UCLA
UCLA Electronic Theses and Dissertations

Title

An Audio-visual Approach to Spanish Trill Acquisition in Adult Learners: A Longitudinal Multilevel Study of Phonetic Development

Permalink

https://escholarship.org/uc/item/7j7558tn

ISBN

9798247932031

Author

Melendez-Ballesteros, Nancy

Publication Date

2026-05-29

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

An Audio-visual Approach to Spanish Trill Acquisition in Adult Learners: A Longitudinal Multilevel Study of Phonetic Development

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Hispanic Languages and Literatures

by

Nancy Meléndez-Ballesteros

2026

© Copyright by

Nancy Meléndez-Ballesteros

2026

ABSTRACT OF THE DISSERTATION

An Audio-visual Approach to Spanish Trill Acquisition in Adult Learners: A Longitudinal Multilevel Study of Phonetic Development

by

Nancy Meléndez-Ballesteros

Doctor of Philosophy in Hispanic Languages and Literatures

University of California, Los Angeles, (2026)

Professor Ji Young Kim, Chair

Imitation, the ability to copy another person's gestures, acoustic features, or speech style, is a cognitive skill present throughout life and plays an important role in language learning and acquisition. Evidence from first language research shows that infants begin imitating speech patterns early in development. This capacity continues into adulthood, particularly during conversational interaction or when individuals are

explicitly instructed to imitate a model. The present study investigates a method for teaching second language (L2) phonology to adult English-speaking learners of Spanish through explicit focused-attention imitation training. The method builds on a pilot study (Meléndez-Ballesteros, 2012) with students enrolled in a second-semester elementary Spanish course. Participants were divided into four groups and performed specific listening and speaking exercises. The current study focuses on the audio-visual exercise of that study. Students were randomly assigned to either a training or non-training group and completed the same audio-visual activities following two separate sets of instructions. The training group received explicit focused-attention imitation instructions, whereas the non-training group repeated the model's speech without explicit imitation instructions. The phonetic training lasted sixteen weeks. The effectiveness of the training program and developmental trends in students' pronunciation were evaluated through acoustic analysis of the segmental duration of the Spanish trill [r], one of the most difficult sounds for L2 learners to acquire. Its duration was measured as a function of time in word-initial and word-medial positions, including intervocalic contexts, where these phones exhibit a phonemic contrast. The patterns of change over time in students' productions of the Spanish trill were measured acoustically at four different time points throughout the semester. A growth-curve model within an HLM framework was used to examine individual learning trajectories over time, revealing variability in learners' development. Although statistically significant group-level differences between the training and the non-training groups were not consistently observed, the longitudinal analyses revealed variability in individual trajectories, with some learners exhibiting changes in production and differing

rates of change from baseline. The present study aims to contribute to the body of research on the explicit teaching of L2 speech sounds, focusing on adult learners of a second language. Unlike most studies of this type, it examines underlying individual differences that may contribute to the effectiveness of the training program by measuring students' performance at four different time points rather than relying solely on the conventional two (i.e., pre- and post-measurements).

The dissertation of Nancy Meléndez-Ballesteros is approved.

Jorge Marturano

Carlos A. Quicoli

Belén Villarreal

Ji Young Kim, Committee Chair

University of California, Los Angeles

2026

Dedication

Para mi inolvidable Wilson,

Esta tesis doctoral es para ti. Si no hubiese sido por tu constancia y paciencia, no habría llegado hasta aquí. Fuiste y siempre serás el amor de mi vida. Jamás olvidaré el apoyo y cariño que me diste a través de todos estos años; cuando necesitaba un empujón, ahí estabas tú con tus chistes y lindas palabras.

Tu manera de ser, tu amor y tu apoyo permanecerán en mí.

To my son, Nicolás, thank you for always pushing me to finish this dissertation. During the last months of the writing process, your hugs, pats on my back, and words of encouragement helped me continue.

A Mami y Papi, gracias por apoyarme siempre y por creer en mí.

Papi, aunque ya no estés aquí, nunca olvidaré lo orgulloso que te sentías de mis estudios y cómo le contabas a todo el mundo sobre esta tesis doctoral. Mami, por tu apoyo incondicional desde Puerto Rico y por siempre estar presente, aun en la distancia.

Acknowledgments

I would like to thank my dear departed professor and friend, Claudia Parodi, who left us too soon. Her undeniable dedication to her students is unforgettable. She was there from the beginning of this research project and truly wanted me to succeed with the proposed teaching methodology. I will never forget her words: “Nancy, acaba esta tesis para que nos hagas ricas :)”. She made me believe in myself, and that is something I now carry into all my classes and interactions with my students and advisees.

I would also like to thank Carlos A. Quicoli for his thoughtful and insightful comments, careful reading of this dissertation, and continued support throughout this process. I am especially grateful that he introduced me to Ji Young Kim, whose guidance became invaluable during the final stages of this dissertation.

I am deeply grateful to Ji Young Kim, who I only wish I could have met the moment she entered the department. I am so grateful she was placed in my path. Her knowledge of rhotics and heritage speakers is remarkable, and her dedication to her students is equally admirable. She met with me outside of working hours to help me organize what I felt was a mess with my data. Her guidance and patience helped me tremendously during the final stages of this dissertation.

To Belén Villarreal, thank you for being there from the very beginning of this project. Your revisions, feedback, and attention to detail were truly helpful. Accessibility was one of the things we both thought was important, so your suggestions on places

where my writing could be more accessible to readers outside the fields of linguistics and statistics were extremely helpful.

I would also like to thank Jorge Marturano for his thoughtful comments, encouragement, and for making time to participate in my dissertation defense during such a busy period. His feedback and enthusiasm toward this work meant a great deal to me.

I would like to thank Maritza Fonseca, whose true love for writing often made me step back and refine sections of this dissertation. I will always remember her words: “Si se puede decir con menos palabras, mejor”. Many times, she helped me see things I no longer noticed after rereading the same chapters.

I would also like to thank Halle Gauthier, a talented artist, for creating the tongue position illustrations included in this dissertation.

To Cristina and Liliana, my dear friends, thank you for being there for me during these last stages of writing. Cristina, hearing you say how proud you were of me after reading the results, really filled my heart. And Liliana, thank you for always telling me, “ya falta poco, tú puedes,” during the moments when writing felt endless.

I would also like to thank Montserrat Reguant and Alejandro Lee, former colleagues and dear friends, for taking the time to read portions of this dissertation throughout the writing process. Their encouragement, feedback, and continued interest in this study meant a great deal to me.

And finally, to Cristina Moon and Lorena Gauthier, thank you for your friendship, encouragement, and for always reminding me to continue and finish this dissertation. Lorena, your “Sí se puede” stickers always arrived at the right moment.

Table of Contents

LIST OF TABLES ..... *xiii*
LIST OF FIGURES ..... *xv*
VITA ..... *xvii*
Chapter 1 ..... *1*
1.1 Introduction **1**
1.2 Second Language Acquisition & Instruction **2**
1.3 L2 Phonology Acquisition in Late Learners **5**
1.3.1 Cognitive Processes in Acquisition of L2 Phonology: The Role of Perception 5
1.4 Age Effect on the Acquisition of L2 Phonology **8**
1.4.1 The Critical Period Hypothesis 8
1.5 CPH and Universal Grammar (UG) **10**
1.6 Counterevidence to the CPH and Loss of Neuroplasticity **12**
1.7 Relevance of the Pilot Study to Current Work **15**
1.7.1 Methodology of the Pilot Study 15
1.8 Organization of the Dissertation **16**
Chapter 2 ..... *18*
2.1 Introduction **18**
2.2 Memory and L2 Language Learning **20**
2.2.1 The Neural Basis of Memory and L2 Phonology **23**
2.3 The Powerful Cognitive Process: Imitation **25**
2.4 Current Study and Convergence Effects **32**
Chapter 3 ..... *33*
3.1 Introduction **33**
3.2 Other Possible Predictions for L2 Acquisition **34**
3.3 Pronunciation Instruction and Phonetic Training **36**
3.3.1 Phonetic Training Focused Studies 37
Chapter 4 ..... *40*
4.1 Introduction **40**
4.2 General Description of Rhotics **40**
4.2.1 Dialectal Variation and Target Form 41

4.3 Spanish Rhotics **42**
4.3.1 Articulatory and Acoustic Properties 43
4.4 Previous Literature on Phonetic Variation and Distribution of Rhotics **46**
4.5 Pilot Study **57**
Chapter 5 ..... ***63***
5.1 Introduction **63**
5.2 Research questions and Hypotheses **63**
5.2.1 Hypotheses 64
5.3 Participants **65**
5.3.1 Participant Background 65
5.4 Consent **67**
5.5 Materials and Methods **68**
5.5.1 Description of Tokens 68
5.5.2 Training Phase 69
5.5.3 Analysis Process 72
5.6 Acoustic Measurements **74**
5.7 Brief Description of the Statistical Analysis **76**
5.8 Hierarchical Data Structure and Statistical Analysis **79**
5.8.1 Rationale and Motivation for Employing an HLM Approach 79
5.8.2 Model Specification 82
5.8.3 Description of Coefficients in the Present Study 87
5.8.4 Centering 89
Chapter 6 ..... ***93***
6.1 Introduction **93**
6.2 Descriptive Overview of the Data **93**
6.2.1 Preliminary Exploratory Data Analysis (EDA) 93
6.2.2 Descriptive Overview of the Full Dataset 95
6.3 HLM Results **99**
6.3.1 Unconditional Means Model 99
6.3.2 Unconditional Growth Model (No Level-2 Predictors) 101
6.3.3 Training Predicting Initial Status 101
6.3.4 Full Model (Training and Predictors of Growth) 102

6.3.4.1 Full Growth Model Including Baseline Duration 104
6.3.4.2 Full Growth Model Including Heritage Speaker Status and Training Condition 106
6.3.5 Additional Analyses 107
6.3.5.1 Distance from Model Speaker 107
6.3.5.2 Distance from Model Speaker Including Heritage Status 109
6.3.5.3 Occlusions as Outcome (Poisson Model) 110
6.4 Summary of Findings 113
Chapter 7 117
7.2 Development Over Time and Individual Variability 122
7.3 Training and Imitation 126
7.4 Nature of the Trill and Rhotic Variability 130
7.5 Longitudinal Contribution 131
7.6 Pedagogical Implications 132
7.7 Conclusion 133
Bibliography 136
Appendices 148
Appendix A. Participant Information 148
Appendix B. Additional Articulatory Illustrations of Rhotic Variation 151
Appendix C. Training Materials and Instructions 152
Appendix D. Preliminary Pilot Study Graphs and Spectrographic Representations. 157
Appendix E. Consent Form 158
Appendix F. Token List 159
Appendix G. Data Structure and Variable Coding 160
Appendix H. Unconditional Means Model (Model 1) 164
Appendix I. Unconditional Growth Model: No Level-2 Predictors (Model 2) 166
Appendix J. Training Predicting Initial Status (Model 3) 169
Appendix K. Full Conditional Model: Training and Predictors of Growth (Model 4A). 172

Appendix L. Full Growth Model Including Baseline Duration (Model 4B) 175

LIST OF TABLES

Table 2.1. Linguistic Characteristics of LAVS Speakers 26

Table 4.1. Summary of Previous Studies on Spanish Rhotics 50

Table 4.2. Summary of Pilot Study Tasks 53

Table 5.1. Tokens Used in the Trill Production Task 68

Table 5.2. Assessment Time Points During the Semester..... 71

Table 5.3. Variables Included in the Hierarchical Linear Model 85

Table 5.4. Interpretation of Coefficients Used in the Hierarchical Linear Model 87

Table 6.1. Mean Trill Duration (ms) by Assessment Time (Exploratory Analysis) 94

Table 6.2. Mean Trill Duration (ms) by Time Point and Group 95

Table 6.3. Unconditional Means Model 99

Table 6.4. Unconditional Growth Model Estimates 100

Table 6.5. Training Predicting Initial Status 101

Table 6.6. Full Model Predicting Trill Duration 103

Table 6.7. Full Growth Model Including Baseline Duration 104

Table 6.8. Full Growth Model Including Heritage Speaker Status and Training Condition 106

Table 6.9. Conditional Growth Model Predicting Distance from the Model Speaker... 107

Table 6.10. Conditional Growth Model Predicting Distance from the Model Speaker Including Heritage Status..... 108

Table 6.11. Poisson HLM Model Predicting Occlusion Counts Over Time..... 109

Table 6.12. Poisson HLM Model Predicting Occlusion Counts as a Function of Time, Trill Duration, Heritage Speaker Status, and Training Condition……..…..111

LIST OF FIGURES

Figure 2.1. Adaptation of Atkinson and Shiffrin's (1968) Information Processing Framework 21

Figure 2.2. Relationship Between Bias and Likelihood of Imitation 31

Figure 4.1. Spectrographic representation of the Spanish trill and tap in carrera 'career'..... 44

Figure 4.2. Position of the tongue tip at the alveolar ridge during the production of the Spanish trill 45

Figure 5.1. Wimba platform used to deliver the audio-visual pronunciation task and record student productions 63

Figure 5.2. Spectrographic representation of the production of a trill by a heritage speaker 70

Figure 5.3. Spectrographic representation of the production of a trill by an L2 speaker 72

Figure 5.4. Preliminary Individual growth trajectories in trill duration at each time point 73

Figure 5.5. Individual growth trajectory at each time point after training 77

Figure 5.6. Individual growth trajectory at each time point. No training 77

Figure 5.7. Sequence of hierarchical linear models estimated in the study..... 77

Figure 6.1. Individual trajectories of trill duration across four time points 96

Figure 6.2. Distribution of trill duration (ms) across time points 98

Figure B1. Bunched rhotic realization 150

Figure B2. Retracted rhotic realization observed in some Caribbean Spanish varieties 150

Figure C1. Language Background Questionnaire 151

Figure C2. Example of an oral assessment task in Horizon Wimba using Voice Authoring 152

Figure C3. Example of a stress perception activity administered through Horizon Wimba 153

Figure C4. Example of a stress perception activity administered through Horizon Wimba 154

Figure C5. Screenshot of the Horizon Wimba audio-visual task used during the Focused-Attention Training (F.A.T.) activities 155

Figure C6. Screenshot of an audio-visual oral activity in Horizon Wimba……155

Figure D1. Preliminary descriptive graph of rhotic productions from the pilot study 156

Figure D2. Spectrographic representation of a learner's trill production before and after training 156

VITA

Education

- Doctoral Candidate, Hispanic Languages and Literatures, University of California, Los Angeles
- M.A., Spanish Linguistics, University of California, Los Angeles

Academic Appointments

- Chair, Language and Culture Department, Mount Saint Mary's University, Los Angeles
- Assistant Professor of Linguistics, Intercultural Communication, and Spanish, Mount Saint Mary's University, Los Angeles

Research and Professional Interests

- L2 phonology acquisition
- Spanish rhotics
- Phonetic training and imitation
- Longitudinal analysis and Hierarchical Linear Modeling (HLM) in L2 phonological development
- Heritage speakers
- Medical Spanish
- Open Educational Resources

Publications

- *Tarea Libre: A First-year Spanish OER Homework Ancillary, Open Educational Resource, with Cristina Moon, Alejandro Lee, and Sarah Harmon, 2023.*
- *Entrada Libre: Intermediate/Advanced Spanish Manual, Open Educational Resource, with Cristina Moon and Alejandro Lee, 2020.*
- Colorado, H. A., Gutiérrez Velásquez, E., & Ballesteros, N. (2022). Bridging the gap between theory and active learning: A case study of project-based learning in introduction to materials science and engineering. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje, 17*, 160–169. https://doi.org/10.1109/RITA.2022.3166862
- Ballesteros, N. (2014). Why teaching pronunciation to Spanish L2 learners matters. *Voices, 2*(1). http://escholarship.org/uc/item/51q604qx
-

Conference Presentations / Workshops / Invited Talks

- Hosted the AATSP/SoCal Fall Conference at Mount Saint Mary's University – Los Angeles, 2023-2024.

- *Tarea Libre*: *A First-year Spanish OER Homework Ancillary*, presented with Alejandro Lee at the AATSP/SoCal Spring Conference, California State University Northridge, 2023.
- ASCCC OERI Launching of *Tarea Libre* with Cristina Moon, Alejandro Lee, and Sarah Harmon, 2023.
- *How We Created a Spanish OER Grammar Manual*, CCCFLC Conference (online), with Cristina Moon and Alejandro Lee, 2021.
- *Spanish OER*, webinar organized by the Academic Senate for California Community Colleges Open Educational Resources Initiative (ASCCC OERi), 2021.
- *Libretexts and H5P: How We Created a Spanish OER Grammar Manual*, Open Education Week and ACTFL Community College SIG Coffee Talk, 2020-2021.
- *L2 Speech Imitation or Phonetic Convergence within a Traditional and Online Teaching Format: Implementing the Latest Technology to Achieve Optimal Learning*, AATSP SoCal Fall Conference at UCLA, 2013.
- *Does it sound native? Assessing students' performance in traditional and online environments*, 94th Annual Conference AATSP, San Juan, Puerto Rico, 2012.
- *Why Teaching Pronunciation to L2 Learners Matters*, University of California, Los Angeles, 2012.

Professional Service

- President, AATSP Southern California Chapter, 2024-2025
- Project Lead, Medical Spanish OER

Chapter 1

Introduction and Background Literature

1.1 Introduction

Many reasons may prevent second language instructors at the college and university levels from integrating pronunciation within a teaching curriculum. These may include lack of enthusiasm towards teaching pronunciation, inexperience in teaching specific phonetic features that help students develop clearer pronunciation, and the existence of an already saturated curriculum that may discourage instructors from adding additional components. In some cases, pronunciation is also considered less important than other language skills (Elliot, 1995) and some instructors may still be firmly committed to the belief that after a certain age full attainment is impossible.

The present project was motivated by a pilot study that was conducted in 2014 (Meléndez-Ballesteros, 2014), which focused on the significance of pronunciation instruction in second language learning, specifically for Spanish L2 learners. The study argued that emphasizing pronunciation in the classroom helps improve the perception and production of a native-like or near-native pronunciation in Spanish as a second language. By focusing on novel phonetic features not present in the student's first language, mainly English, learners could develop greater phonetic accuracy and approximate more native-like

pronunciation patterns, which helps them communicate effectively and confidently in Spanish, particularly in an increasingly globalized world where effective communication is crucial.

The pilot study showed that students who completed the imitation exercises demonstrated improvement in their productions of Spanish rhotics compared to those who did not. The present study aims to expand on this research by exploring the explicit teaching of L2 speech sounds to adult learners. It also attempts to unveil the individual differences that contribute to the effectiveness of the training program. To do so, students' performance is assessed at four distinct time points rather than through traditional pre- and post-measurements. This allows the analysis to capture individual developmental trajectories over time rather than relying solely on group averages. Moreover, it addresses new research questions that emerged from further statistical and acoustic analyses of the pilot study data.

To this end, addressing these challenges and developing effective pronunciation training programs requires a general understanding of second language acquisition and instruction. The next section discusses the main concepts of second language acquisition, how pronunciation should be included in the teaching curriculum, and the cognitive processes involved in language learning.

1.2 Second Language Acquisition & Instruction

During the last thirty years or so, much of the research in the field of second language acquisition (SLA) and second language instruction for the adult

learner of Spanish focused mainly on the teaching of general aspects of a second language (L2) (i.e., grammar, vocabulary, orthography, etc.) without paying much attention to pronunciation. This inadvertently contributed to difficulties in acquiring aspects of the L2 phonological system, thus resulting in speakers who exhibit moderate to heavy foreign-accented speech in Spanish.

Within the last few decades, foreign language teaching methods have been employed in classrooms that have not rendered favorable results on the accuracy or full attainment of the targeted phonological system. Many of these methods, whose goal is to teach general communication skills, do not account for the diverse learning styles L2 learners bring into the classroom. These communicative-based methods have been considered too dogmatic and cannot be used in different learning settings (e.g., Brown, 2001; Kumaravadivelu, 2003, 2005).

It is generally accepted within the teaching field that to be proficient in any language, a language learner should attempt to be fluent and accurate in various aspects of the language, including morphosyntax and phonology. In the context of the present study, accuracy refers specifically to phonological accuracy. It reflects the degree to which a learner's production approximates target-like pronunciation, specifically in the production of the Spanish trill. In other words, if a student is said to be fluent but not accurate, it means that the student can manage the general aspects of a language (i.e., expressing ideas, maintaining interaction, and using appropriate expressions), but may have

pronunciation-related issues that can obstruct the communication process. On the other hand, if the student is accurate but not fluent, it means that the student can produce sounds that are closer to the target language but may still struggle to express themselves with ease and maintain the flow of the conversation effectively.

Until recently, little was said about developing second language training programs for the adult population focusing on *pronunciation*. Although some have been incorporated, they differ in the scope of training. In other words, they differ in the determinant of optimal proficiency, how the training should be delivered to adult L2 students, and how long the pronunciation programs should be, among other things.

