere Alsius i Torrent: Banyoles (Girona). Title: Oral abilities in L2 classrooms. Sessions: 7 sessions of 50 min in 14/03/2016, 15/03/2016; 16/03/2016, 17/03/2016  
- 3-month stay abroad (Università degli Studi di Padova; Host: Dr. Maria Grazia Busà)  
- Writing of the Ph.D. project  
- Submission of the Ph.D. project  
- Poster presentation at the conference *Language as a form of Action*. Institute of Cognitive Sciences and Technologies (ISTC): Roma (Italy), June 21-23. Title: Do beat gestures help preschool children to recall and understand discourse information? Coauthors: Vilà-Giménez, I., Kushch, O., Borràs-Comes, J., & Prieto, P. |
| July 2017 | - Oral presentation of the Ph.D. project  
- Poster presentation at the conference *International Congress for the Study of Child Language* (IASCL): Lyon (France), July 17-21. Title: Prominence in speech and gesture help preschoolers to recall and comprehend information. Coauthors: Vilà-Giménez, I., Kushch, O., Borràs-Comes, J., & Prieto, P. |
| August 2017 to December 2017 | - Analysis and writing of Studies 1 and 2  
- Design and data collection for Study 3  
- Oral presentation at the conference *European Second Language Association* (EuroSLA): Reading (UK), August 30 - September 2. Title: The effects of encouraging rhythmic beat
gestures in second language reading pronunciation.

Coauthor: Prieto, P.

- Poster presentation at the conference *Architectures and Mechanisms of Language Processing* (AMLAP): Lancaster (UK), September 7-9. Title: Does encouraging the production of beat gestures enhance L2 pronunciation? Coauthor: Prieto, P.

- Poster presentation at the conference *Architectures and Mechanisms of Language Processing* (AMLAP): Lancaster (UK), September 7-9. Title: Do beat gestures and prosodic prominence enhance preschoolers' recall and comprehension of discourse information? Coauthors: Vilà-Giménez, I., Kushch, O., Borràs-Comes, J., & Prieto, P.

| January 2018 to March 2018 | - Writing of Study 3 |
| April 2018- June 2018 | - Potential 3-month stay abroad to collaborate in Study 4 (preparation of the linguistic materials for Study 4). Host: Dr. Veronica Sardegna. |
| | - Attendance to the congress International Society for Gesture Studies (ISGS) in Cape Town (South-Africa). |
| July 2018 – December 2018 | - Data collection and analysis for Study 4 |
| January 2019 – May 2019 | - Writing of Study 4 |
| | - Creation of educational tool based on the intervention study |
| | - Attendance to the International Congress of Phonetic Sciences (ICPS) in Melbourne |
(Australia). 4-10 August, 2019.

<table>
<thead>
<tr>
<th>May 2019-August 2019</th>
<th>- Writing of the Ph.D. dissertation</th>
</tr>
</thead>
<tbody>
<tr>
<td>September 2019</td>
<td>- Dissertation defence</td>
</tr>
</tbody>
</table>

7. SELECTED REFERENCES


This study focuses on describing a pronunciation training for an intensive English program on communication for different levels. In their article, they specify which segmental and suprasegmental elements should be taught at different levels. For example, for beginners they recommend teaching basic phonetics features such as vowel length, consonants and final consonants and clusters, as well as practising the alphabet. For intermediate students the segmental part and basic intonation, intonation and sensitize to stress-timing and stress perception for suprasegmentals. They also carried out several interviews to English teachers to understand the status of pronunciation instruction practice. With the program that they propose students learn that pronunciation needs to be learnt from the beginning and that it is necessary for intelligibility and comprehensibility. Moreover, they state that pronunciation should not be taught separately from the rest of the language instruction and should be integrated in each lesson.


Sardegna (2009) performed a longitudinal study on the long-term effectiveness of pronunciation instruction, specifically on suprasegmentals. In her study, she assessed
the progress of the students in English primary phrase stress (when a word in the sentence is more stressed than the others), constructions stress (when a multi-word, like compound nouns, is more stressed than the others) and word stress (when a word is stressed more than the others) has during a semester of a pronunciation course. Results showed that the instruction of pronunciation learning strategies was effective for the improvement of the read discourses in primary stress, construction stress and word stress.


Gordon & Darcy (2016) performed a classroom study to assess the effects of short-term pronunciation instruction. They investigated the effects of explicit and non-explicit pronunciation instruction on comprehensible speech. They had three different groups of participants in the training test: pronunciation instruction on suprasegmental features; pronunciation instruction on four vowel sounds; and without explicit instruction. Results showed that the group that was instructed on suprasegmental features had a higher comprehensibility ratings.


Derwing, Munro & Wiebe 1998 explored three types of pronunciation teaching methods on fluency, comprehensibility and accentedness ratings. They tested three methods: segmental, global (i.e., suprasegmentals) and no pronunciation-specific instruction. Results show that while segmental and global training have improvements in comprehensibility and accentedness in a sentence-pronunciation exercise, only global group improve with spontaneous speech.

The study by Gluhareva & Prieto (2016) aimed to assess whether observing rhythmic gestures, specifically rhythmic beat gestures, can reduce the accentedness of the Catalan learners of English when producing a discourse. In a between-subject design, participants in the beat condition had to watch several videorecorded English discourses while a native speaker was producing beat gestures; participants in the no-beat condition watched the same videorecorded discourses produced without gestures. The pre-test and post-test sample accentedness analysis revealed that the training with beat gestures reduced the accentedness ratings of the participants.
8. GENERAL REFERENCES