Some educators believe *that form should not be the focus of instruction but rather of* overall understanding, which follows Levis's *Intelligibility Principle* (2005). On the other hand, due to the differences in L2 adult needs, other educators prefer to focus on form to approximate a more native-like production. In either case, this must be addressed because many adult L2 students within foreign language classrooms are highly interested in learning how to speak the L2 clearly, with no noticeable accent, so that they can use it at a higher educational or professional level (i.e. business, medical field, etc.). However, research on accent, intelligibility, and comprehensibility has shown that the presence of a foreign accent does not necessarily lead to miscommunication or difficulty understanding speech (e.g., Munro & Derwing, 1995; Derwing & Munro, 1997; Trofimovich & Isaacs, 2012). In other words,

L2 learners may still retain a noticeable accent while remaining intelligible and comprehensible to listeners. Nevertheless, pronunciation instruction can help learners improve clarity and communicative confidence in professional interactions with students, patients, business associates, and so on. However, before one can develop an effective pronunciation training program that can assist adult learners, it is important to understand the cognitive processes that may prevent them from successfully acquiring the target language. One such factor is the role of memory and imitation in speech learning, which will be discussed in detail in Chapter 2. A more detailed discussion of pronunciation training programs and their effectiveness will be presented in Chapter 3.

1.3 L2 Phonology Acquisition in Late Learners

This section provides an overview of the issues in acquiring L2 phonology, focusing on adult learners and their unique challenges and strategies. It will explore how cognitive processes, such as attention and perception, play an important role in acquisition. Moreover, it will discuss various models that explain the phonological acquisition process, focusing on adult learners' difficulties in perceiving and producing L2 sounds accurately.

1.3.1 Cognitive Processes in Acquisition of L2 Phonology: The Role of Perception

To address the challenges of learning L2 phonology in late learners, many researchers conducted studies to test whether certain phonetic features or

individual segments pose difficulties, specifically whether late L2 learners can process and accurately produce 'new' phones. Only those studies related to the present study's topic will be discussed. Although the Critical Period Hypothesis (CPH) is an important concept in this context, it will be discussed in detail in section 1.4.

A widely held view is that sounds are filtered when perceived; only those pertinent features like the speaker's L1 are kept, and the rest are blocked. In other words, when receiving an L2 speech signal, the L2 learner will generally choose the one closer or easiest to produce from their repertoire without focusing on specific and distinct features. Thus, when producing a sound, the target will not be the L2 but rather the L1 sound that learners perceive as the most similar to the L2 phone. This, of course, questions the issue of perceptual acuity. When there is *Focused Attention,* perceptual acuity helps the listener discriminate the sounds from the incoming speech signal so that they can be easily understood and reproduced later. For this reason, 'perception' has brought many researchers to concur that if individuals are unable to perceive or discriminate an L2 sound, they will not be able to produce it accurately (Yrttiaho, N.S., 1939; Barry, W., 1989; Flege, J. E., 1991, 1995). In addition, some studies have shown that the closer (similar) the L2 sound is to the one in the L1, the more difficult it will be to acquire due to the non-existence of a phonetic category for it. These similar L1 & L2 sounds will be merged into one category when producing the L2 target, as suggested by a cognitive mechanism called Equivalence Classification that is

part of Flege's Speech Learning Model (SLM) (1987, 1992 & 1995). Other researchers, on the other hand, have stated that it is not the case of how difficult it is to acquire L2 phones but the rate of acquisition. Similar phones are acquired faster than dissimilar ones due to a 'markedness' constraint as specified in the Similarity Differential Rate Hypothesis (SDRH) by Major and Kim (1996). According to the authors, this constraint determines what accelerates or decelerates the rate of acquisition. If a phone is marked, meaning it occurs less frequently in the L1, it will take longer for the L2 learner to acquire it. On the other hand, if it is unmarked, the rate of acquisition will be faster. Another model used to explain the difficulty of acquiring L2 phonology that is like Flege's SLM is Best's (1994, 1995) Perceptual Assimilation Model (PAM). This model postulates that the perceptual similarity between specific L1 and L2 sounds will influence how an L2 learner discriminates the differences between pairs of phonemically distinct L2 sounds in perception discrimination tasks. In Best and Tyler (2007), it is further established that perceptual similarity may affect the learning of L2 pronunciation. Some researchers have further explored this interaction and the similarity of the L1 and L2 sound systems and concluded that some models fail to demonstrate when and how this categorization process of similar and dissimilar phones will occur (Brown, 2000). This is precisely the idea of *focused attention*, which will be elaborated further in this

study. Despite the approach taken, the overall goal is to find a system that forges a cogent link between perception and production.[1]

1.4 Age Effect on the Acquisition of L2 Phonology

Given the complexities involved in acquiring L2 phonology, it is essential to consider the impact of age on this process. The Critical Period Hypothesis (CPH) provides a framework for understanding how age influences language acquisition, particularly in terms of morphology, phonology, and syntax. The following section delves into the CPH and its implications for second language learning, focusing on the differences between children and adult learners.

1.4.1 The Critical Period Hypothesis

It is a conventional belief that children and adults differ significantly in the natural acquisition of the general aspects of a language, specifically morphology, phonology, and syntax; thus, the 'age' effect becomes the primary variable used in many studies. For this reason, in the fields of Second Language Acquisition and Linguistics, to explain the major differences between children and adult learners of a first and second language, it is common practice to allude to the Critical Period Hypothesis (CPH). Although there are many variations or adaptations of the CPH, the idea was initially

[1] This idea or connection is comparable to the one that the internationally known otolaryngologist Dr. Alfred Tomatis (1920-2001) discovered back in the mid-twentieth century, the connection between the ear and the voice. He emphasized that we can only produce what the ear can hear. He originally discovered it in 1947, but it was not until ten years later that he made it public.

introduced in the works of Penfield & Roberts (1959), where it was suggested that learning a first language naturally has a cut-off point, roughly at the age of nine. Lenneberg (1967) further elaborated the CPH by studying children who suffered damage to the left hemisphere of the brain and concurred that there is a window of opportunity that an individual must automatically and naturally acquire a language, beginning early in a child's life with an age onset of two, but ending at puberty. He contended that due to the hemispheric lateralization of the brain, that is, a cognitive process that establishes the dominance of language in the left hemisphere as more language is acquired, the natural language learning process ends. It is worth noting that prior to this critical period, either hemisphere could potentially serve as the language processor. As the left becomes more dominant, the right loses its role in language processing. If a language is learned after this 'critical period,' attainment of an individual's first language or second language will be minimal. Lennerberg (1967) claimed that when dealing with brain damage, only those who suffered it before puberty were likely to recover their language skills completely. In the literature, it has been argued that this optimal time for learning a language in a naturalistic setting coincides chronologically with an abrupt decline in the brain's plasticity (Penfield & Roberts, 1959) and other neurological changes that hinder language learning for an adult (Klein, 1986; Scovel, 1988; Long, 1980). Accordingly, due to the loss of plasticity of the brain at puberty, mastery of a first or second language is difficult or requires other processes. This has been supported by studies in L2 learning, where it has been documented that

children will usually outperform adult learners (Scovel 2000), to be precise, the learning process for children may seem effortless, faster and more on target the earlier they are introduced to a second language. Hence, the widespread view "the younger=the better" for learning an L2.

1.5 CPH and Universal Grammar (UG)

The biological approach to explaining language learning resonates with what Chomsky (1965) had hypothesized in Universal Grammar (UG), the idea that genes play a significant role in human development, especially in language, at least in its initial state because it contains innate principles that induce natural language development in any language system. In other words, humans are born with a set of rules called 'Universal Grammar' hard-wired into the brain that can be accessible at any point in a lifetime as long as language exposure is from an adequate environment, which feral children from the previous section lacked. With UG in mind, Chomsky (1965) posited the idea of a Language Acquisition Device (LAD), an innate and language-specific component. It works under a set of 'principles' and 'parameters', the latter set by each specific language that exists and the former which are common to all languages. Lenneberg (1967) accepted the innateness of language in his CPH but added the specific time frame for language development that opposes the tenets set forth by UG. Much later, Steven Pinker (1994) posited his view on the age limit in language acquisition, including its underlying causes, as stated below:

> “In sum, acquisition of a normal language is guaranteed for children up to the age of six, is steadily compromised from then until shortly after puberty, and is rare thereafter. Maturational changes in the brain, such as the decline in metabolic rate and the number of neurons during early school-age years, and the bottoming out of the number of synapses and metabolic rate around puberty, are plausible causes.” (p. 293)

The Critical Period Hypothesis set forth seminal research that to this day questions the validity of the underlying mechanisms or capacities that are specific to human adults when learning a language under natural or non-natural exposure (i.e., school). Since adopting the strict version of the CPH limits the explication of how individuals can learn a language after puberty, or whichever the age limit is given, another name has been adopted, with a more relaxed definition, namely the ‘sensitive period.’ It states that the decrease in the brain's malleability is gradual without stating precisely when it begins or ends. Thus, this implies that a language can still be learned at a later age, but probably with difficulty or by other mechanisms. Since the time of the maturational theories that purportedly predicted language learning, a vast amount of research in second language acquisition and instruction has emerged that, in addition to ‘age effects,’ has considered other variables such as the length of exposure and length of residency, among others. These variables will be discussed in later chapters. For now, it is important to look briefly into another body of research and case studies subsumed within the field of psychology that may provide cogent evidence for the possibility of a

late learner acquiring an L2 because ‘plasticity’ is in fact available throughout one’s lifetime.

1.6 Counterevidence to the CPH and Loss of Neuroplasticity

Psychologists believe that genes and the environment continuously shape the brain from the moment we take our first breath of air to the day our lungs cease to do so. Hence, the idea that a ‘critical period’ exists, one that blocks restructuring of the brain and learning after puberty, is not a widely accepted doctrine within the field of psychology. As John J. Ratey, a clinical professor of psychiatry at Harvard Medical School, states in his book A User’s Guide to the Brain: Perception, Attention, and the Four Theaters of the Brain (2002), “…we, the owners---to the extent that our genes allow it---can actively shape the way our brains develop throughout the course of our lives” (p. 17). He further asserts that through our life experiences, emotions, thoughts, and actions, the restructuring of our brain is possible.

This was evident in one well known case to which he refers, the case of Temple Grandin, an autistic scholar, who through the act of rehearsal --- repetitive acts that helped her focus on her daily schedule while working at her aunt’s cattle ranch at the age of sixteen, succeeded at controlling her rage and inability to respond positively to the surrounding environment, which in turn had previously impeded her from interacting with others. The simple sound of a telephone ringing or cars passing by would cause her great confusion, which is typical in some types of autism. Thus, what this seems to

indicate is that the ability to learn and restructure disruptive patterns in the brain is possible through other mechanisms. In Temple's case, this involved training herself to focus on the task at hand.

Another case study, also included in Ratey's 2002 book, that supports the presence of plasticity beyond the 'critical period' is the case of Rickie Flach, a young woman who, at twenty-three, was unable to perceive the world as a person with normal vision would. Her visual input was similar to when one looks through a peephole of a door --- one can normally see what is in front, zoomed-in, but anything around is out of focus. In her case, the image in front of her would also become distorted after a few minutes. She struggled with this vision impairment for twenty years, from the age of three. When she was finally diagnosed properly, she was given a special pair of glasses that helped her focus for longer periods of time. After six months, she was able to see without them. The visual part of the brain was able to reorganize through training as was explained by Dr. Ratey (2002).

Similar to how Lenneberg (1967) used children with brain impairments to explain or support the CPH, the aforementioned case studies can be used to do just the opposite ---They can in turn support the idea that language learning or relearning, with a conscious effort, can still take place in early adulthood and continue thereafter.

Other researchers in the same cognitive field as the one presented above have similarly challenged the dogma that the adult brain is hard-wired, and that due to this it is incapable of restructuring itself when receiving new

stimuli. Jeffrey M. Schwartz, M.D., a research professor in psychiatry at UCLA, and Sharon Begley, an award-winning science columnist, in their book The Mind & the Brain: Neuroplasticity and the Power of Mental Force, 2003, provide counterevidence to the critical period in question. Within this book, it is shown how neuroplasticity is in fact present in adulthood. By researching numerous clinical cases of OCD patients, the authors found that these patients, just as Temple, as a result of focusing their attention to other stimuli at the time the negative ones materialized in their minds, were able to control the negative behaviors that haunted them all their lives; thus, this once again signaled that an adult's brain is not fixed. Change is possible at any age. With a conscious effort to focus on a task at hand, the brain can reorganize itself; thus, learning takes place. Schwartz (2003) puts it best when affirming the existence of plasticity in adulthood:

"The fact that (even) adults are able to learn and that learning reflects changes in synapses tells us that the brain retains some of its early dynamism and malleability throughout life. […] Thus, the power of willful activity to shape the brain remains the working principle not only of early brain development, but also of brain function as an ongoing, living process." (p. 130)

The above clinical studies suggest that, through focused attention and repeated practice, the brain can rewire itself when receiving new stimuli. This idea serves as the basis for the theoretical framework adopted in the present study---Speech Motor Control Theory (henceforth, SMCT), which will be discussed in detail in Chapter 3.

1.7 Relevance of the Pilot Study to Current Work

The preliminary findings from the pilot study, which initially included measuring data at two-time points, have led to modifications and expansions in the present study. Based on the insights gained, data collection methods were adjusted, two additional time points to measure data were added, and an advanced statistical analysis was employed to identify the actual individual and group growth trajectories related to the training of *rhotics* at different time points. In the 2014 study, data collection and analysis involved two assessments: acoustic measures and ratings for accents by native speakers of Spanish. Acoustic measures were used to assess the production of targeted phonemes, while ratings by native Spanish speakers assessed the learners' accents. The assessments were conducted before and after the completion of specific imitation tasks.[2]

1.7.1 Methodology of the Pilot Study

A preliminary pilot study was conducted to examine how native Spanish and English speakers perceived the Spanish trill [r] and tap [ɾ] produced by Spanish L2 learners. The productions were evaluated using a rating scale based on listeners'

[2] This topic was first presented at an Academic Symposium in March 2012, the University of California at Los Angeles in April 2012, the 94th Annual Conference AATSP in San Juan, Puerto Rico in July 2012, and the AATSP SoCal Fall Conference at UCLA in Fall 2013.

perceptions of the sounds produced. The findings from this pilot study helped motivate the present dissertation by highlighting the variability involved in L2 rhotic production and the need to further examine pronunciation development over time. Additional methodological details regarding the pilot study are discussed in Chapter 5.

1.8 Organization of the Dissertation

In the next chapter, Chapter 2, the relationship between the cognitive process 'imitation,' which is subsumed in the theoretical framework that will be adopted, namely Speech Motor Control Theory (SMCT), and 'memory,' will be discussed in detail. It will begin with an overview of the SMCT followed by its connection to imitation, what the latter entails in general, how it is manifested in speech, perception and production factors affecting imitation, the differences between direct and non-direct imitation in speech, and how the former is more beneficial to the adult L2 learner. The discussion will follow with an illustration of how honing this skill, imitation, can improve how sounds are retained in and retrieved from the memory and how this, in turn, can facilitate the ultimate learning of an L2 sound.

In Chapter 3, research on experiments focusing on implementing speech training programs that include imitation in adult instruction will be discussed. This chapter reviews previous work on pronunciation training and speech learning programs that have attempted to improve L2 phonological acquisition.

Chapter 4 provides a detailed discussion of the Spanish rhotics. It begins with a general overview of rhotics and their articulatory and acoustic properties. Additionally, this chapter includes a review of studies on the production and distribution of Spanish rhotics in native, heritage, and non-native speakers. It also discusses how these sounds vary depending on phonological context, speaker background, and task type.

Chapter 5 presents the methodology of the current study, including the participants, data collection procedures, coding methods, and statistical analyses employed in the investigation. As part of the analytical framework, a growth-model approach within an HLM framework is used to examine students' development over time and individual growth trajectories in the production of the Spanish trill.

Chapter 6 presents the results of the study. It includes the descriptive and statistical analyses, as well as the growth model within an HLM framework used to examine students' development over time and the effectiveness of the training program.

Finally, Chapter 7 will provide a discussion of the study---review of the research questions, a summary of the results, any implications for integrating the audio-visual speech training task in an L2 curriculum, concluding remarks, and future directions for the present study.

Chapter 2

Cognitive Processes in L2 Phonology Learning: Imitation and Memory within the SMCT Framework

2.1 Introduction

In this research, the Speech Motor Control Theory (SMCT) will be adopted to explore the relationship between cognitive and motor tasks. This section will illustrate how these cognitive processes can influence motor actions and speech production by emphasizing the role of focused attention and observation.

A substantial amount of evidence within behavioral and neurophysiological studies has established that there is a strong correlation between cognitive and motor tasks, that is, anything that requires mental effort (i.e., *focused attention*, as I will refer to it, on a particular task) will have some influence over an individual's action (Hauert, 1986; Wulf et al., 1999; Blanding & Proteau, 2000; Heyes & Foster, 2002) even if the learner observes the action without actually performing it. Neuroimaging recordings in Nishitani and Hari (2000) and Järveläinen et al. (2001), as well as transcranial magnetic stimulation (TMS) in Fadiga et al. (1995), reveal that motor-related cortical areas are activated during action observation and action execution. Of these researchers,

Fadiga's (1995) experiment has often been cited as evidence for the existence of a 'mirror' system, demonstrating that the simple observation of

an action can trigger activation of a corresponding motor act. By applying TMS to the left frontal cortex, motor-evoked potentials (MEPs) were produced in distal muscles and were enhanced during an observation task. MEPs are neural responses recorded from muscles following activation of the motor cortex and are commonly used as an index of motor system excitability. Of the four randomly assigned tasks, only those subjects who were assigned to the grasping observation task showed increased motor excitability. If observation alone activates motor representations, then imitation remains available beyond early childhood. This finding, the interaction between cognitive and motor tasks, shows that explicitly instructing students to imitate as closely as possible the L2 tokens produced by the model in the audio-visual task requires their *focused attention* (cognitive act) on the visual stimuli; therefore, helping them produce (motor act) the tokens more L2-like. Moreover, it suggests that young and adult L2 learners may share more learning abilities than traditionally assumed under strong interpretations of the Critical or Sensitive Period Hypothesis. The ability to 'imitate' should be considered an innate property since it has been established in L1 research that even newborns with typical development (Meltzoff, 1999; Kuhl & Meltzoff, 1996) and those with Down syndrome (Heimann et al. 1998) can imitate the lips and tongue movement of others in their vicinity. The ability of L2 adult learners to perform the task is consistent with the idea that some degree of neuroplasticity remains available in the adult brain.

It is worth noting that the interaction above-mentioned is located within Broca's area located on the frontal lobe of the brain's left hemisphere, which is commonly known as the area where the production of speech sounds is processed. It has been referred to as the "orchestrator of the human mirror neuron system" (Nishitani & Hari, 2000, p. 913); that is, the same cortical region is activated when an individual observes an action as well as performs it (Wohlschlager & Bekkering, 2002).

It would not be surprising for second language researchers to counter-argue that mere observation and performance of a speech act implies that language has been fully acquired, but this is not the case. It requires that the audio-visual stimuli be manipulated and represented appropriately in another area of the brain, the Short-Term Memory (S.T.M), so that it can be transferred to a more permanent storage, the Long-Term Memory (L.T.M), for successful retrieval later. For this to happen, a rehearsal process, a training phase, will have to occur. The following section will delineate the memory's role in L2 phonetic training and ultimate language acquisition.

2.2 Memory and L2 Language Learning

Learning a second language entails accumulating an unfamiliar linguistic system that may or may not be like that of the second language (henceforth, L2) learner. It is assumed that learners carry specific linguistic knowledge of their first language within the language storage area that could interfere with the proper or complete attainment of a second language. This storage area is

what we know as ‘memory,’ a cognitive process that is always active, interacting with many neurons in various brain regions, especially within the hippocampus, a part of the temporal lobe that plays a vital role in adult learning. It is well known that long-term storage and successful retrieval of the new knowledge acquired, whether retrieved immediately or much later, requires deciphering and manipulating it. The memories that we have stored permit us to make sense of the world around us. In other words, any new information received is analyzed or interpreted based on our prior knowledge stored within this processing system. Previous experiences serve as a basis for comparing the incoming information- it is categorized in a way that makes sense to us (i.e., a group of similar items). Once fully understood, it is stored for future retrieval and reference. Thus, the process occurs repeatedly, as seen in the schematic design in Figure 2.1, which is an adaptation of the Atkinson and Shiffrin (1968) information processing framework.

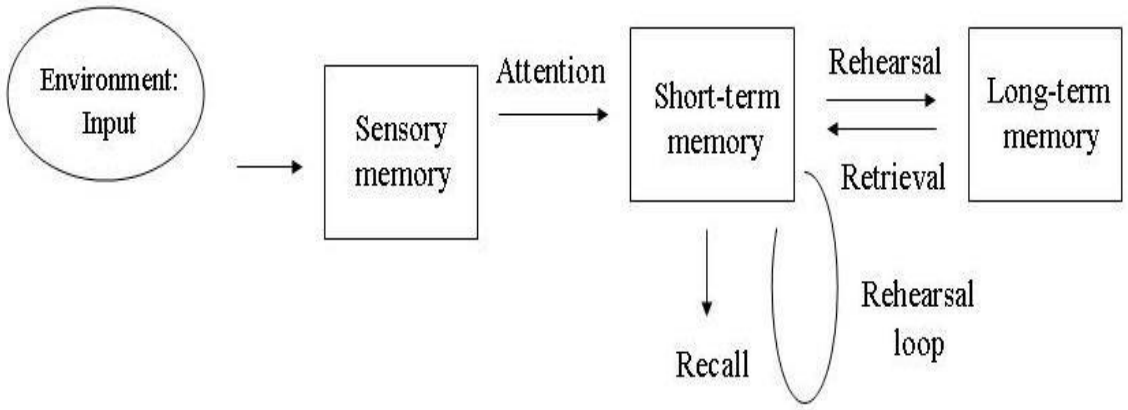


Figure 2.1. *Adaptation of Atkinson and Shiffrin’s (1968) Information Processing*

Framework

The diagram depicts how past experiences in long-term memory (L.T.M.) generally aid learning. The sensory memory receives new experiences, events, and the like, holds briefly the raw unanalyzed information---the auditory information in the echoic memory and the visual in the iconic memory, and only the information that receives focused attention will be passed on to the next area, namely the short-term memory (S.T.M) for further processing. It is sent to the permanent storage---L.T.M when fully understood and manipulated correctly. It is of utmost importance to note that to retain the new information longer and make it available for recall at any time, there must be a period of active rehearsal. The rehearsal loop serves as the venue for this.

When learning something new, it is only natural to compare it with past experiences and then analyze it. In this study, the L2 learners follow this process with the articulatory features of the incoming speech signal. It is assumed that when hearing a new sound system, they will search within their L.T.M. for the similarities and/or differences of other sounds (comparing), analyze them, and decide which one(s) they will or will not use when reproducing the new system. However, is it accurate to say that these pre-existing sounds will remain unchanged for the rest of an individual's life? Is it possible to modify, delete, or even create new ones? It is likely due to the brain's neuroplasticity and ability to adapt over time. This may help explain

how learning remains possible throughout adulthood. Each day presents people with new life experiences that shape their thoughts and plan future events. They can learn new skills, terms, and sounds, among other things. In this dissertation, one of the goals is to discuss in detail another underlying mechanism involved in processing the incoming speech signal, namely 'imitation.' Doing so aims to shed light on the potential obstacles that L2 adult learners face when learning a foreign language.

2.2.1 The Neural Basis of Memory and L2 Phonology

Within the cognitive fields, it has been postulated that the most affected cognitive processes in the adult or aging brain are 'attention' and 'memory' (Glisky, 2007). Proponents of the CPH will most likely assert that it is due to the loss of 'plasticity' within the areas that represent them, the temporal lobe and the prefrontal cortex. However, recent research in the neuroscience field has revealed that neurogenesis, that is, the ability of the brain to create new cells, has been found within the hippocampus, a part of the temporal lobe in the adult brain (Mirochnic, Wolf, & Staufenbiel, 2009; Klempin & Kempermann, 2007). This finding is relevant to L2 research because it suggests a possible mechanism through which adults may continue to learn a second language, particularly the link between *focused attention* and ultimate learning.

Knowing how the act of 'attending' an action and 'memory' interact will shed light on how it is possible for an adult learner to attain another language after the 'window of opportunity' has closed. Broadbent's Selective Listening

to Speech chapter (1958) has extensively been referred to in many studies to illustrate how specific physical properties of stimuli, speech stimuli in L2 research, are selected and transferred to the Working Memory (W.M.). Attending carefully to a particular visual stimulus, while consciously blocking other forces, helps reduce the amount of information to be manipulated in the Working Memory; thus, allowing easier transfer to Long Term Memory for later retrieval. In other words, the more concise the information coming into the L.T.M., the longer it will be retained and by extension successfully retrieved whenever needed. When dealing with the audio component, the stimuli are temporarily held in the phonological loop (part of the S.T.M.) with its temporary memory traces until they are called upon through a covert rehearsal process. The more times the memory traces are retrieved for rehearsal, the more likely they will solidify and become permanent. Thus, this implies that an L2 learner may have another vehicle to acquire an L2 sound successfully. Although it was not originally included as part of the phonological loop rehearsal process, overt rehearsal was demonstrated to improve L2 learning, as was seen in language teaching studies (Iacoboni, 2008; Ellis & Beaton, 1993).

As seen in the previous sections, the ability to learn a language after a set period is possible through the implementation of other cognitive mechanisms---imitation and memory enhancement. Although learning a second language earlier may be beneficial, it does not appear to be impossible

later in life. The key mechanism available throughout one's lifetime, imitation, must be honed as explained in the following section.

2.3 The Powerful Cognitive Process: Imitation

Imitation is a cognitive process involved in many forms of learning. Individuals use it to acquire cultural norms, emotional expressions, motor skills, and first and second languages. This dissertation focuses specifically on the imitation of speech sounds in second language (L2) learning among adult learners.

Understanding how imitation works when producing a speech sound is essential to identifying the challenges an adult L2 learner encounters. This section will delve into the mechanisms involved in imitating speech sounds and the obstacles adult learners face in this process.

Successful imitation implies that learners pay close attention, *focused attention* as discussed in 2.2, to the sounds being heard and seen (i.e., using articulatory gestures) to produce the new sounds as closely as possible to the speaker. Due to the adult learners' cognitive load[3], unless they are guided or directed, it can be difficult to focus their attention on specific articulatory or acoustic features; hence, the inability to learn a second language effectively.

[3] In this context, the phrase 'cognitive load' is not used as it is used in the Cognitive Load Theory by Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory.* New York: Springer. In this study, it is used to explain the cluttered mind that adults typically possess and is typically absent in younger human beings.

This can be applied to second language learners and classroom settings in general.

Much evidence points to imitation as a fundamental element or mechanism in the language acquisition process. Beginning at a very young age, approximately 12 weeks after birth, infants begin to imitate speech patterns from the input they receive within their environment (Kuhl & Meltzoff, 1996; Kuhl et al., 1997; Kuhl, 2004). Kuhl and Meltzoff (1996) observed that infants imitated verbally the vowels /a/, /i/, and /u/ produced by an adult speaker. They used acoustic analysis and phonetic transcription by trained listeners to provide evidence that the infants matched the modeled vowels. Unlike studies on neonatal facial imitation, which have caused considerable debate (e.g., Anisfeld, 1996; Jones, 2009), this work emphasizes the role speech imitation plays in early speech development. Although some researchers suggest that imitation diminishes as children become more fluent in their first language, it has been well documented that the ability to imitate continues throughout an individual's lifetime, particularly in conversational settings. During these conversational interactions, speakers can imitate their interlocutors at different linguistic levels, such as rate of speech (Giles et al., 1991), speech style (Kappes et al., 2009), intonation (Goldinger, 1998), regional accent (Delvaux & Soquet, 2007), and at the lexical and grammatical levels (Branigan et al., 2000; Pickering & Garrod, 2004), among others. It is a recurring behavior we all possess, of which we are not quite cognizant.

In adulthood, imitation in non-instructed speech is often considered an unintentional act known as phonetic convergence. It is an “accommodation phenomenon” that can occur when speakers are interested in facilitating the communication exchange with other interlocutors, thus creating a common linguistic ground between speakers (Giles et al., 1991; Sato et al., 2013). This accommodation does not occur in every conversation with every speaker. There are several factors that can lead to full convergence, some or none (i.e. sociolinguistic factors). In Parodi (2009), it was found that speakers within the Hispanic community in Southern California, particularly the greater Los Angeles area (the counties of Los Angeles, San Bernardino, Orange, and Ventura), which comprise numerous Hispanic dialects, were inclined to converge with the koiné found in this area. The koiné, which she refers to as the Los Angeles Vernacular Spanish or LAVS, results from the leveling of various Mexican rural dialects transported to the city (Parodi, 1999b). According to Parodi, this vernacular is well established due to the continuous immigration of Mexican speakers and other Hispanics within the area, the high degree of bilingualism, and the number of Hispanic communities that speak it. It is worth noting that, although it is spoken by most of the communities within the greater Los Angeles area, and does have native speakers, it exhibits features that are often stigmatized by monolingual speakers of Spanish, who may regard it as an “incorrect” form of Spanish (Parodi, 2010). This perception likely stems from the fact that LAVS is a **koineized variety**, meaning it emerged from the contact and mixture of various Spanish dialects

brought by immigrants to the region. Over time, this contact led to the formation of a stable variety that can be learned natively. Koineization can result in the development of distinct phonological, lexical, and syntactic features, some of which may not align with prestige norms. Table 2.1 depicts just a few of the linguistic features that Parodi found to be characteristic of the speakers of LAVS.

Table 2.1

Linguistic Characteristics of LAVS Speakers

	'Standard' Spanish	*LAVS*
Generalization of the 2nd person singular form, preterit tense	*-iste* *<comiste>* *'You ate'*	*-istes* *<comistes>* *'You ate'*
Diphthongization	<peor> [pe-or] 'worse' <toalla>	<peor> [pjór] 'worse' <toalla>

	[to-á-ya]	[twá-ya]
	‘towel’	‘towel’
Neutralization of the Spanish tap and trill. Both are pronounced as two quick taps. She labels them as ‘ŕ’	[bá-rio]	[bá-ŕio]
	<barrio>	<barrio>
	‘neighborhood’	‘neighborhood’

According to Parodi’s findings, only certain levels of speech convergence are salient among the adult speakers that have moved to the areas where this variety exists, namely the lexical and the grammatical. As for the phonological level, these adult speakers maintain their original phonological systems (Parodi, 2004). This may be due to maturational factors, as she and others have claimed (Lenneberg, 1967), explained in the first couple of sections but merits reiterating here, whereby an individual’s ability to learn a language comes to a halt once a hypothesized “window of

opportunity" to do so comes to an end, generally on or before the onset of puberty. Others, on the other hand, claim that the inability to imitate can be attributed to other factors, such as the degree of social biases and preference towards a speaker or group of the speech to be imitated (Babel, 2010). In Babel's study, results showed a definite relationship between the willingness to imitate and the social bias the imitator possessed towards the other speaker. The study consisted of New Zealanders (the potential imitators) and an Australian speaker who served as the model to be imitated. Consistent with what Babel hypothesized, the less biased, the more likely the participant would imitate the behavior of that speaker. Conversely, the higher the bias, the less likely the New Zealander would imitate the Australian model. Figure 2.2 demonstrates the relationship between an individual's bias and the likelihood of imitation (or convergence). A lower level of bias leads to a higher likelihood of imitation, while a higher level of bias results in a lower likelihood of imitation.

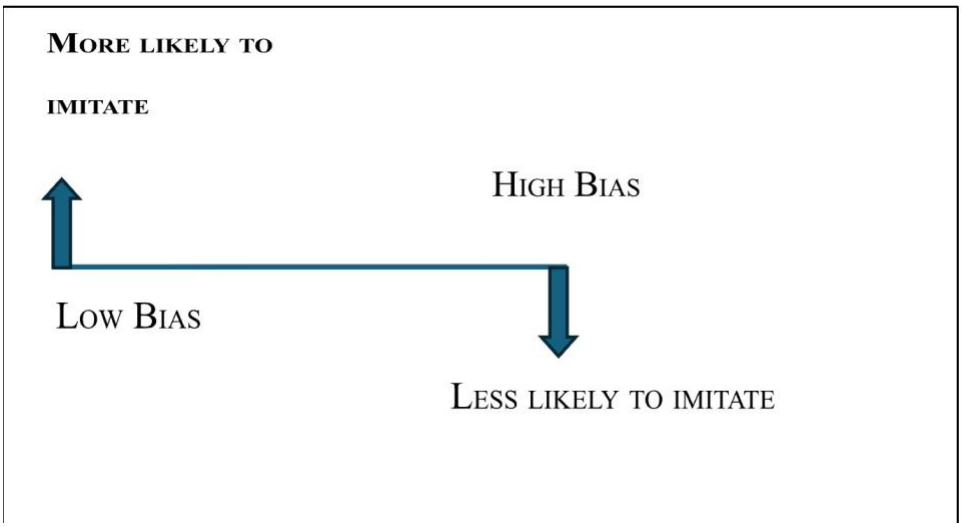


Figure 2.2. *Relationship Between Bias and Likelihood of Imitation*

Another potential factor for how much a speech model will be imitated can be related to gender and likeness. In Babel (2012), the findings showed that the degree of convergence depended on how attracted the female listener (potential imitator) was to the male speaker. The more attractive the female participants found the male model to be, the more likely they would imitate his vowels. In other studies, it was observed that typically if listeners, who are in the process of producing what they have just heard, are not linguistically close to or do not identify themselves with the speaker or particular group (Giles & Coupland, 1991) and want to maintain a social distance (Bourhis & Giles, 1977), divergence, the opposite effect, will be evident. If, on the other hand, they feel they belong to the speaker's group or want to decrease the social distance, they are more inclined to imitate the speech patterns of that person (Bourhis & Giles, 1977; Giles, 1973). It is safe

to say, thus far, that the degree of convergence or divergence will vary depending on the speakers involved.

2.4 Current Study and Convergence Effects

As aforementioned, speech imitation or phonetic convergence is an unintentional process manifested during a conversational setting. It is a natural process available throughout one's lifetime that needs to be honed during the later stages of life. In many studies that discuss the convergence process, the possibility of not converging due to an individual's inability (i.e., lack of imitation skill) to do so has not been considered. It seems that honing this skill applies to learners after a certain age and everyone who does not possess it. People with no cognitive, auditory, or speech deficiencies possess this skill; they are unaware of it. Not everyone can be a great imitator, regardless of age. However, with training, this skill can be perfected.

Based on this, understanding the role of memory in second language (L2) learning helped to move on to other essential cognitive processes involved in language acquisition, such as phonetic convergence. It also emphasized focused attention and the capacity to imitate speech patterns. This allowed us to see more clearly the causes of adult learners' challenges with perceiving and producing the sounds of a new language. The present study builds on these insights by examining how imitation can be used as a deliberate training mechanism to support L2 phonological development in adulthood.

Chapter 3

Perspective on Second Language Acquisition

and

Speech Training Programs

3.1 Introduction

The introductory chapter discussed the Critical Period Hypothesis (CPH), with 'age' being one of the contributing factors in acquiring L2 phonology. According to proponents of the CPH, due to maturational constraints on brain plasticity, the optimal age for learning phonology and other elements of a language is early in life, generally before puberty (Penfield & Roberts, 1959; Lenneberg, 1967; Curtiss, 1977; Pinker, 1994).

While some interpretations of the CPH distinguish between what is biologically driven (i.e., L1 acquisition) and what is learned (e.g., later forms of adult L2 learning), the present findings suggest that adult learners retain a significant ability to develop an L2 phonology over time.

The discussion that follows revisits the role of 'age' in L2 acquisition by examining related constructs such as exposure, formal instruction, and structured training. It also introduces the framework adopted in the present study. Building on Chapter 1, an innovative method for evaluating L2 phonological acquisition is presented: an audio-visual imitation approach combined with acoustic measurements of the Spanish trill [r] and longitudinal statistical modeling (HLM) to assess development over time. The findings aim

to demonstrate the effectiveness of a training method and contribute to discussions on the role of age in L2 phonology acquisition.

To understand how age effects have been interpreted in the literature, it is necessary to examine related constructs such as age of exposure, length of exposure, and age of arrival. In doing so, traditional views on the role of age can be revisited.

3.2 Other Possible Predictions for L2 Acquisition

Regardless of the approaches adopted in the literature, language experience is often interpreted as supporting the adage that 'earlier = better' in L2 phonology acquisition. In addition, successful language production is often attributed to learners' ability to process 'new' phonological categories accurately, despite similarity issues, as discussed in the previous section. This is echoed in the few studies that follow. Although these studies examine different variables, such as age of exposure, age of arrival, and length of exposure, they often suggest that maturational factors play a central role in ultimate attainment. Beginning with age of exposure, Au, Knightly, Jun, and Oh (2002) argued that passive exposure to a language early in childhood can be beneficial, particularly in adult heritage speakers. Their results showed that when heritage speakers relearn their first language as an adult after being dormant for some time, there are linguistic benefits (phonetic and phonological) that can lead to production in adulthood (Knightly, Jun, Oh & Au, 2003), and perceptual advantages (Tees & Werker, 1984) even if the

language has not been used since childhood. Genesee (1978) and Hatch (1983), among others, contend that when a child is introduced early to L2 instruction, the advantage that the child has over the adult learner is due to the 'length of exposure,' not age. The earlier the instruction begins, the more time the learner must practice the language; hence, the more time to become proficient in it. As it pertains to research on 'age of arrival' (AOA), its purpose is to determine whether there are any benefits in attaining L2 proficiency the younger individuals are when they arrive to the L2-speaking community, that is, before the onset of puberty (Oyama, 1978; Patkowski, 1980; Johnson & Newport, 1989; Birdsong & Molis, 2001). After puberty, the attainment of certain components of the L2 system (morphology and phonology) will be challenging, except vocabulary learning (Herschensohn, 2007). Moreover, significant results have shown a negative correlation between the age of arrival and level of L2 proficiency. The later in life the L2 learner arrives to the L2 environment, the less proficient the learner will be in the target language. On the other hand, the younger the L2 learner arrives to the L2 environment, the higher the L2 proficiency. But studies have suggested that to ensure that age effects are truly being measured and not the rate of acquisition, residency should be a minimum of five years (Snow, 1983), or if necessary ten years (Dekeyser, 2000). Thus, these findings suggest that the amount of exposure to the L2 may be a determining factor. This variable is also central to research conducted within more formal instructional settings. In what follows, similar positions are presented, but these focus on exposure within classroom

settings. Researchers within the L2 teaching field have claimed that the amount of instruction that L2 learners receive will be a determinant of the level of proficiency in the target language.

3.3 Pronunciation Instruction and Phonetic Training

The previous few sections have focused on the exposure of a foreign language within a more natural environment. The studies in this section will focus on 'exposure' as well, but in the form of instructed study, the amount of 'input' received in a classroom. However, it is important to understand that when talking about exposure in general, many tend to limit it to cases where there is natural interaction between the L2 learner and speakers of the L2, input received from media sources, and school. But, there is more to 'input' as Flege (in Piske, 2009) states:

> "[Input includes] all L2 vocal utterances the learner has heard and comprehend, including his own, regardless of whether these utterances have been produced correctly by L2 native speakers, or incorrectly by other non-native speakers of the L2" (p. 175).

The last part of Flege's definition of 'input' will thus imply that the accurate attainment of a language is inhibited if there is no structured formal exposure to it. For this reason, implementation of a good pronunciation method in an L2 classroom is needed. Moreover, more studies are focusing on using explicit phonetic training helps students hone their pronunciation skills.

3.3.1 Phonetic Training Focused Studies

Previous research has consistently shown that phonetic training can lead to measurable improvements in both perception and production (e.g., Saito & Plonsky, 2019; Sakai & Moorman, 2018; Lee, Jang, & Plonsky, 2015). Developing phonetic training tasks requires thoughtful planning by the instructor, ensuring that most, if not all, students' needs are met. Some believe this can be accomplished by implementing training tasks that focus on segmental elements of pronunciation, that is, specific features within words--- i.e. vowels and consonants (Elliot, 1995; Díaz-Campos & Morgan, 2002; Gonzalez-Bueno & Quintana-Lara, 2011).

In Elliot (1995), the objective of the study was to determine the efficacy of implementing explicit training of twenty Spanish phones, 10-15 minutes of 21 class periods. Two random classes of Spanish were selected to participate in the experiment, but only one received the treatment (training). The interaction between successful training of the phones and subject's attitude/concern for pronunciation accuracy were examined. Pre- and post-assessments served as determinants of successful change in production. The results for the control group were insignificant, whereas those who served in the experimental group showed significant improvement with the training. These findings are consistent with earlier work on high-variability phonetic training (HVPT), which has shown that exposure to multiple speakers and varied input can facilitate phonetic learning (Logan, Lively, & Pisoni, 1991).

In the study conducted by Diaz-Campos and Morgan (2002), English-speaking students in a Spanish course, were successfully trained to differentiate between two palatal realizations that normally cause them difficulty, namely the non-sibilant fricative [ʝ] and the approximant [j]. In addition to segmental training studies, a growing body of research has examined the role of imitation and repetition in L2 phonological development. These approaches often require learners to reproduce a model speaker's production, sometimes with audio-visual support, in order to enhance both perceptual and articulatory accuracy. Studies have shown that such training can facilitate improvements in pronunciation, especially when learners are encouraged to pay close attention to the phonetic detail of the input (e.g., Hardison, 2003; Saito, 2013; Foote & Trofimovich, 2014). These findings are particularly relevant to the present study, which adopts an imitation-based approach to phonetic training. Other studies have found that formal pronunciation training predicted successful production of phones that are not part of the L2 learner's phonological system, such as the Spanish trill 'rr' (Meléndez-Ballesteros, 2014). In this study, seventy-four students, who were enrolled in a second semester elementary Spanish course in a Southern California college, participated in an imitation training task. Half of the students were instructed to imitate, as closely as possible, the production of an audio-visual model who pronounced the tokens in questions. The other half were assigned other types of listening and speaking tasks (three to be exact) without explicitly being told to imitate. Preliminary results revealed that those

who completed the imitation task successfully produced the targeted phones. As these segmental studies, and many others reveal, the act of an L2 learner to focus and imitate a particular phonological feature requires what I will refer to as *Focused-Attention Training* (F.A.T.), which should be continually honed (see Appendix C for an example of the audio-visual task used in the present study).

Although phonetic training has been shown to improve L2 pronunciation, perception and production outcomes are not always uniform across learners or training conditions. In general, learners do not exhibit consistent or linear patterns of development, and patterns may vary in both direction and magnitude across individuals and time. In more recent research, the influence of the structure of training itself on learning outcomes has been examined. For example, Saito and Chen (2025) show that the distribution of training sessions plays a significant role, with spaced training leading to greater gains and retention than massed practice. These findings suggest that not only the presence of training, but also how it is implemented, can affect the extent to which learners benefit from explicit instruction. While these studies highlight the importance of training design, they primarily focus on group-level outcomes (i.e., averages across learners) rather than individual trajectories over time.

Chapter 4

The Spanish Trill in L2 and Heritage Speakers

4.1 Introduction

In chapter 3, we explored various perspectives on second language acquisition and the importance of structured formal exposure to the target language and pronunciation methods in an L2 classroom. The language training programs discussed demonstrated that by explicitly instructing students to focus on distinctive features of the speech signal, improvement in L2 phonology is possible. This is only feasible if instructors develop phonetic training tasks that consider most, if not all, students' needs.

The current chapter delves into the specific sound analyzed in this study, the Spanish trill. It will provide a general description, cross-linguistic variations, and articulatory and acoustic properties of rhotics. By examining the rhotics closely, the goal is to provide language instructors with a method that will aid in the development of more effective language teaching strategies and improve their understanding of phonological acquisition in L2 learners.

4.2 General Description of Rhotics

Rhotics are a class of sounds that have multiple realizations typically represented by the letter 'r' in languages where they occur. These sounds are described by their unique and challenging articulatory and acoustic properties that vary across languages and dialects. In general, rhotics can be produced

with the tongue's apex (tip) or blade moving towards the alveolar ridge, curling towards the palate, or going as far as the uvula. These different productions will depend on the language, dialect, and the phonetic context in which they occur.

Across languages, the sounds commonly labeled "r" vary widely in their articulation. In American English, the phoneme /r/ is typically realized as a postalveolar approximant [ɹ], occurring in word-initial and postvocalic positions (Ladefoged & Johnson, 2015). In these distributions, there is no contrast between a tap and a trill. The alveolar tap [ɾ] is observed in intervocalic contexts as in *letter* and *ladder*. This sound results from the flapping of /t/ and /d/ rather than from allophonic variation of /r/. By contrast, in Spanish, the tap [ɾ] and the trill [r] represent a phonemic contrast (e.g., *pero* 'but', *perro* 'dog') (Hualde, 2005). This cross-linguistic difference is significant for understanding the challenges faced by English-dominant learners acquiring Spanish rhotics.

4.2.1 Dialectal Variation and Target Form

Studies focusing on the production of the rhotics in non-classroom settings have traditionally focused on the successful production of these sounds with respect to a target form, most often defined in terms of acoustic properties such as duration (Quilis, 1993; McDonough & Johnson, 1997; Recasens, 1986, among others) and number of lingual contacts (Bradley & Willis, 2012; Ladefoged & Maddieson, 1996; Lavoie, 2001; Lindau, 1985; Solé 2002). The expected duration and the number of lingual contacts, if any, however, can

vary across dialects of Spanish. Features such as these are normally used to make the tap/trill distinction (the ‘r’ vs ‘rr’) but are insufficient to characterize the variants that are present in U.S. Spanish, particularly in the greater Los Angeles area. Hence, if we turn our attention to a second language learner of Spanish in a classroom setting within the greater Los Angeles area, assessing their production becomes a more complex task. This variability complicates the selection of an appropriate reference point.

For purposes of this dissertation, a reference mean was calculated by averaging the trill durations produced by the native Colombian model speaker and the heritage speakers included in the dataset. The average duration of the trill in the present dataset is approximately 110 ms, roughly 25 ms longer than the value reported for Peninsular Spanish in Quilis (1993). This difference likely reflects dialectal and contextual variation rather than deviation from a target form. This value serves as the point of reference against which learner development is evaluated in the current study.

Establishing a clearly defined reference mean and examining how learners’ productions approach that value over time requires a statistical model capable of capturing longitudinal change.

4.3 Spanish Rhotics

Due to their articulatory and aerodynamic complexity, Spanish rhotics (the r’s) present one of the greatest pronunciation challenges to first and second language learners. It has been documented that children learning Spanish as

their first language typically reach near-complete acquisition of these sounds by around age 7 (Bosch, 1983). Rhotics can be articulated as two distinct phones in Spanish, the single tap ‘r’, phonetically transcribed as [ɾ], and the trill with two to three taps to the alveolar ridge [r]. The Spanish trill occurs in word-initial or word-medial syllable-initial position after the alveolar consonants /n/, /s/, and /l/ (e.g., [r]ío ‘river’; al[r]ededor ‘around’ or ‘surrounding’; En[r]ique; Is[r]ael). The trill may also occur after a tautosyllabic consonant (e.g., consonant clusters such as b[r]oma ‘joke’), whereas the tap is commonly found in word-final position (e.g., habla[ɾ] ‘to talk’). Variation between the tap and trill may additionally occur in word-medial coda position and in word-final position before a consonant or a pause (e.g., ser poeta ‘to be a poet’; cantar## ‘to sing’), although these contexts are most commonly realized as a tap. The tap and trill acquire a contrastive feature in intervocalic position; thus, if not pronounced as expected, it can lead to confusion or miscommunication (e.g., *ca*[ɾ]o ‘expensive’ vs. ca[r]o ‘car’; pe[ɾ]ito ‘expert’ vs. pe[r]ito ‘small dog’).

4.3.1 Articulatory and Acoustic Properties

The Spanish trill is among the most difficult sounds to acquire because it requires very precise aerodynamic conditions, specifically the coordination of airflow and tongue movement (Ladefoged & Maddieson, 1996). To produce the Spanish trill, a substantial amount of air must be built up and directed over the

tongue so that it can vibrate. In addition to this, the tongue must be curled so that the apex (the tip) can make contact with the alveolar ridge.

Acoustically, the Spanish trill is characterized by a series of rapid interruptions in the speech signal caused by repeated contact between the tongue tip and the alveolar ridge. These contacts appear in the waveform and spectrogram as brief closures and releases, often accompanied by changes in intensity and periodicity. Figure 4.1 presents an example of a Spanish trill [r] and tap [ɾ] displayed in Praat, including the waveform, spectrogram, and TextGrid annotation. Since the trill involves multiple articulatory contacts (occlusions), it typically exhibits a longer duration than the tap.

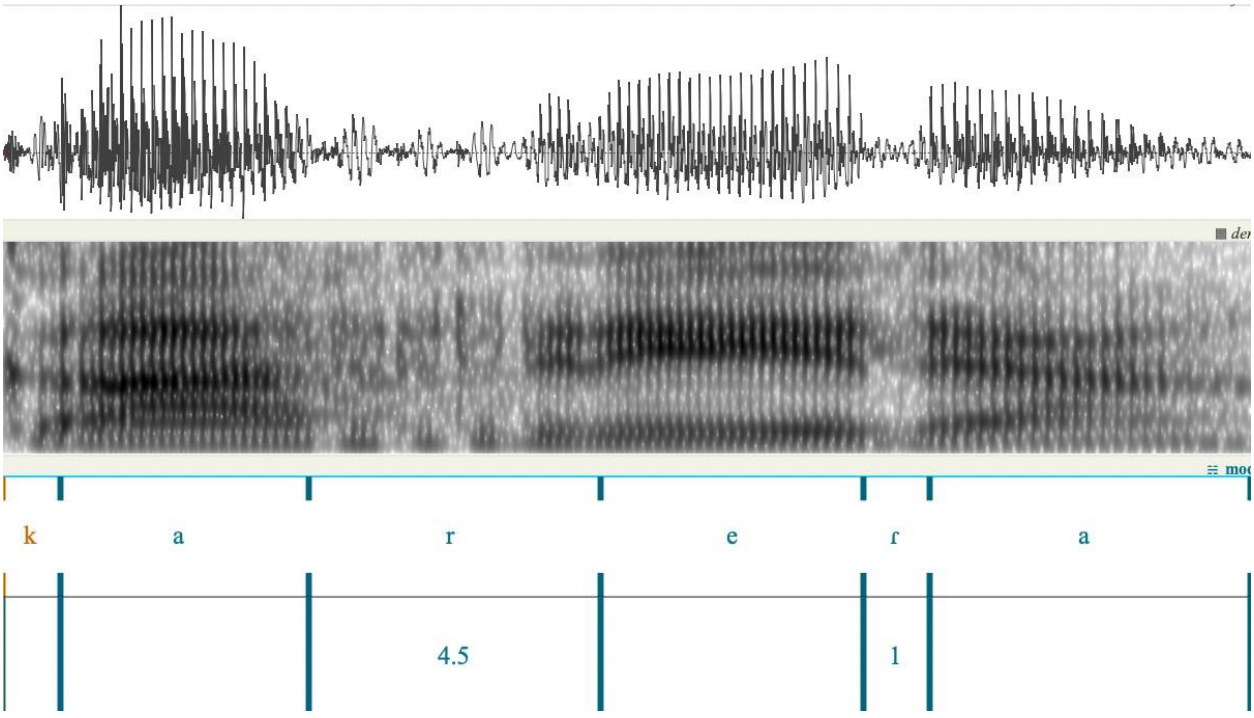


Figure 4.1. *Spectrographic representation of the Spanish trill and tap in carrera 'career'.*

***Note**:* The numbers indicate the occlusions associated with each rhotic.

In addition to its acoustic features, successful trill production also depends on precise tongue placement. Figure 4.2 below illustrates the ideal place (i.e., the alveolar ridge) for the successful production of the Spanish trill. By positioning the tip of the tongue at this precise point, the trill can be produced. Additional figures illustrating different types of rhotic realizations are included in Appendix B.

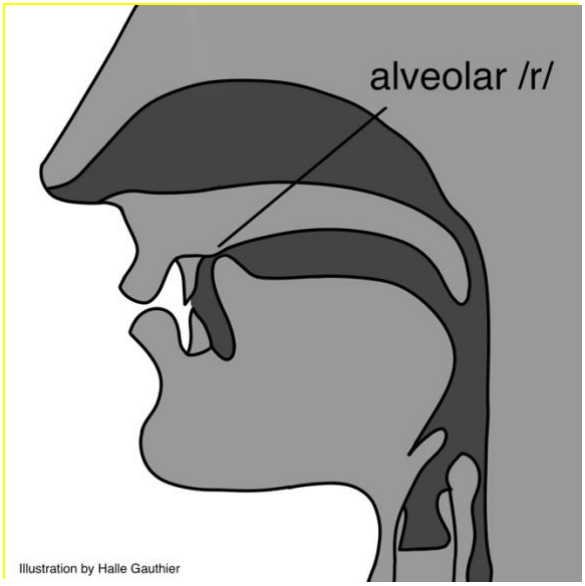


Figure 4.2. *Position of the tongue tip at the alveolar ridge during the production of the Spanish trill.*

Having outlined the articulatory characteristics of the Spanish trill, the next section reviews relevant studies on L2 and heritage Spanish rhotic production and identifies gaps in the literature.

4.4 Previous Literature on Phonetic Variation and Distribution of Rhotics

Considering how difficult it is to produce the trill and how much it can vary, even within the same speaker, prior research has reported mixed results when looking into speaker background, task type (i.e. spontaneous/natural speech vs. controlled elicitation), and phonological context, such as whether the trill occurs in word-initial or medial position. As a result, several studies have examined the Spanish trill [r] in depth and have analyzed phonetic variation as well as the distribution of variants across phonological environments and proficiency levels. Whereas some studies primarily focus on describing acoustic variability (e.g., Carballo & Mendoza, 2000; Henriksen, 2015), others examine how variants are distributed across contexts or proficiency levels (e.g., Amengual, 2016; O'Rourke & Potowski, 2016). A smaller subset of studies treats variation and distribution as part of a developmental reorganization process (e.g., Rose, 2010).

However, much of this research relies on cross-sectional comparisons that provide only snapshots of learner performance, rather than capturing how the same individuals change over time. For example, Rose (2010) examines the development of intervocalic taps and trills by grouping learners according to proficiency level instead of following individual learners longitudinally. Since each proficiency group consists of different individuals, the observed patterns reflect aggregated outcomes rather than developmental change within the same speaker. It is important to note that learners at each level bring different

language backgrounds, instructional experiences, and sociolinguistic influences that may shape the group averages. As a result, these comparisons offer a view of how variants are distributed across proficiency levels but do not provide direct evidence of how a learner's phonetic system reorganizes over time. While Rose (2010) shows that learners gradually associate different rhotic variants with their appropriate phonological contexts, studies of native Spanish demonstrate that rhotic production itself is highly variable.

Bradley and Willis (2012) investigate phonetic variation in syllable-initial Spanish rhotics in Veracruz Mexican Spanish. Their focus is on how the trill /r/ and tap /ɾ/ maintain phonological contrast despite extensive phonetic reduction, that is, when the articulatory gesture is weakened. For example, they report that trills may be produced with only one or two lingual contacts followed by an approximant phase, which they transcribe as [ɾɹ] or [rɹ]. Even in these reduced realizations, trills remain acoustically longer than taps, allowing the contrast between /r/ and /ɾ/ to be maintained. Using acoustic analysis of rhotic tokens elicited through a wordless picture-book narrative, the authors show that taps and trills exhibit a wide range of realizations, including approximants, fricatives, and other reduced forms. More importantly, they demonstrate that the contrast between the two rhotics is maintained primarily through segmental duration rather than through a specific number of lingual contacts. Even when trills display fewer contacts than traditionally expected, they remain significantly longer than taps, indicating that temporal

cues play a central role in maintaining the contrast. This study provides evidence that Spanish rhotic contrast can be maintained through gradient phonetic cues, particularly duration, rather than through a fixed number of lingual contacts. However, because the analysis is cross-sectional and based on spontaneous narrative speech, it does not address how individual speakers' rhotic productions reorganize over time or whether instructional intervention can influence temporal aspects of trill production.

Studies on heritage speakers further examine how this variability emerges during the development of Spanish trill production. Repiso-Puigdelliura G and Kim JY (2021) examine the production of the trill /r/ among child and adult heritage speakers of Spanish. Results show that although adult heritage speakers produced the trill in a more target-like manner than child heritage speakers, many still diverged from monolingual native baselines. Their analysis suggests that trill development may progress from a single lingual constriction to frication and eventually to multiple lingual constrictions. Importantly, some adult heritage speakers retained variants associated with earlier stages of development, which resulted in increased variability in their trill inventories. These findings highlight that trill development may remain unstable even into adulthood and that variability persists across speakers. However, the study relies on cross-sectional comparisons and therefore does not track how individual speakers' productions evolve over time. Additional work on heritage and bilingual

Spanish further highlights the variability of rhotic production and the influence of linguistic experience on the realization of the tap–trill contrast.

Amengual (2016) examines the Spanish tap–trill contrast among heritage speakers and L2 learners, with particular attention to the role of language dominance in shaping rhotic production. The task involved reading sentences within a carrier phrase rather than free reading. The study measures trill duration and number of occlusions across Spanish-dominant heritage speakers, English-dominant heritage speakers, and English-dominant L2 learners. The results reveal systematic differences across groups: Spanish-dominant heritage speakers produce more canonical trills with longer durations and multiple occlusions, whereas English-dominant heritage speakers and L2 learners rely more heavily on non-canonical variants while still maintaining the phonological contrast through duration. These findings highlight the heterogeneity of heritage speakers and demonstrate that rhotic acquisition is influenced by speakers' linguistic experience and language dominance rather than being determined solely by categorical accuracy. At the same time, the study remains cross-sectional and task-specific, which raises questions about how individual learners' rhotic systems evolve over time and whether training can influence the phonetic realization of trills.

While Amengual (2016) examines the acoustic realization of the tap–trill contrast in controlled production tasks, sociophonetic research has also explored how rhotic variation emerges in dialect contact situations. O'Rourke

and Potowski (2016) investigate phonetic accommodation in Spanish dialect contact situations in Chicago, focusing on the realization of word- and syllable-initial /r/ alongside coda /s/. The study analyzes sociolinguistic interview data from multiple generations of Mexican, Puerto Rican, and mixed-heritage speakers and shows that trill realization varies as a function of generation, interlocutor, and social alignment. While some groups show limited accommodation, others exhibit patterned shifts in rhotic realization depending on social context. These findings reinforce the view that rhotic production is highly variable and shaped by social and contextual factors, consistent with Rose's (2010) emphasis on distributional rather than categorical phonological patterns. However, the study does not examine phonetic development over time within speakers, nor does it model how rhotic systems reorganize as a function of learning or training, as its focus is sociophonetic accommodation rather than developmental change.

Research on bilingual production focuses on how language experience shapes the realization of Spanish rhotics. Cummings Ruiz and Montrul (2020) investigate how bilingual speakers produce Spanish tap and trill rhotics and focus on the role of age of bilingual onset and language input. They use acoustic measures such as duration and number of apical occlusions during storytelling and picture-naming tasks, and compare sequential and simultaneous heritage speakers, L2 learners, and monolingual Spanish speakers.

The results show that sequential bilingual heritage speakers pattern more closely with monolinguals in their rhotic production. Simultaneous bilinguals, however, exhibit greater variability and less target-like realizations, although they still maintain a distinction between tap and trill categories. The authors interpret this variability as reflecting incomplete acquisition rather than phonological loss or transfer from English. This interpretation aligns with Rose's (2010) view that rhotic acquisition is a gradient process characterized by dynamic reorganization rather than categorical success or failure. Since the study is cross-sectional, it does not provide insight into how these phonological systems reorganize within individual speakers over time or how they might respond to targeted instruction.

The studies reviewed here show that rhotic variability emerges not only in bilingual and L2 populations but also among native Spanish speakers, suggesting that such variation is inherent to the rhotic system itself. Lastra and Martín Butragueño (2006) analyze the phonetic and sociolinguistic variation of Spanish rhotics in Mexico City, focusing on the assibilation of tap and trill rhotics. The study draws on sociolinguistic interview data and provides evidence of multiple rhotic realizations, including canonical trills, fricatives, and assibilated forms, and shows that these variants are systematically distributed according to phonological context, speech style, and social factors such as gender, age, and social class. Most importantly, the authors demonstrate that even native speaker trill production is highly variable and cannot be reduced to

a single articulatory target. This finding strongly supports Rose's (2010) argument that phonological categories must be understood as distributions of phonetic variants rather than a single canonical realization. However, since the study is sociophonetic rather than developmental, it does not address how such distributions emerge, reorganize, or stabilize within individual speakers across time or learning conditions.

Similarly, Mazzaro and González de Anda (2020) investigate rhotic assibilation in Chihuahua and El Paso Spanish and examine its phonetic realization and sociolinguistic conditioning in a bilingual contact setting. They use acoustic and perceptual analyses of rhotics that are elicited through reading, narrative, and conversational tasks. The study identifies a wide range of rhotic variants, including assibilated, partially assibilated, approximant, and canonical tap and trill forms. The results show that assibilation is the most frequent realization in phrase-final position and is influenced by phonological context, task type, gender, and generation. Bilingualism with English does not significantly predict variant choice. These findings further demonstrate that trill production is inherently variable and shaped by contextual and social factors, reinforcing Rose's (2010) claim that rhotic categories consist of structured distributions rather than uniform articulatory targets. The study focuses on sociophonetic variation across speakers and contexts and therefore does not model developmental change or examine how individual speakers' rhotic systems reorganize over time or under specific learning conditions.

In addition to sociophonetic variation across dialects, research on first-language acquisition also provides important insight into the development of Spanish trill production. Carballo and Mendoza (2000) examine the acoustic characteristics of Spanish trill production in monolingual Spanish-speaking children aged three to nine, with particular attention to developmental differences associated with speech intelligibility. The study uses acoustic measures such as segment duration, spectral properties, and the number of lingual contacts and shows that younger children and children with articulatory difficulties often produce reduced or tap-like realizations in trill contexts, characterized by fewer lingual contacts and shortened duration. The findings highlight the articulatory complexity of the trill and its late acquisition. This pattern suggests that the role of speech motor control plays an important role in developmental outcomes. While the study provides important baseline evidence that trill production develops gradually and is subject to considerable variability in childhood, development is largely framed in terms of accuracy relative to an adult target. Unlike Rose's (2010) distributional approach, the study does not examine how variant distributions reorganize across contexts. Additionally, it does not model longitudinal change within individual children. This limits its relevance for understanding phonetic reorganization over time.

Overall, these studies demonstrate that Spanish rhotics exhibit considerable phonetic variability across speakers, dialects, and bilingual contexts. However, much of this research relies on cross-sectional designs that compare groups of speakers rather than examining how the same

individuals' productions change over time. As a result, it remains difficult to determine how trill production develops longitudinally within individual learners. Table 4.1 below summarizes the studies reviewed in this section.

Table 4.1

Summary of Previous Research on Spanish Rhotics Development and Variation

Study	Participants	Task Type	Analytic Focus	View of Variation / Development	Key Findings	Limitation
Rose (2010)	L2 learners of Spanish	Reading and elicited produc-tion	Distribution of rhotic variants across phonological contexts	Distributional / gradient	Learners gradually associate rhotic variants with appropriate contexts	Cross-sectional grouping; does not track individual development
Bradley & Willis (2012)	Native speakers (Veracruz Spanish)	Wordless picture-book narrative	Acoustic properties and phonetic reduction of rhotics	Gradient phonetic variation	Tap–trill contrast maintained primarily through duration rather than number of contacts	Cross-sectional; no developmental analysis

Repiso-Puigdelliura & Kim (2021)	Child and adult heritage speakers	Production tasks	Develop-mental stages of trill realization	Develop-mental / variable	Trill development progresses through intermediate articulatory stages	Cross-sectional; does not track change within speakers
Amengual (2016)	Heritage speakers and L2 learners	Controlled sentence reading	Acoustic measures (duration and occlusions)	Experience-dependent variation	Language dominance influences trill realization and variability	Cross-sectional and task-specific
O'Rourke & Potowski (2016)	Heritage speakers in Chicago	Socio-linguistic interviews	Sociophonetic accommoda-tion	Socially conditioned variation	Rhotic realization varies by generation, interlocutor, and social alignment	Focus on sociophonetic variation rather than phonetic development
Cummings Ruiz & Montrul (2020)	Heritage speakers, L2 learners, monolinguals	Story-telling and picture-naming	Acoustic measures of taps and trills	Gradient bilingual development	Sequential bilinguals pattern closer to monolinguals; simultaneous bilinguals show greater variability	Cross-sectional design

Lastra & Martín Butragueño (2006)	Native speakers (Mexico City)	Socio-linguistic interviews	Assibilation and rhotic variation	Socio-phonetic variation	Multiple rhotic variants distributed across contexts and social factors	Not developmental
Mazzaro & González de Anda (2020)	Speakers in Chihuahua and El Paso	Reading, narrative, conversa-tional tasks	Assibilation and sociophonetic conditioning	Context-dependent variation	Assibilation common in phrase-final position; influenced by context and social factors	Cross-sectional sociophonetic analysis
Carballo & Mendoza (2000)	Spanish-speaking children (3–9 years)	Production tasks	Acoustic properties of trill development	Develop-mental acquisition	Younger children produce reduced or tap-like trills	Development framed relative to adult targets; not distributional

In addition, while several studies provide evidence of variability in trill realization, fewer examine how specific training tasks may facilitate phonetic change at the token level. This gap is particularly relevant given theoretical accounts that view phonetic development as gradual, variable, and sensitive to both input and practice. Addressing these limitations requires not only longitudinal data but also carefully designed elicitation tasks capable of capturing fine-grained phonetic change.

To this end, a pilot study was conducted prior to the present research. The pilot study served as an exploratory step to evaluate the feasibility of imitation-based pronunciation training tasks and to help shape the methodology used in the present study. The following section briefly describes this pilot and its role in the development of the research design.

4.5 Pilot Study

The pilot study aimed to examine whether imitation-based pronunciation training could help improve the production of Spanish rhotics [r] and [ɾ], particularly the trill, the double 'r', among adult L2 learners of Spanish. It was conducted with students enrolled in Elementary Spanish courses at the same Southern California institution from which the participants in the main study were later drawn. The pilot sought to evaluate the effectiveness of several training tasks, determine which tasks best captured learners' phonetic performance, and assess whether measurable change could be observed following the training tasks.

A total of seventy-four students participated in the pilot study during the Spring 2012 semester. The participants included both non-native speakers of Spanish and heritage speakers from diverse linguistic backgrounds. Students completed a series of listening and speaking exercises through Horizon Wimba, an online voice platform used for oral assignments. Four elicitation and training tasks were designed to elicit rhotic productions. These tasks combined perception and production activities in different ways: (a) dictation

combined with shadowing activities, (b) dictation without shadowing, (c) listening and repetition tasks, and (d) audio-visual pronunciation task using video prompts through the Horizon Wimba platform. Table 4.2 summarizes the four pilot study tasks and their primary instructional focus.

Table 4.2

Summary of Pilot Study Tasks

Group	***Task Type***	***Activity***	***Focus***
A	Dictation + Shadowing	Listen to audio and simultaneously repeat	Perception + Production
B	Dictation	Listen to audio and transcribe	Perception
C	Repetition	Listen to audio and repeat recordings	Production
D	Audio-Visual Pronunciation Task	Observe the model speaker and complete the task using either imitation or repetition instructions	Production + Visual Cues

The shadowing activities were adapted from exercises commonly used in interpretation training and oral proficiency development. Similar to the present study, the targeted vocabulary and pronunciation activities were incorporated into the students' regular oral homework assignments, which also included more communicative and open-ended speaking tasks. Approximately every two weeks, as new chapters were introduced, students practiced vocabulary items and oral activities containing Spanish rhotics.

Additional oral homework activities incorporated through the Horizon Wimba platform included pronunciation-identification exercises, stress-placement activities, and oral responses to recorded prompts. Students listened to recorded

questions, pronunciation models, or stress-placement options and then responded orally or selected the appropriate pronunciation from multiple choices. For example, students listened to different stress realizations of words such as publicó and identified the correct syllabic stress placement. In this sense, the preliminary study helped identify which pronunciation-training tasks most effectively elicited rhotic productions and learner attention to pronunciation.

Following the training phase, perceptual ratings by native speakers of Spanish were collected to evaluate changes in learners' productions. Native Spanish and English speakers rated the Spanish trill [r] and tap [ɾ] produced by Spanish L2 learners using a rubric scale ranging from 0 to 5. Since the raters were not phoneticians, phonetic symbols were avoided whenever possible, and orthographic examples were used to guide the ratings. While the pilot study yielded some evidence of short-term improvement, particularly in tasks involving imitation of a native model, the results also revealed considerable variability across productions. Preliminary descriptive graphs from the pilot study suggested that learners demonstrated greater difficulty with intervocalic /r/ productions than with other rhotic environments. In particular, the intervocalic trill showed lower rates of accurate production across learners, which further motivated the present dissertation's focus on the Spanish trill. Examples of the questionnaire materials, Horizon Wimba task screenshots, sample prompts, and training activities used during the pilot phase of the study are included in Appendix C. Additional descriptive graphs

and spectrographic representations from the pilot study are included in Appendix D.

Based on these findings, two of the four tasks were identified as the most effective in eliciting clearer and more consistent rhotic productions, namely Tasks C (repetition) and D (audio-visual task using video prompts). Unlike the remaining tasks, which emphasized perception, dictation, and response-based activities, these tasks required learners to repeat productions immediately after hearing a native speaker model. Both tasks contributed to the development of the study; however, the present study ultimately focused on the audio-visual task (Task D), which was later implemented using two separate sets of instructions. All participants completed the same audio-visual activities, but the term training in the present study specifically refers to the focused-attention imitation component in which learners were explicitly instructed to pay attention to and imitate the articulatory gestures (e.g., tongue and lip movements) and pronunciation of the model speaker. This focused-attention component was incorporated into the Focused-Attention Training (F.A.T.) framework introduced earlier.

The pilot study primarily employed a pre- and post-training design. Students completed the activities over multiple sessions throughout the semester, and comparisons focused on productions collected before and after the training phase. The tasks were incorporated into the students' online oral assignments through the Horizon Wimba platform and were designed to target

both perception and production of Spanish rhotics in different speaking contexts.

While the pilot study initially focused on whether imitation-based tasks could facilitate more native-like productions of Spanish rhotics, the present dissertation study narrows its focus specifically to the Spanish trill. Additionally, it adopts a more longitudinal and variability-centered perspective informed by subsequent methodological and theoretical developments in L2 phonology research. One of the primary revisions involved expanding the number of assessment sessions to examine development over time and permit the use of HLM. In addition, the dissertation refined the elicitation and training materials based on the pilot study findings. Greater emphasis was also placed on individual learner trajectories rather than solely on pre- and post-training comparisons. Rather than viewing pronunciation development strictly in terms of native-like attainment, the present study examines how learners' productions fluctuate and develop over time at the individual level. In this sense, the pilot study served primarily as a foundation for refining the training tasks and research design later implemented in the dissertation.

The pilot study relied primarily on descriptive observations, pre- and post-training comparisons, questionnaire responses, and visual representations of learners' productions rather than the longitudinal statistical analyses later used in the present study. These exploratory analyses helped identify general production patterns and informed the development of the research design adopted in the dissertation. Nonetheless, the pilot study played a crucial role in

refining the experimental design and establishing the foundation for the longitudinal approach adopted in the present study

Chapter 5

Methodology

5.1 Introduction

This chapter presents the methodology conducted in the present study. It includes the research questions, participant information, training procedures, data collection process, acoustic and statistical analyses that were used to examine the development of Spanish trill production over time.

5.2 Research questions and Hypotheses

The focus of the study presented in this dissertation is the role of imitation in Spanish phonology acquisition within an instructional setting. The audio-visual task designed for this study explicitly instructs learners to imitate the speech of a model presented on a computer screen, as illustrated in Figure 5.1. The task serves as a training tool for adult learners of Spanish who may vary in their ability to imitate speech.

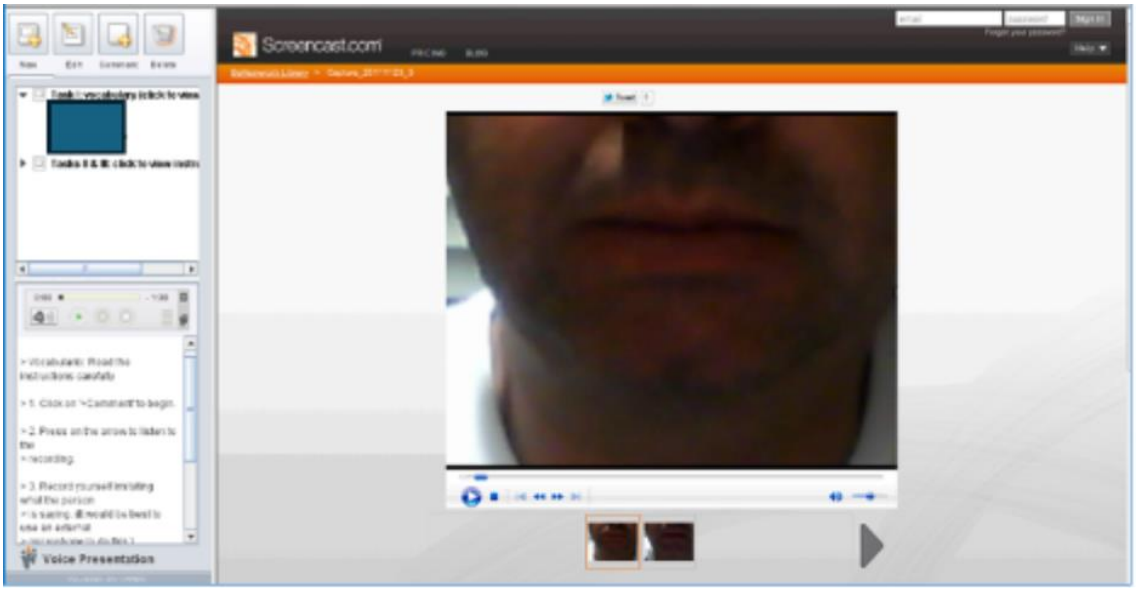


Figure 5.1. *Wimba platform used to deliver the audio-visual pronunciation task and record student productions.*

With this in mind, the study aims to address the following general questions: To what extent is the ability to imitate available to late learners? If so, can second language learners and heritage speakers draw on this cognitive ability to approximate the phonology of Spanish later in life? More specifically, the study addresses the following questions:

RQ1 (effect of time): How does trill duration change over time across the four assessment points?

RQ2 (effect of training): Can audio-visual tasks with explicit focused-attention imitation instructions support learners' ability to produce Spanish phones that may not be fully available in their phonological repertoire, such as the Spanish trill /r/?

RQ3 (effect of heritage status): How do heritage speakers and second language learners compare in their production of the Spanish trill over time as a function of training condition?

5.2.1 Hypotheses

As mentioned earlier, tracking individual growth at four different time points is central to this study. For this reason, the following hypotheses are tested.

1. Is there any change in trill duration over time for each student?

H_0: There will be no significant change in trill duration over time.

H_1: Students' trill durations will move toward the target duration defined for the trill in this study over time.

2. What is the rate of change for each individual learner? To what extent does focused-attention training (F.A.T.) influence learners' development over time? Additionally, which group exhibits a greater rate of change as a function of F.A.T.: the imitation group or the repetition group?

H_0: There will be no significant difference in the rate of change over time between the training and non-training groups.

H_1: The training group will exhibit a greater rate of change over time than the non-training group.

3. Do heritage speakers and L2 learners differ in terms of their rate of change over time as a function of F.A.T.?

H_0: There will be no significant difference in the rate of change over time between heritage speakers and L2 learners as a function of F.A.T.

H_1: Heritage speakers and L2 learners will exhibit different rates of change over time as a function of F.A.T.

These hypotheses are tested using the HLM approach described later in this chapter.

5.3 Participants

5.3.1 Participant Background

Eighty students, who were enrolled in an elementary Spanish course in a Southern California university during the data collection period, were given

listening and speaking tasks as part of their homework assignments. This cohort is part of the broader data collection from 2012 to 2015. Most of the participants' ages ranged from 18 to 23 years. At the time of the implementation of the phonetic training task and data collection, they had been taking Spanish for two semesters. There were non-heritage speakers of Spanish, of diverse ethnic backgrounds, as well as heritage speakers of Spanish, most of Mexican descent. The students were randomly assigned to two groups: A) non-training group or B) training group. Heritage and non-heritage speakers were represented in both groups.

For the current study, a subset of 40 students was selected to be thoroughly analyzed. Among the learners, 9 non-heritage speakers and 12 heritage speakers were assigned to the training condition, whereas 10 non-heritage speakers and 9 heritage speakers were assigned to the non-training condition. The decision to analyze just 40 students was based on several factors. Firstly, there was an uneven number of heritage versus non-heritage speakers in the original cohort, which could have presented a biased or inaccurate analysis. Secondly, the acoustic and statistical analyses required for this study were quite cumbersome. By handling a balanced subset, it ensured the findings were reliable, valid, and meaningful. Although the data were collected between 2012 and 2015, the present study in developing an effective pronunciation training program for learners of Spanish remains applicable, making the findings highly relevant. Research in L2 phonology

acquisition has always been an important area of study when proposing new methodologies, and it is constantly evolving.

5.4 Consent

During the data collection phase, students were given listening and speaking exercises designed to enhance their oral communication skills. The exercises were created to improve the students' proficiency in phonology and oral expression, which are critical components in the fields of Linguistics and Second Language Acquisition.

Informed consent was obtained from all participants prior to their inclusion in the study. The consent process included a statement in the course syllabus informing students that their work might be used for future research, published, or presented at conferences. Additionally, a detailed consent form was provided to the students, which they signed and dated (See Appendix E).

The consent process ensured students were fully informed about the manner their class work would be used in the study. It gave them options to protect their anonymity and respected their preferences. It is worth noting that the data collected during the data collection period as part of class oral assignments is now being used for secondary data analysis in the current study.

5.5 Materials and Methods

As discussed in the pilot study section, four different tasks were used. However, through the course of the study, these were reduced to two, which proved to be more effective for the purposes of the training and data collection procedures. The final implementation focused on the tasks described in the following sections.

5.5.1 Description of Tokens

The stimulus set consisted of lexical items containing the Spanish trill in different phonological environments. The target items included words containing the intervocalic trill /r/, such as *carrera* 'career', *carretera* 'road', *carro* 'car', *perro* 'dog', *corrida* 'run', and *torre* 'tower'. Additional items included words containing the trill in word-initial position, such as *recámara* 'bedroom', *rizado* 'curly', *regar* 'to water', and *rubio* 'blond'.

Although several trill contexts were present in the stimulus set, the acoustic and statistical analyses in this dissertation focus specifically on the production of the Spanish trill. The complete stimulus set used in the pronunciation task, including the different phonological environments, is presented in Table 5.1.

Table 5.1

Tokens Used in the Trill Production Task

Word-Medial Trill	***Word-Initial Trill***
carrera 'career'	*rizado* 'curly'
carro 'car'	*regar* 'to water'
perro 'dog'	*rubio* 'blond'

carretera 'road' *corrida 'run'* *torre 'tower'*	*recámara* 'bedroom'

These items were embedded within regular homework assignments that contained additional vocabulary items unrelated to trill production. These additional words served as distractors so that participants would not focus exclusively on the target sounds. Only the tokens containing trills were extracted and analyzed for the purposes of the present study. Each participant produced six target tokens at each assessment point.

Recordings were collected at four different time points during the semester, resulting in 24 tokens per student. With 40 participants included in the analysis, the dataset consisted of a total of 960 observations. These repeated observations are nested within individual learners, making them suitable for analysis using HLM. In other words, multiple observations are produced by the same learner across time, creating the hierarchical structure required for multilevel modeling.

5.5.2 Training Phase

Over the course of a 16-week semester (starting in week 4 and including finals week), subjects in the training group participated in a pronunciation training (F.A.T.) regimen that consisted of watching a video of a native speaker of Spanish pronouncing isolated words containing the Spanish trill in both initial and medial positions and imitating the model speaker as closely as possible. They were instructed to focus their attention on the articulatory gestures of the model speaker without being explicitly informed that the trill was the primary target sound. The non-training group, on the other hand,

watched the same video but received a different set of instructions. They were asked to listen and record themselves repeating what they heard. No emphasis was placed on imitating the gestures of the model speaker.

The stimuli were audio-visual in nature and consisted of a native speaker of Spanish was video recorded pronouncing the target words in isolation. These were presented through the Wimba platform (see Figure 5.1), where students accessed them as part of their weekly oral production homework assignments and during the four assessment time points.

The audio-visual pronunciation task was integrated into the weekly oral homework assignments throughout the semester. Although both groups completed the task, the training group received explicit focused-attention imitation instructions. Since the weekly oral homework was aligned with chapter topics covered throughout the semester, the lexical items varied across the course. Additional Wimba tasks are included in Appendix C.

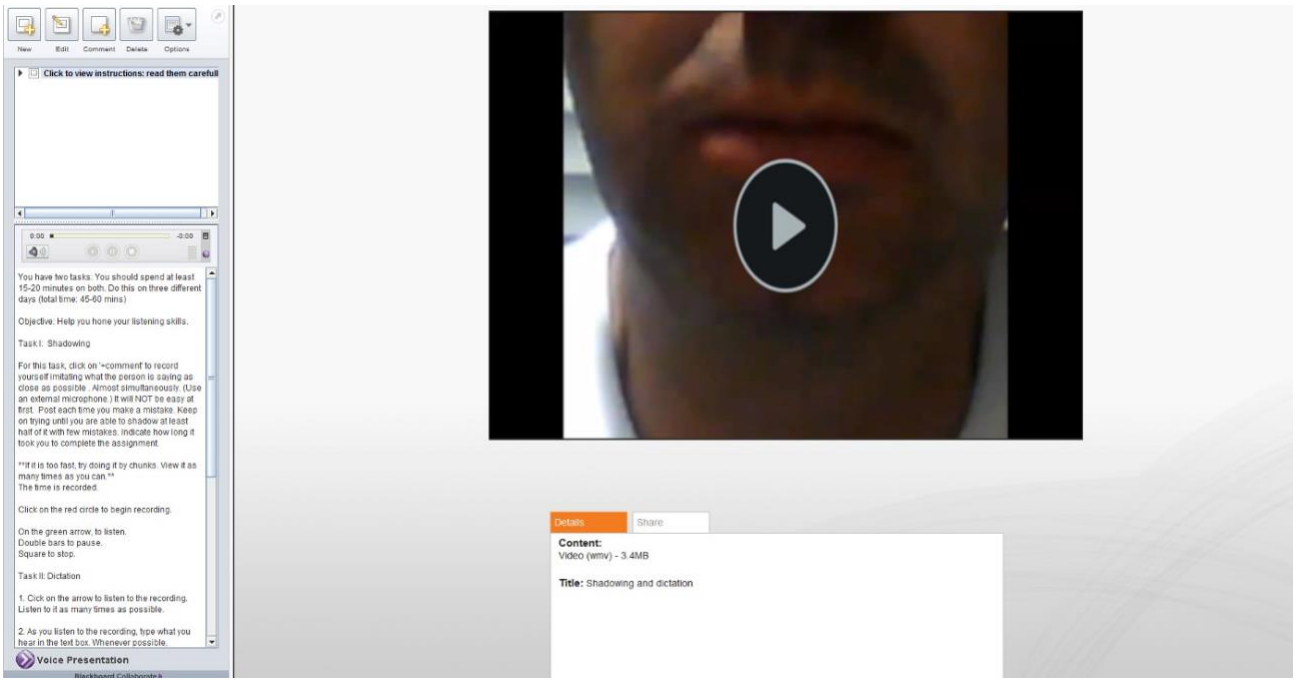


Figure 5.2. *Wimba platform audio-visual pronunciation with explicit imitation instructions for the training group.*

Students' imitations or repetitions of the native speaker were recorded using Wimba and downloaded as .wav files with a sampling frequency of 16 kHz. The recordings were analyzed acoustically in Praat version 5.3.76 (Boersma & Weenink, 2014). Students used their own computers and headsets to complete the recordings outside of class as part of their weekly homework assignments. They were instructed to complete the recordings in a quiet environment in order to minimize background noise.

Training on how to do this was provided to ensure instructions were followed correctly.

Measurements of the trill were taken every four weeks starting in week four, which was the baseline point. Thus, the assessments occurred at Time 0 (Week 4), Time 1 (Week 8), Time 2 (Week 12), and Time 3 (Week 16). The training phase took place between the weeks of assessment. Originally, Time

was coded sequentially (1-4), representing the four time points at which the data were collected. To avoid the variability that students bring to a second level language course, the time variable was centered such that Week 4 of instruction became Time 0, representing the baseline after receiving similar instruction (see Table 5.2). Thus, the intercept now represents learning at this stage rather than at the very beginning of the course.

Table 5.2

Assessment time points during the semester

Time Point	***Week of Semester***	***Function in Study***
Time 0	Week 4	Baseline measurement
Time 1	Week 8	First post-training assessment
Time 2	Week 12	Second post-training assessment
Time 3	Week 16	Final assessment

5.5.3 Analysis Process

To obtain the measurements needed for statistical analysis, participants' productions of the trill /r/ were segmented acoustically. The segmental duration, defined as the length of time it takes to produce the sound, shown at the bottom of the spectrogram, was saved in an SPSS file. Figures 5.3 and 5.4 illustrate representative segmentations of the token *rizado* 'curly' produced by a heritage speaker and an L2 speaker, respectively, using waveform,

spectrogram, and TextGrid boundaries in Praat. The trill was identified primarily through the presence of multiple lingual contacts visible in the waveform and spectrogram, as well as brief interruptions and, in some cases, frication noise between contacts. In some productions, however, it was difficult to determine the exact boundaries of the trill. In these cases, additional acoustic cues, including intensity, voicing, pitch (f0), and periodicity, were used to assist in the segmentation process. For example, regular repeating wave patterns, visible voicing energy in the lower portion of the spectrogram, and the pitch tracking line in Praat helped distinguish speech sounds from pauses, recording artifacts, or background noise and helped identify the boundaries between the trill and the surrounding vowels. Although these acoustic features were not analyzed quantitatively in the present study, they were used as supplementary cues during segmentation. *Duration* served as the key outcome in the Growth Model that will be explained in later sections.

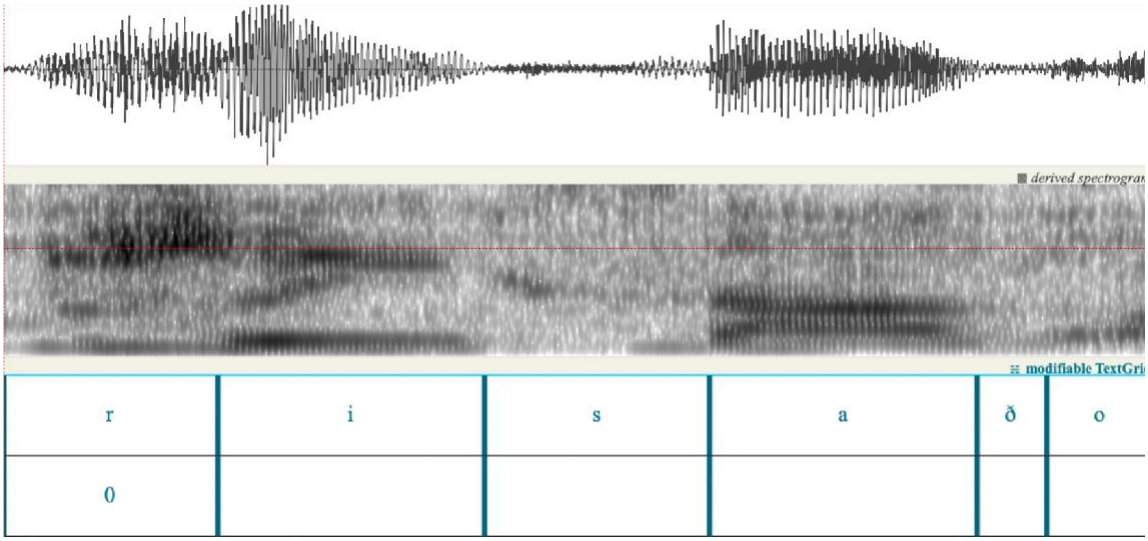


Figure 5.3. *Spectrographic representation of the production of a trill by a heritage speaker*

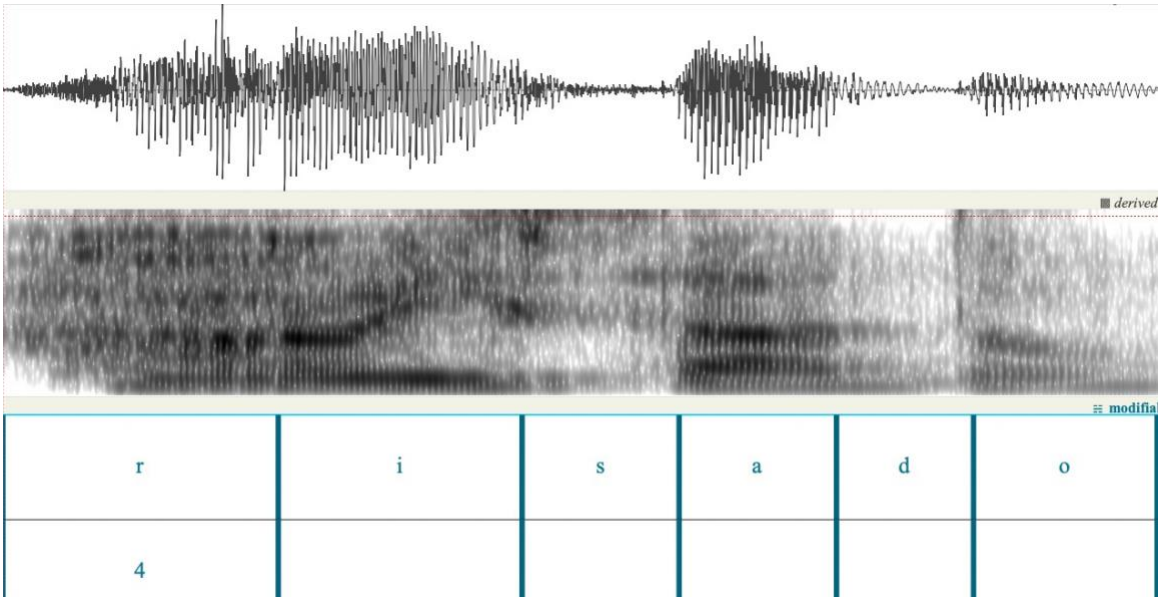


Figure 5.4. *Spectrographic representation of the production of a trill by an L2 speaker*

5.6 Acoustic Measurements

Studies focusing on the production of the trill in non-classroom settings have traditionally emphasized the successful production of these sounds with respect to a target form, often defined in terms of acoustic properties such as duration (Quilis, 1993; McDonough & Johnson, 1997; Recasens, 1986, among others) and number of lingual contacts (Bradley & Willis, 2012; Ladefoged & Maddieson, 1996; Lavoie, 2001; Lindau, 1985; Solé, 2002). The expected duration and the number of lingual contacts if any, however, can vary across dialects of Spanish. Features such as these are typically used to distinguish between the tap and trill, the /r/ vs /r/ respectively, but are

insufficient to characterize the variants that are found in US Spanish, particularly the Los Angeles area.

Assessing the production of the trill by language learners of Spanish in a classroom setting within the Los Angeles greater area presents a challenge. Should one compare the student's production to every existing dialect, or should it be compared to the variation within their community? For purposes of this dissertation, since the trill is the most difficult to produce and is absent in the L2 learners' sound repertoire, the average duration for it was calculated by taking the average measurements of Spanish heritage students' and the model's productions. This average will serve as the reference point for evaluating trill production in the present study. The heritage speakers used for these reference measurements were not participants in the experimental groups. The average duration of the Spanish trill across the heritage speakers and the model speaker production of the trill was approximately 110 ms, about 25 ms more than the one found in peninsular Spanish (Quilis, 1993).

Although duration served as the primary acoustic measure in the present study, the number of lingual occlusions was also coded as an additional characteristic of trill production. Since occlusion patterns may vary considerably across speakers and dialects, they were treated primarily as a supplementary measure rather than as the principal outcome variable.

In cases where the spectrographic analysis did not show clear occlusions, the production was interpreted as lacking the multiple lingual contacts characteristic of the trill. For this reason, these tokens were not

assigned a value for the number of occlusions, and only the occlusion variable was coded as missing in the dataset. However, these productions were retained in the analysis, as learners often compensated through increased duration. As a result, duration remained the primary measure used to assess development in trill production.

In studies that focus on training students to become more proficient or native-like in a second language, not only is it important to measure their outcomes by how close they have reached a global (standard) measure of a particular feature, but it is equally important to choose a statistical method that demonstrates the efficacy of the L2 speech training task(s) by tracking students' progress at four different time points during their training phase. A Hierarchical Linear Model would be the ideal approach.

5.7 Brief Description of the Statistical Analysis

Prior to conducting the main HLM analyses, exploratory data analyses were performed. This preliminary step is important in longitudinal multilevel research because it allows the researcher to examine the structure, variability, and overall distribution of the data before building the multilevel models. These analyses provided an overview of the distribution of the data, helped identify any potential errors, and allowed for a preliminary assessment of the relationships among the variables. In addition, these preliminary analyses helped determine whether the data exhibited the repeated and nested structure necessary for HLM, since multiple observations were collected from the same learners across time. Descriptive statistics,

scatterplots, boxplots, and other exploratory procedures were examined to evaluate variability across learners and assessment points and to inspect the overall distribution of the data prior to model estimation.

For the exploratory analysis presented in this section, a subset of twenty-five participants was selected to visualize general patterns in the data during the initial stages of statistical analysis. These participants were selected because they had observations across all assessment points, which allowed for clearer visualization of longitudinal patterns and variability over time. The HLM analysis presented later uses the full dataset of forty participants measured at four time points, although some observations are missing.

Initially, scatterplots and descriptive statistics were examined to visualize general patterns in the data across assessment points. More comprehensive descriptive analyses using the full dataset are presented in Chapter 6.

Figure 5.5 illustrates the individual average trajectories in trill duration at each time point. Overall, the trajectories suggest variability in learners' production patterns across assessment points, with some students showing increases and others decreases over time. The fact that the slopes are not flat for most students suggests that some change may be occurring over time. This variability further supports the use of a longitudinal hierarchical modeling approach.

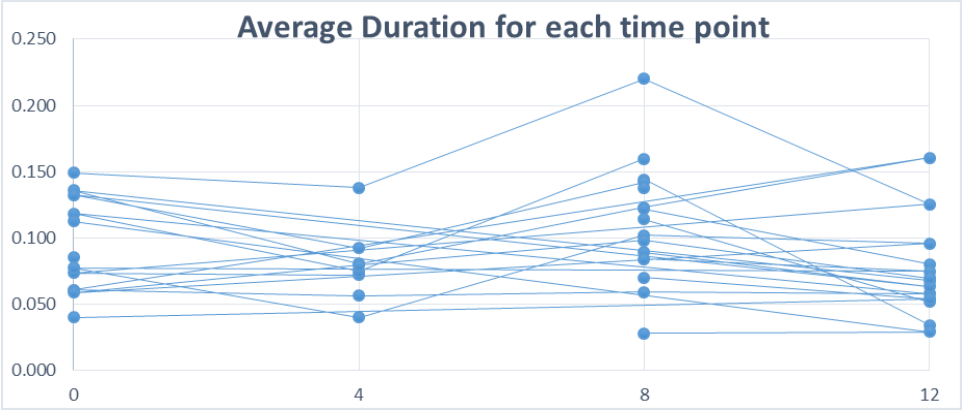


Figure 5.5. *Preliminary Individual growth trajectories in trill duration at each time point*

Figures 5.6 and 5.7 illustrate individual growth trajectories for two representative learners, one from the training group and one from the non-training group.

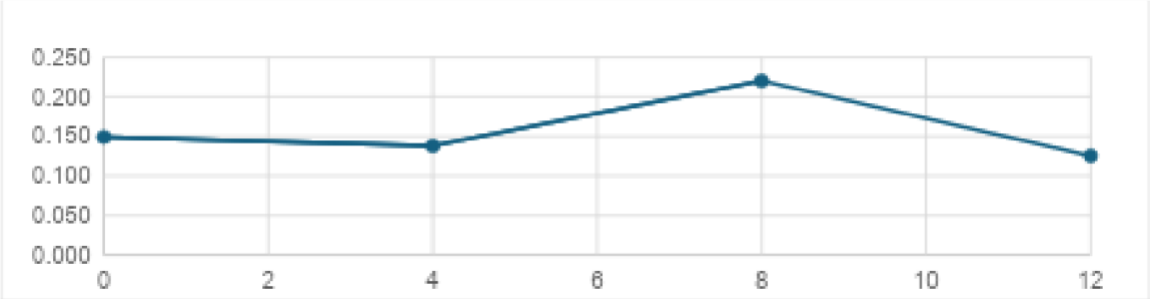


Figure 5.6. *Individual growth trajectory at each time point after training*

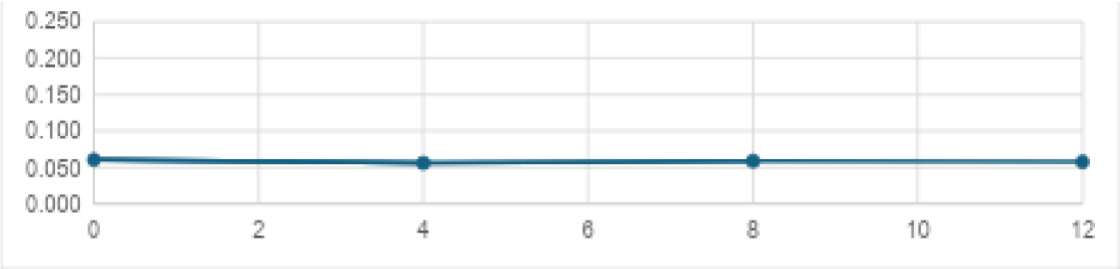


Figure 5.7. *Individual growth trajectory at each time point. No training.*

5.8 Hierarchical Data Structure and Statistical Analysis

5.8.1 Rationale and Motivation for Employing an HLM Approach

In the present study, the data consist of multiple trill tokens produced by the same learners across several time points, resulting in observations that are nested within individuals. HLM provides an appropriate framework for analyzing this type of data (Raudenbush & Bryk, 2002), as it accounts for relationships among observations within the same learner and can accommodate missing observations without modifying the structure of the dataset.

HLM was employed to analyze patterns in trill production at four different time points. Unlike traditional statistical approaches that assume observations are independent, HLM accounts for the fact that observations are related at two levels: within and between speakers, as is typically the case in classroom-based data.

Given this variability, studies that focus on assessing program success typically do so by averaging the results collected from the students. In other words, students are grouped and given a particular treatment or assessment, and effectiveness is measured based on the average of all students' outcomes (Level-2). Final (post) results are compared to a baseline (pre) by using a fixed parameter regression model. Any increment will be considered 'success' and any decrement, a failure. Additionally, results are computed without considering the individual variance that exists within the groups (Level-2),

such as gender, socioeconomic status (SES), self-esteem, motivation, among others, that can serve as predictors of a particular outcome.

This study examines how learners' production of the Spanish trill develops over time, as opposed to reaching a single target form. Though previous work has examined longitudinal development in L2 pronunciation (e.g., Nagle, 2019), it has not typically combined acoustic analysis with multilevel modeling of individual trajectories. For this reason, an approach that can account for both change over time and variability across learners is required.

Speakers learn at different rates; some may improve, plateau, or show little or no improvement over time. These differences can influence the overall averages used in traditional statistical analyses. In the present study, Level-1 captures within-individual change in trill production across time and tokens, while Level-2 accounts for differences between learners in their overall patterns of production.

To fully understand the rationale for employing an HLM approach, it is important to consider the variables and different scenarios at Levels 1 and 2. At Level-1, the model captures within-individual variation across time and tokens, including variables such as the position of the trill and the preceding or following vowels. Taking into account the Level-1 variables, we can elucidate how these factors influence the outcome of the intervention, namely the duration of the trill.

At Level-2, the model accounts for differences between learners, including characteristics such as heritage status, as well as the training condition

(Imitation vs no imitation). In the present study, the imitation group received explicit focused-attention imitation instructions, whereas the other group repeated the stimuli without explicit imitation instructions. This allows for the analysis of how different types of interventions and learner characteristics influence the overall outcomes.

In this study, HLM is used to account for both within- and between-group variance. It provides a clearer picture of what specific elements influence the results, including individual trajectories of change. Within the education, business, social work, and health fields, this approach is commonly used for accounting for these differences. It is referred to in the literature as mixed-effects models, random effects (Singer, 1998; Laird and Ware, 1982), and is commonly known as a multilevel linear model within the sociology field (Goldstein, 1995). HLM is widely used in intervention research. In this dissertation, all students completed the audio-visual tasks; however, only the imitation group received explicit focused-attention imitation instructions directing them to focus their attention on the articulatory gestures and pronunciation of the model speaker. As discussed earlier, whether students receive these focused-attention imitation instructions may influence the overall success of attaining the phones in question. However, it is important to note that this study will not focus solely on the overall success of the task. It will also analyze the trajectory of individual change in attaining the trill. For this purpose, a linear growth model will be used to analyze the repeated measures. This method for analyzing longitudinal data will shed light

on what happens between the pre- and post-assessments that are normally used in L2 research. It will help reduce noise that can affect statistical precision by analyzing data on multiple time points (Raudenbush, 2001) --- four in this study.

In addition, since not all subjects participate in each of the four assessments for several reasons --- the data of those that have taken the assessments at different time points or of those who have missing observations, will still be included in the analysis process. The inclusion of all subjects with missing data is advantageous due to what Hedeker and Gibbons (2006) explain: "[…] (a) by including all data, the analysis has increased statistical power, and (b) Complete-case analysis may suffer from biases to the extent that subjects with complete data are not representative of the larger population of subjects" (p. 48) For these reasons, HLM serves as the primary analytical approach in the present analysis.

In this study, HLM is employed to analyze the acquisition of the trill at multiple time points during a semester in second language (L2) learners. By using a growth-curve model within the HLM framework, we can unveil the true growth trajectory of students' production of the Spanish trill. This approach is necessary to understand how effective the training program is and how individual performances differ.

5.8.2 Model Specification

Since the exploratory data analysis suggests a general trend between the initial status

(baseline) and each subsequent time point, a model such as the one below was fitted to the data. This model represents a longitudinal structure in which repeated observations are nested within individuals and allows us to estimate each learner's rate of change from week 4 to week 16 of the semester. Given that the dataset contains repeated observations nested within individual learners and an unequal number of tokens across time points and lexical items, HLM was selected as the primary statistical approach. The hierarchical linear model specified for this study is shown below.

Summary of the Model Specified

Outcome Variable: *Duration*

Modeling duration as a function of weeks:

Level-1 Model (within students)

$$DURATION_{ti} = \pi_{0i} + \pi_{1i} (WEEK_{ti}) + e_{ti}$$

Where:

- π_{0i} = true initial status for student i

(The intercept represents the student's baseline trill duration when the time variable *week* is not centered.)

π_{1i} = true rate of improvement for student *i*

(This parameter represents the rate at which trill duration changes as *week* increases.)

Level-2 Model (between students)

$$\pi_{0i} = \beta_{00} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + r_{1i}$$

Where:

- β_{00} = mean initial status across students *(The average trill duration at baseline.)*
- β_{10} = mean rate of change across students *(The average rate of change in trill duration over time.)*
- r_{0i} = individual-level variability (*The extent to which individual learners differ from the average baseline trill duration.)*
- r_{1i} = random effects representing variation between students

The specification of this model requires identifying the variables that operate at each level of the hierarchical structure. In hierarchical linear models, variables are organized according to the level at which they vary (Raudenbush & Bryk, 2002). Level-1 variables capture within-student variation and include predictors that may change across tokens or time points for the same learner. In the present study, variables such as week, trill position, following vowel, stress, and duration vary across observations produced by the same student and therefore belong to Level-1.

Level-2 variables represent characteristics that describe the learner and remain constant across observations. In this study, variables such as training

condition and heritage status characterize the students themselves and therefore operate at Level-2.

Table 5.3 summarizes the Level-1 and Level-2 variables included in the hierarchical linear model. At Level-1, the variable WEEK represents the different time points at which measurements were taken and allows us to track progress over time.

The variable DURATION represents the length of time it takes for each student to produce the trill sound at each time point and serves as the primary outcome measure. In addition, the variable OCCLUSIONS represents the number of tongue contacts observed during trill production. Other Level-1 variables describing the phonological context of the trill tokens are also included in the hierarchical linear model.

The dataset consists of all tokens produced by the participants across the four assessment points. Since not all lexical items were elicited at every time point and some productions were excluded due to recording quality, the number of observations varies across tokens and participants. This statistical approach accommodates such unbalanced datasets and allows all available tokens to be included in the analysis.

At Level-2, the variables TRAINING and HERITAGE capture group characteristics that may influence student performance. TRAINING indicates whether a student received training (0 = repetition without explicit imitation instructions, 1 = focused-attention imitation training), and HERITAGE

identifies heritage speaker status (0 = heritage speaker, 1 = non-heritage speaker).

Together, these variables allow us to examine variation in student performance at the individual and group levels and to evaluate how the training intervention influences learners' development over time.

Table 5.3

Variables Included in the Hierarchical Linear Model

Level	*Variable*	*Description*
Level -1 (within students)	TIME	Time variable representing the four assessment points during the semester (Weeks 4, 8, 12, and 16).
	DURATION	Duration of the trill production measured acoustically in milliseconds. This serves as the outcome variable in the model.
		Level-1 (within students)
	OCCLUSIONS	Number of tongue contacts observed during trill production.

	POSITION	Position of the trill within the word (e.g., intervocalic or word -initial).
	FOLLOWING.VOWEL	Vowel following the trill, included to account for phonological context effects.
	STRESS	Stress of the syllable containing the trill.
Level -2 (between students)	TRAINING	Indicates whether the student received training (0 = repetition without explicit imitation instructions, 1 = focused -attention imitation training) .
	HERITAGE	Indicates heritage speaker status.

5.8.3 Description of Coefficients in the Present Study

In this study, the coefficients associated with the growth-curve within the hierarchical linear model provide insight into patterns of change in students' production of the Spanish trill over time. These parameters describe relationships between variables operating at two levels of the model: Level-1 (within students) and Level-2 (between students).

The current study measures the trill duration as a function of time across the four assessment points collected during the semester. Students' productions were measured acoustically at these time points throughout the semester. The

coefficients of the model therefore allow us to examine individual trajectories of change and variation across learners.

Table 5.4 summarizes the interpretation of each coefficient used in the model and introduces the centering procedures discussed in the next section.

Table 5.4

Interpretation of Coefficients Used in the Hierarchical Linear Model

Level	***Coefficient***	***Description***
Level-1: Within-student (e.g., Time centered)	π_{0i} = True initial status	Represents the starting point or the true initial status for each student, indicating their baseline performance.
	π_{1i} = True rate of change	Represents the rate of change at which each student's trill duration changes at each time point (*weeks*). This parameter corresponds to the slope of the student's growth trajectory.
Level-2: Betweengroup (e.g. imitation vs non-imitation group)	β_{00}= Mean initial status	Represents the average baseline trill duration across students after accounting for the nested structure of the data.

β_{10}= Mean rate of change	Represents the average rate of change in trill duration across students as the unit of measurement (*week*) increases.

5.8.4 Centering

Students entering a second-level language course often begin the semester with varying levels of proficiency. Some students have recently completed the previous course, while others may have placed into the course through a placement examination. These differences can make it difficult to establish a common starting point for measuring progress. For this reason, the training in the present study began in week 4 of the semester. The first weeks of instruction allowed students to receive the same classroom input before the training phase began. As a result, week 4 serves as the baseline measurement point for the growth model.

To ensure that the HLM model's parameters were interpreted meaningfully, time was centered at Week 4, the first assessment point following comparable classroom instruction for all learners. In HLM, centering changes the reference point from which growth is measured rather than

changing the data itself. As a result, Week 4 became Time = 0, allowing the intercept (π_{0i}) at Level 1 to represent each learner's baseline trill duration at that assessment point rather than at an arbitrary value of zero. Similarly, the Level-2 intercept (β_{00}) represents the average baseline trill duration across learners. The slope parameters (π_{1_i} and β_{10}) represent the rate of change in trill duration over time at the individual and group levels, respectively.

Centering time at Week 4 also helped reduce collinearity between the intercept and slope parameters, making the model estimates more reliable and easier to interpret. The centered time variable was coded as follows: Week 4 = 0, Week 8 = 1, Week 12 = 2, and Week 16 = 3 (see Table 5.2).

A set of hierarchical models were developed and fit using the HLM program to assess within-learner change and between-learner differences in change over time. The sequence of models estimated in the present study is summarized in Figure 5.7. It shows each step taken to examine variance, change over time, and the extent to which that change varies as a function of training and learner characteristics.

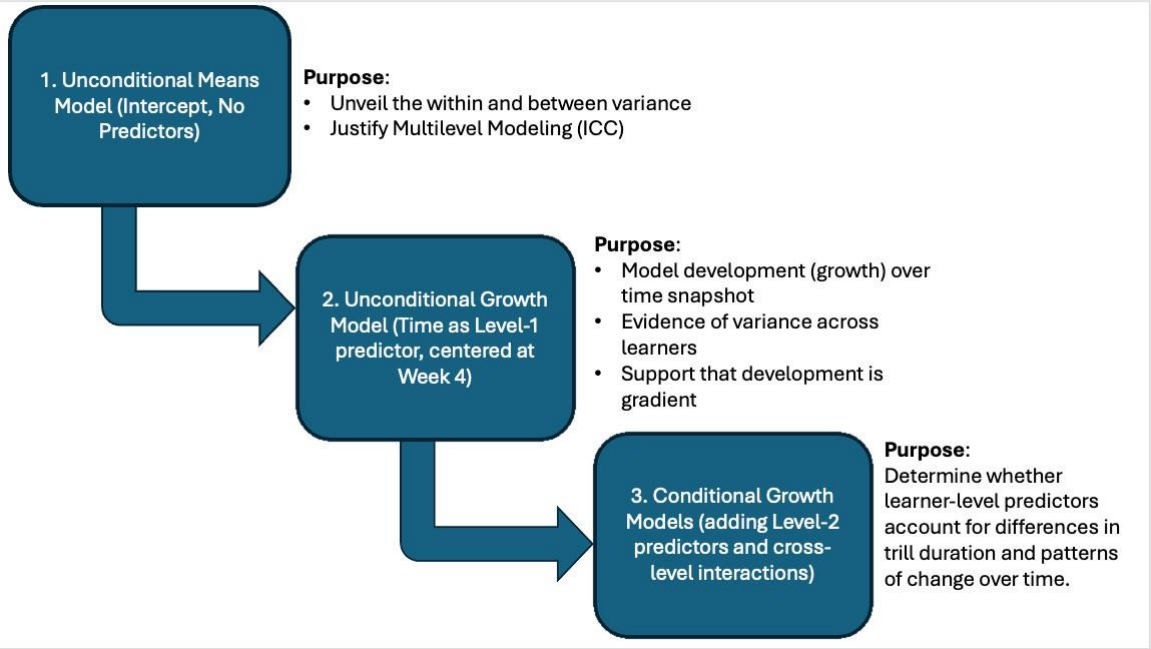


Figure 5.7. *Sequence of hierarchical linear models estimated in the study.*

Note. The final stage of the model-building process includes step-by-step conditional growth models examining training condition, baseline duration, heritage speaker status, distance from the model speaker, and occlusion production.

The first model, an unconditional means model (Model 1), contains no predictors and serves to separate the variance in trill duration into within-learner (Level-1) and between-learner (Level-2). This model was used to justify the use of a multilevel modeling framework. The Intraclass Correlation Coefficient (ICC) was computed to assess the percentage of variance that could be attributed to the within learner level.

Once it was confirmed by the ICC that a hierarchical structure was warranted, an unconditional growth model (Model 2) was estimated that

included time centered at Week 4 as a Level-1 predictor. This model captured overall development in trill duration over time and provided an estimate of individual growth trajectories, addressing the question of whether trill duration changes over time.

A series of conditional models were estimated. First, training condition was added as a Level-2 predictor of the intercept. Subsequent models included learner-level predictors to capture differences among learners. Specifically, training condition and heritage speaker status were added at Level 2 to determine whether these variables account for variance in initial status and patterns of production over time. In addition, cross-level interactions were assessed, particularly the interaction between time and training group, to examine whether the rate of change in trill duration varies as a function of training condition. Additional analyses also examined distance from the model speaker and occlusion production over time.

Having outlined the data collection and analysis procedures, as well as providing a detailed account for implementing HLM to measure the trill duration at four distinct time points, the next chapter focuses on presenting results of the present study.

Chapter 6

Results

6.1 Introduction

Statistical analyses were conducted to examine the trajectories of development in the production of the Spanish trill over time. The main outcome variable, trill duration, was analyzed at four time points during a 16-week semester.

The results are presented in phases, beginning with a descriptive overview of the data, followed by multilevel analyses that address the research questions of the study. The order of the analyses conducted reflects the sequence of models described in Chapter 5. Descriptive patterns are presented to provide an overview of general trends and variability across learners. Next, the results of the hierarchical linear models are introduced, beginning with the unconditional means model used to assess the proportion of variance attributable to the learner level. The final models examine change over time and differences between groups, including the effect of training on trill production.

6.2 Descriptive Overview of the Data

6.2.1 Preliminary Exploratory Data Analysis (EDA)

To examine patterns of development across time, descriptive analyses were conducted using both summary statistics and visualizations of individual

trajectories, as well as the distribution of trill duration across learners. Prior to the full analyses presented in this chapter, a preliminary exploratory analysis was conducted using a smaller subset of participants. Although the number of subjects at this time was insufficient to draw any definitive conclusions, the data did tend to show a general trend among the nonheritage speakers.

For this preliminary exploratory analysis, twenty-five students were assessed at the four different time points, and out of these, nine participants were heritage speakers and were excluded from this preliminary exploratory analysis because their production appeared to differ from those of the non-heritage speakers. However, heritage speakers are included in the full HLM analysis presented later in this chapter, and their productions will be examined separately in a subsequent analysis.

The heritage speakers tended to produce higher trill durations at the beginning, possibly due to hypercorrection, and progressively decreased over time. The nonheritage speakers, on the other hand, tended to show the opposite pattern. To observe a more linear upward or downward trend, the separation of these two types of learners was warranted. For this preliminary exploratory analysis, the tables, graphs, and descriptive analyses include only the non-heritage speakers. However, both heritage and non-heritage speakers were included in the full HLM analyses presented later in this chapter.

Table 6.1 shows the average trill duration for each assessment point. Overall, the results suggest some degree of change over time.

Table 6.1

Mean Trill Duration (ms) by Assessment Time (Exploratory Analysis)

Week	N	Mean	Std. Deviation
0	9	83	.032384
4	5	66	.017622
8	12	99	.036718
12	11	70	.035664
Total	37	82	.034830

Note. In this exploratory analysis, the time variable is centered so that the baseline assessment collected in Week 4 of the semester corresponds to Time 0. The remaining assessment points correspond to Time 1, Time 2, and Time 3, representing Weeks 8, 12, and 16 of the semester, respectively. Duration values were extracted from Praat and are reported in milliseconds in this table, ranging from approximately to 66–99 ms across the assessment points.

The variability observed during this preliminary stage is consistent with the rationale for employing HLM to account for variability both within and between learners over time.

6.2.2 Descriptive Overview of the Full Dataset

The full dataset was then examined by group (trained vs. non-trained) across the four assessment points. Table 6.2 shows mean trill duration and standard

deviations at each time point by group. We observe an increase in mean duration from the initial time point to Time 2, followed by a decrease at the final measurement. Differences between the trained and non-trained groups are minimal at each time point. At the same time, the relatively large standard deviations indicate substantial variability within both groups, suggesting that learners differ considerably in their production of trill duration. The magnitude of the standard deviations remains relatively stable across time points, indicating that variability persists throughout the study. This suggests that group-level averages may mask important differences in individual developmental trajectories, which will be further analyzed in this study.

Table 6.2

Mean Trill Duration (ms) by Time Point and Group

Time (Centered at Week 4)	***No Training Mean (SD)***	***Training Mean (SD)***
0	*103.47 (38.47)*	*102.66 (41.31)*
1	104.95 (33.11)	100.56 (35.89)
2	116.98 (35.18)	113.94 (35.45)
3	100.68 (36.27)	101.35 (32.90)

Note. *Time 0 corresponds to Week 4 (baseline after initial instruction). Values represent trill duration in milliseconds.*

Figure 6.1 shows the progression of trill duration across the four time points and illustrates both individual and overall trends.

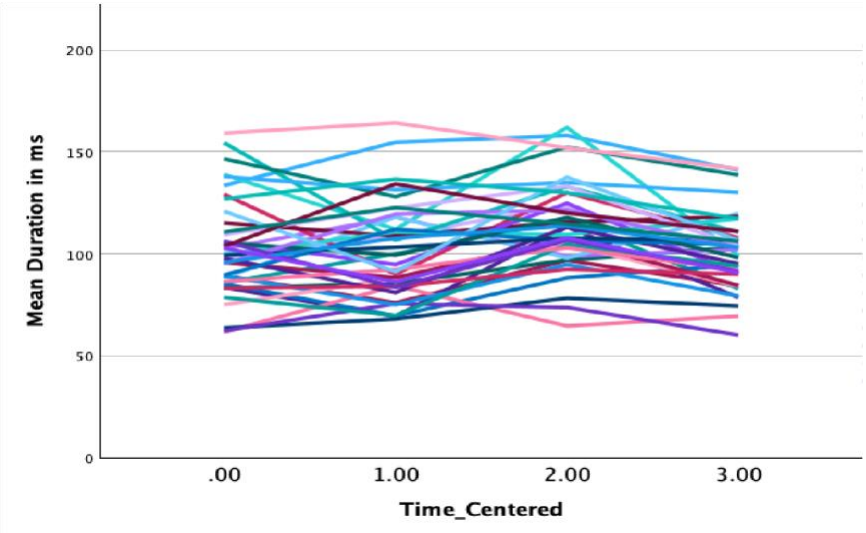


Figure 6.1. *Individual trajectories of trill duration across four time points*

As depicted in the spaghetti plot, one can see quite a bit of variability both within and across individual trajectories. The changes from the first assessment point to the second are not consistent across learners. Trill durations show both increases and decreases across learners. At Time 3, durations tend to be highest across trajectories. At the final point, several values decrease, although some increase slightly or remain stable.

Development is therefore not consistent across learners and varies over time. While a general pattern of change can be observed, individual learners differ considerably in the magnitude and direction of their

development. The spaghetti plots further show that, although both trained and non-trained groups appear to follow a similar overall pattern, differences between groups are not immediately evident at the descriptive level. Since an effect of training is not clearly observable in the raw trajectories, further statistical analysis is needed to formally examine change over time and differences between groups.

To further examine the distribution of trill duration across time, Figure 6.2 presents boxplots for each time point. As shown, the data are spread at all time points, and overlap across time points is evident, with similar ranges of values observed at each time point. The outliers at the first two assessment points appear more widely dispersed, whereas those at the final time point are more closely clustered around the central tendency. This suggests a slight reduction in the spread of extreme values. In other words, the trill /r/ productions at the final time point appear somewhat less dispersed than those observed earlier in the semester. However, variability remains evident across all time points.

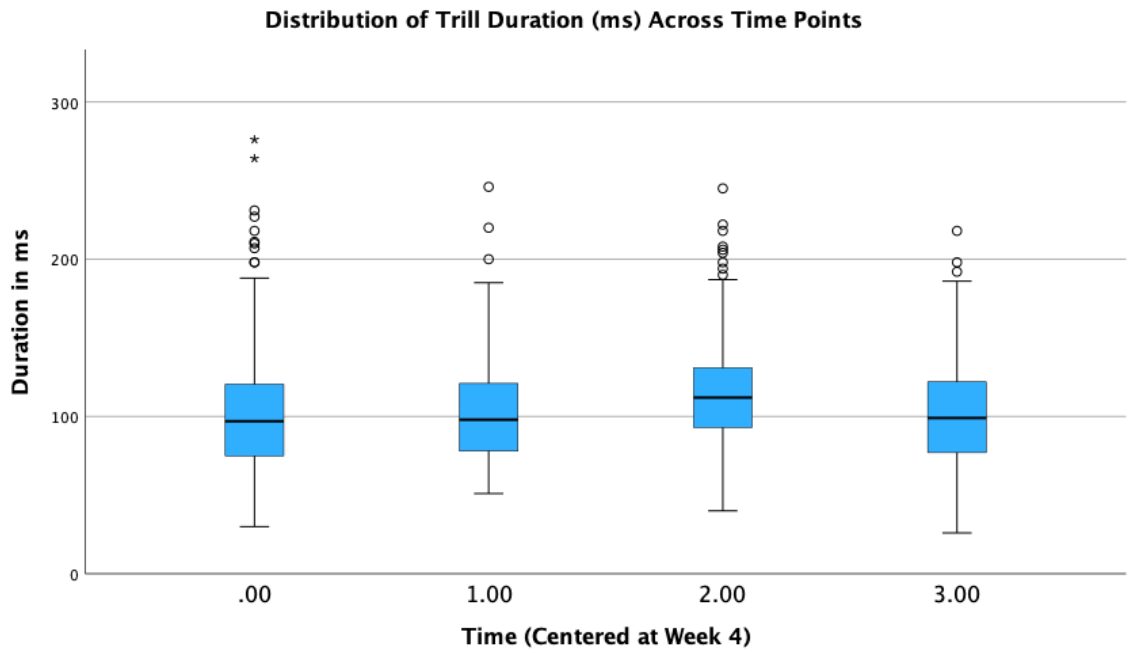


Figure 6.2. *Distribution of trill duration (ms) across time points.*

While the descriptive analyses presented in this section suggest variability across learners and over time, they do not allow for formal evaluation of change or group differences. In other words, descriptive patterns alone cannot be determined by the fact whether these differences are statistically reliable. Multilevel modeling is therefore employed.

6.3 HLM Results

6.3.1 Unconditional Means Model

To begin the model-building process and support the implementation of HLM, an unconditional means model (Model 1) was run. This model contains no predictors and serves to partition the variance in trill duration into within-learner

and between-learner components. This model was used to determine the extent to which variation in trill duration is attributable to differences between learners, as opposed to variation within learners across time.

The results indicate that a portion of the variance in trill duration is associated with differences between learners (Level-2), while the remaining variance reflects variation within learners (Level-1) over time. The Intraclass Correlation Coefficient (ICC) was calculated to estimate the proportion of total variance attributable to the learner level. The ICC indicated that approximately 26% of the variance in trill duration is due to differences between learners (τ_{00} = 0.00035), with the remaining variance reflecting within-learner variation across time (σ^2 = 0.00101). Estimates for fixed and random effects are presented in Table 6.3.

Table 6.3

Unconditional Means Model

Parameter	*Estimate*	*SE*	*p-value*
Fixed Effects			
Intercept (γ_{00})	0.104624	0.003136	< .001
Random Effects			
Intercept variance (τ_{00})	0.00035		
Residual variance (σ^2)	0.00101		
ICC	0.26		

These results suggest that observations are not independent, meaning that the repeated observations come from the same learner over time rather than from entirely separate individuals. Due to this nested structure, a multilevel modeling framework is appropriate for the present data.

6.3.2 Unconditional Growth Model (No Level-2 Predictors)

Having established the distribution of variance in the first model, a second model (Model 2) was run, adding time as a Level-1 predictor to examine change in trill duration over time.

The results indicate that the effect of time was not statistically significant (γ_{10} = 0.00024, *p* = .702), suggesting that, on average, trill duration did not show a consistent linear increase across time points, despite variability in individual trajectories. Estimates for fixed and random effects are presented in Table 6.4.

Table 6.4

Unconditional Growth Model Estimates

Parameter	*Estimate*	*SE*	*p-value*
Fixed Effects			
Intercept (γ_{00})	0.10425	0.00369	< .001
Time (γ_{10})	0.00024	0.00064	.702
Random Effects			
Intercept variance (τ_{00})	0.00035		
Residual variance (σ^2)	0.00101		

6.3.3 Training Predicting Initial Status

To determine whether learners differ in their initial trill duration (Time 0 = Week 4) as a function of training condition, a third model (Model 3) was estimated in which training was added to the intercept as a Level-2 predictor.

The results of Model 3 indicate that the effect of training condition on the intercept was not statistically significant (γ_{01} = -0.002028, *p* = .749), suggesting that the training and non-training groups did not differ significantly in their average trill duration at the baseline assessment (Time 0). The estimated intercept (γ_{00} = 0.105312, *p* < .001) indicates that the average trill duration at baseline for the reference group (non-training group, coded as 0) was approximately 105 ms. Estimates for fixed and random effects are presented in Table 6.5.

Table 6.5

Training Predicting Initial Status

Parameter	Estimate	SE	p-value
Fixed Effects			
Intercept (γ_{00})	0.105312	0.005253	< .001
Training (γ_{01})	-0.002028	0.006304	.749
Random Effects			
Intercept variance (τ_{00})	0.00036		
Residual variance (σ^2)	0.00101		

6.3.4 Full Model (Training and Predictors of Growth)

Building on Model 3, which examined whether training condition predicted learners' initial trill duration at baseline, Model 4A represents a full model that includes training and predictors of growth, including a cross-level interaction between training and time. It addresses the key research questions of the current study: whether learners differ in their development, whether training affects the development of individual trajectories,

and whether these patterns vary across learners. Later models further examine whether these developmental patterns vary as a function of heritage speaker status.

The results of Model 4A indicate that the interaction between training condition and time (centered at Week 4) was not statistically significant (γ_{11} = 0.000648, *p* = .608), which suggests that the rate of change in trill duration did not differ between the training and non-training groups. In other words, the rate of change over time (slope) did not differ between the training and non-training groups. Both groups appeared to follow similar developmental patterns across the semester. The main effect of time was also not statistically significant (γ_{10} = -0.000094, *p* = .915), indicating that trill duration did not show a consistent linear change over time across learners.

Additionally, the effect of training condition on the intercept remained non-significant (γ_{01} = -0.003018, *p* = .686), suggesting that learners assigned to either condition (trained or non-trained) did not differ in their initial trill duration at Time 0 (centered at Week 4). The estimated intercept (γ_{00} = 0.105827, *p* < .001) indicates that the average trill duration at baseline for the reference group (non-training group, coded as 0) was approximately 106 ms. Estimates for fixed and random effects are presented in Table 6.6. For full model specifications and output, see Appendices H–L. These findings should be interpreted with caution, given the considerable variability observed in learners' initial trill durations and developmental trajectories.

Table 6.6

Full Model Predicting Trill Duration

Parameter	Estimate	SE	p-value
Fixed Effects			
Intercept (γ_{00})	0.105827	0.005693	< .001
Training (γ_{01})	-0.003018	0.007415	.686
Time (γ_{10})	-0.000094	0.000879	.915
Training × Time (γ_{11})	0.000648	0.001264	.608
Random Effects			
Intercept variance (τ_{00})	0.00036		
Residual variance (σ^2)	0.00101		

6.3.4.1 Full Growth Model Including Baseline Duration

Based on the non-significant results of Model 4A, which addresses the question of whether there is a training effect over time, further analysis was warranted. Model 4B was estimated to determine whether students' starting point could influence developmental trajectories. The results indicate that development depends on where learners begin. In this sense, the benefits of training appear to be more prominent for learners starting furthest from the target model. Conversely, learners who are already closer to the target tend to show more limited changes in their developmental trajectories.

The results of Model 4B, presented in Table 6.7, indicate that the baseline trill duration significantly predicted learners' developmental trajectories across time. The effect of baseline duration on the intercept was statistically significant (γ_{01} = 0.001140,

$p < .001$), indicating that learners' initial trill durations were associated with their overall production patterns. Additionally, the interaction between baseline duration and time was statistically significant (γ_{11} = -0.000093, p = .005), suggesting that learners with lower initial duration values exhibited greater increases over time, whereas learners with higher initial duration values, possibly reflecting evidence of hypercorrection, showed more limited change across the semester.

The main effect of time was also statistically significant in this model (γ_{10} = 0.009692, p = .005), indicating that trill duration increased over time once learners' starting points were taken into account. In other words, after including the differences observed at the starting point, students tended to show increases in trill duration across the semester. However, neither the effect of training on the intercept (γ_{02} = -0.000797, p = .647) nor the interaction between training and time (γ_{12} = 0.000497, p = .661) were statistically significant. The non-significant intercept effect was not unexpected, as learners had not yet undergone training at the baseline assessment point. The nonsignificant interaction between training and time suggests that the effects of imitation training are not uniform across learners but may depend on learners' initial production levels and distance from the target model.

The variance component for the intercept approached zero (τ_{00} = 0.00000), indicating that much of the variability between learners' initial status was accounted for by the predictors included in the model. The residual variance remained at σ^2 = 0.00096, reflecting continued variability within learners across time points.

Table 6.7

Full Growth Model Including Baseline Duration

Parameter	***Estimate***	***SE***	***p-value***

Fixed Effects			
Intercept (γ_{00})	-0.014625	0.005234	.008
Baseline Duration (γ_{01})	0.001140	0.000049	< .001
Training (γ_{02})	-0.000797	0.001727	.647
Time (γ_{10})	0.009692	0.003470	.005
Baseline × Time (γ_{11})	-0.000093	0.000033	.005
Training × Time (γ_{12})	0.000497	0.001132	.661
Random Effects			
Intercept variance (τ_{00})	0.00000		
Residual variance (σ^2)	0.00096		

6.3.4.2 Full Growth Model Including Heritage Speaker Status and Training Condition

Unlike the previous models, this conditional growth model included both heritage and non-heritage speakers. Since heritage speaker status may contribute to variability in learners' development trajectories, an additional conditional growth model was estimated that included heritage speaker status and training condition as Level-2 predictors. Results did not reveal statistically significant differences between heritage speakers and L2 learners at baseline (γ_{01} = -0.004178, *p* = .587) or in their rate of change over time (γ_{11} = 0.000170, *p* = .889). Similarly, training condition was not a significant predictor of either initial trill duration (γ_{02} = -0.003462, *p* = .652) or rate of change over time (γ_{12} = 0.000674, *p* = .580). Estimates for fixed and random effects are presented in Table 6.8. These findings suggest that neither heritage speaker status nor training condition systematically explained learners' developmental trajectories across the semester. However, substantial variability remained across learners, indicating that developmental patterns were not uniform across individuals.

As will be discussed further in Chapter 7, variability across heritage speakers suggests that additional factors may have influenced learners' trajectories over time.

Table 6.8

Full Growth Model Including Heritage Speaker Status and Training Condition

Parameter	Estimate	SE	p-value
Fixed Effects			
Intercept (γ_{00})	0.108032	0.007204	< .001
Heritage Speaker Status (γ_{01})	-0.004178	0.007621	.587
Training (γ_{02})	-0.003462	0.007605	.652
Time (γ_{10})	-0.000185	0.000984	.851
Heritage Speaker Status × Time (γ_{11})	0.000170	0.001217	.889
Training × Time (γ_{12})	0.000674	0.001218	.580
Random Effects			
Intercept variance (τ_{00})	0.00037		
Residual variance (σ^2)	0.00101		

In addition to the primary models examining trill duration, further analyses were conducted to explore related aspects of learners' production.

6.3.5 Additional Analyses

6.3.5.1 Distance from Model Speaker

The first additional analysis examined whether learners moved closer to the model speaker over time, using distance from the model speaker as the outcome variable. A multilevel model was estimated with time (centered at Week 4) at Level-1 and training condition at Level-2.

The results indicate that the effect of time was not statistically significant (γ_{10} = 0.000186, p = .763), suggesting that distance from the model speaker did not change

consistently over time. The interaction between the training condition (trained vs nontrained) and time was also not statistically significant (γ_{11} = -0.001092, *p* = .278), indicating that it did not influence the rate at which learners moved closer to the model speaker. In addition, no significant differences were observed between groups at baseline (γ_{01} = 0.005477, *p* = .258). Estimates for fixed and random effects are presented in Table 6.9. However, the absence of statistically significant group-level differences does not necessarily indicate the absence of individual developmental changes.

As will be discussed further in Chapter 7, variability across learners suggests that additional factors may influence development in ways that are not fully captured through group-level statistical analyses alone.

Table 6.9

Conditional Growth Model Predicting Distance from the Model Speaker

Parameter	***Estimate***	***SE***	***p-value***
Fixed Effects			
Intercept (γ_{00})	0.045557	0.003551	< .001
Training (γ_{01})	0.005477	0.004769	.258
Time (γ_{10})	-0.000186	0.000616	.763
Training × Time (γ_{11})	-0.001092	0.001005	.278
Random Effects			
Intercept variance (τ_{00})	0.00018		
Residual variance (σ^2)	0.00086		

Note. The dependent variable represents the absolute difference between learner trill duration and the model speaker's trill duration.

6.3.5.2 Distance from Model Speaker Including Heritage Status

A conditional growth model that includes heritage speaker status was estimated to determine whether heritage and non-heritage learners differed in their baseline distance from the model speaker and in their patterns of change over time. Heritage speaker status and training condition were included as Level-2 predictors of both the intercept and slope. Results are presented in Table 6.10.

Results indicate that heritage speaker status did not significantly predict change over time (γ_{11} = -0.000699, *p* = .525), which suggests that heritage and non-heritage learners did not differ significantly in their developmental trajectories across the semester. Training condition also did not significantly predict change over time (γ_{12} = 0.001194, *p* = .270). Although heritage speaker status approached significance as a predictor of baseline distance from the model speaker (γ_{01} = 0.008643, *p* = .077), this effect did not reach conventional significance levels. Significant variability remained across learners at the intercept level (τ_{00}= 0.00017, *p* < .001).

Table 6.10

Conditional Growth Model Predicting Distance from the Model Speaker Including Heritage Status

Parameter	*Estimate*	*SE*	*p-value*
Fixed Effects			
Intercept (γ_{00})	0.040986	0.004393	< .001
Heritage Status (γ_{01})	0.008643	0.004758	.077
Training (γ_{02})	0.006435	0.004729	.182
Time (γ_{10})	0.000186	0.000985	.850
Heritage Status × Time (γ_{11})	-0.000699	0.001098	.525
Training × Time (γ_{12})	-0.001194	0.001081	.270

Random Effects	
Intercept variance (τ_{00})	0.00017
Residual variance (σ^2)	0.00086

Note. The dependent variable represents the absolute difference between learner trill duration and the model speaker's trill duration.

6.3.5.3 Occlusions as Outcome (Poisson Model)

In addition to the distance from the model speaker, a further analysis was conducted to investigate whether the number of occlusions changed over time. A Poisson HLM model was estimated to account for the count nature of this outcome variable. Time, centered at Week 4, was entered as a Level-1 predictor with a random intercept specified at the learner level.

In Table 6.11, results indicate that the effect of time was not significant (γ_{10} = 0.004547, p = .524), suggesting that occlusion counts, or the number of times the tongue made contact with the alveolar ridge during trill production, did not change systematically across the semester. However, significant variability was observed between learners in overall occlusion production (τ_{00} = 0.03873, p < .001).

Table 6.11

Poisson HLM Model Predicting Occlusion Counts Over Time

Parameter	***Estimate***	***SE***	***p-value***
Fixed Effects			
Intercept (γ_{00})	0.835019	0.044569	< .001
Time Centered (γ_{10})	-0.004547	0.007140	.524
Random Effects			

Intercept variance (τ_{00})	0.03873

Note. *In Poisson models, the Level-1 variance is not estimated separately; therefore, no residual variance (σ²) is reported*.

To further illustrate the distribution of occlusions at each time point, a boxplot was created. As shown in Figure 6.3, occlusion counts exhibit variability across time points, including the differences in spread and the presence of outliers, with no clear pattern of increase over time.

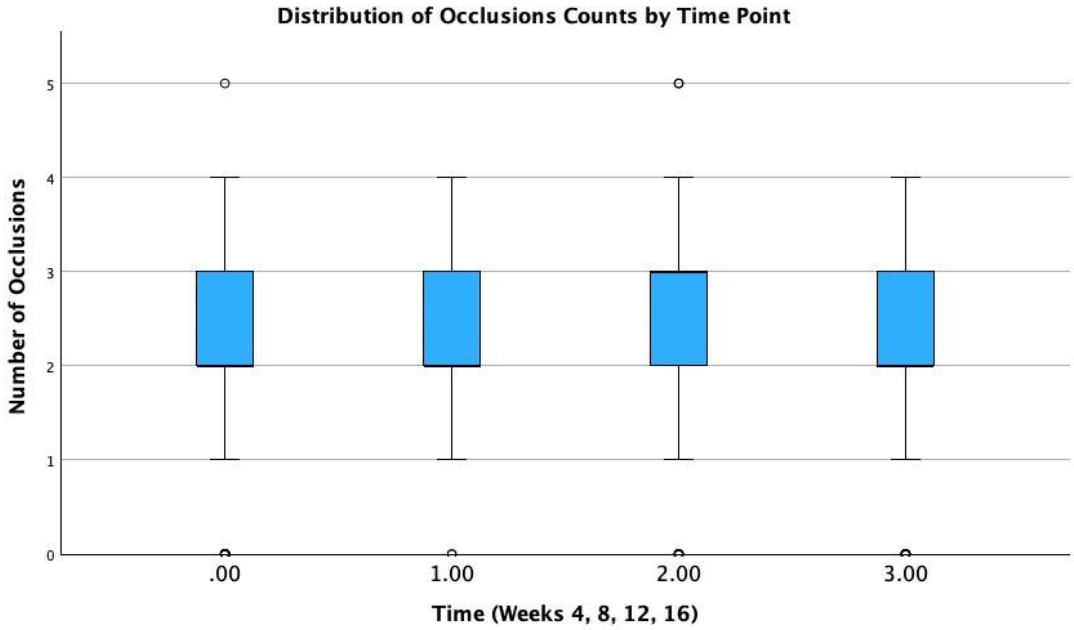


Figure 6.3. *Distribution of Occlusions by Time Point*

An additional Poisson HLM model including heritage speaker status, training condition, and trill duration as predictors was also estimated. Table 6.12 presents the results of the Poisson HLM model predicting occlusion counts as a function of time, trill duration, heritage speaker status, and training condition. The results indicate that the

effect of time became statistically significant in the expanded model (γ_{10} = -0.024584, *p* = .026), suggesting a slight decrease in occlusion counts across the semester. This decrease may also reflect the variability previously observed among some heritage speakers, who appeared to hypercorrect their productions during the earlier assessment points. Trill duration significantly predicted occlusion counts (γ_{20} = 5.476051, *p* < .001), indicating that longer trill durations were linked to greater numbers of occlusions.

Heritage speaker status was also a significant Level-2 predictor (γ_{01} = -0.165828, *p* = .010), whereas training condition did not significantly predict occlusion production or developmental change over time. The variance component for the intercept approached non-significance (τ_{00} = 0.00405, *p* = .076), suggesting reduced variability between learners after accounting for the predictors included in the model.

Table 6.12

Poisson HLM Model Predicting Occlusion Counts as a Function of Time, Trill Duration, Heritage Speaker Status, and Training Condition

Parameter	***Estimate***	***SE***	***p-value***
Fixed Effects			
Intercept (γ_{00})	0.320586	0.078452	< .001
Heritage Speaker Status (γ_{01})	-0.165828	0.060940	.010
Training (γ_{02})	0.005348	0.057425	.926
Time (γ_{10})	-0.024584	0.011040	.026
Heritage Speaker Status × Time (γ_{11})	0.025571	0.014605	.080
Training × Time (γ_{12})	0.019236	0.013972	.169
Trill Duration (γ_{20})	5.476051	0.479322	< .001
Random Effects			

Intercept variance (τ_{00})	0.00405	.076

6.4 Summary of Findings

The present study examined the development of Spanish trill production among adult English-speaking learners over the course of a 16-week semester. Using a longitudinal design with four assessment points, this study aimed to determine whether learners exhibited changes in trill duration over time and whether these changes were influenced by training conditions (training vs non-training). Additional analyses were conducted to further examine learners' proximity to a model speaker, using a study-specific reference value for trill duration, as well as patterns in the production of occlusions. Across all analyses, the findings indicate that development was not uniform across learners and was not systematically explained by time or training condition.

Descriptive analyses revealed considerable variability in trill duration both within and across learners at all four time points. While group means suggest a slight increase in duration from the initial time point to the intermediate assessments, followed by a decrease at the final measurement, this pattern is not consistent across individuals. The spaghetti plots illustrate that learners follow different trajectories, with some showing increases in duration over time, others showing decreases, and others remaining relatively stable. At the descriptive level, no clear differences are observed between the trained and non-trained groups, and variability remains evident across all time points. The boxplots further show substantial overlap in the distribution of trill

duration, indicating that learners' productions do not converge toward a single pattern over time.

The unconditional means model confirmed that a portion of the variance in trill duration is attributable to differences between learners. The Intraclass Correlation Coefficient indicated that approximately 26% of the total variance was associated with the learner level, with the remaining variance reflecting within-learner variability across time. This finding supports the use of a multilevel modeling framework and highlights the importance of accounting for individual differences when examining phonetic development.

The unconditional growth model, which included time as a Level-1 predictor, did not reveal a statistically significant effect of time on trill duration. This indicates that, on average across learners, trill duration did not show a consistent linear increase across the four time points. Although descriptive patterns suggest some fluctuation in duration over time, these changes are not systematic when modeled statistically.

Subsequent models examined whether training condition influenced learners' initial status or rate of development. The model predicting initial status indicated that there were no statistically significant differences between the training and non-training groups at baseline. Similarly, the full model, which included training, time, and their interaction, did not reveal significant effects of training condition on either the intercept or the rate of change over time. The interaction between training and time was also not statistically significant, indicating that learners in the training and non-training groups did not differ in their developmental trajectories. Although statistically significant group-level differences between the training vs non-training groups were not observed,

variability in individual trajectories suggests that learners did not respond uniformly to the focused-attention imitation training.

An additional conditional growth model (Model 4B), which incorporated learners' baseline trill duration, showed that their starting points (the trill duration at the first time point) influenced the development across the semester. Specifically, learners with lower initial trill duration values showed greater increases over time, whereas learners who were already closer to the target model tended to show smaller changes or none. These findings suggest that developmental patterns may depend in part on learners' initial production levels and distance from the target model, even when group-level training effects are not statistically significant.

In addition to the primary analyses of trill duration, further models were estimated to examine related aspects of learners' production. The analysis of distance from the model speaker was conducted to determine whether learners moved closer to the model over time and whether this change was influenced by training condition. The results indicated that neither time nor training condition, nor their interaction, had a statistically significant effect on distance from the model speaker. An additional model including heritage speaker status suggested that heritage speakers may have been somewhat closer to the model speaker at baseline than non-heritage learners, although this effect did not reach conventional significance levels. Heritage speaker status also did not significantly predict developmental trajectories over time. Overall, these findings suggest that learners did not systematically converge toward the model speaker over time. Variability in learners' proximity to the model speaker was observed across individuals.

A final analysis examined the number of occlusions produced in trill realizations using a Poisson HLM model. This analysis was conducted to determine whether learners exhibited changes in the articulatory properties of trill production over time. The results indicated that the effect of time was not statistically significant, suggesting that the number of occlusions did not change systematically across the semester. However, significant variability was observed between learners in overall occlusion production. The boxplot of occlusion counts further supports these findings, showing similar distributions across time points, with no clear pattern of increase or decrease. In the expanded Poisson model, trill duration and heritage speaker status significantly predicted occlusion counts, suggesting that articulatory variability differed across learners and production patterns.

Accounting for both within- and between-learner variability, the findings across all models indicate that trill production did not follow a consistent pattern of development on average across learners. No statistically significant effects of time, training condition, or their interaction were observed for trill duration, distance from the model speaker, or occlusion counts. At the same time, substantial variability was observed both within and between learners across all analyses. These findings suggest that development in L2 trill production does not follow a uniform or linear trajectory and is not explained by time or training condition on average across learners. These results are addressed in the following chapter, where their implications for L2 phonological development and focused-attention imitation training are discussed.

Chapter 7

Discussion

7.1 Discussion Overview

The present study focused on training adult English-speaking learners in L2 phonology, specifically the Spanish trill. The task design consisted of a longitudinal audio-visual (imitation-based) task that was created and incorporated into the students' regular weekly oral homework throughout the semester. Students were randomly assigned to either the training (imitation) or non-training group (non-imitation), with the former given explicit instructions to follow the model speaker as closely as possible, while the latter repeated without those explicit instructions. To examine trill development across condition and time, a series of statistical analyses was conducted at the individual (Level-1) and between (Level-2) levels. The results showed that trill production did not exhibit a statistically significant pattern of change over time across learners. Moreover, no statistically significant differences were observed between the training and non-training groups. However, the analyses revealed considerable variability (e.g., different trill durations) both within and between learners across all models and outcome measures. Thus, these findings suggest that learners do not develop the trill in the same manner or at the same rate. This pattern is consistent with research on gradient phonological development (e.g., Rose, 2010) and variability in Spanish rhotic production (e.g., Bradley & Willis, 2012). Based on this, it is evident that development cannot be fully captured by analyses that rely solely on aggregated group-means. In

other words, the focus is on computing a single mean score per group and then comparing the groups.

Moreover, this study aligned with a considerable body of research that showed how explicit phonetic training, particularly imitation-based approaches, can help learners improve their pronunciation (e.g., Hardison, 2003; Saito, 2013; Foote & Trofimovich, 2014). Like prior studies, the current training task required learners to focus their attention on a model speaker, including articulatory gestures, to produce tokens containing the Spanish trill as closely as possible. It is important to note that this emphasis on directing learners' attention to a specific aspect of the speech signal is what distinguishes imitation-based tasks from passive forms of exposure and becomes central to the interpretation of the present findings. In the present study, a focus on attention guided the interpretation of developmental patterns. Emphasis was placed on the variability across individual learners and the lack of consistent developmental change overall.

As introduced in Chapter 3, *Focused-Attention Training (F.A.T.)* refers to the learners' effort to direct their attention to and imitate specific articulatory features, a process that may develop and be refined over time. In the present study, F.A.T. helps interpret how learners engaged with the imitation task, particularly in how they focused their attention and attempted to reproduce the Spanish trill. While imitation-based phonetic training has been examined in prior research, here F.A.T. is adopted as a working descriptive term grounded in these observations. Surprisingly, although the instructions to imitate were specifically embedded within the task for the training group, imitation was observed in some members of the other group (non-training group). Thus,

this suggests that imitation can occur even without explicit instruction. However, when learners were asked to imitate, their productions often reflected a higher level of effort, at times and for some, resulting in exaggerated realizations of the Spanish trill. The variability observed across learners may therefore reflect differences in how attention was directed during the task instead of simply noting the presence or absence of imitation. In this sense, F.A.T. helps explain both exaggerated early productions and non-linear patterns of development as part of the learning process, rather than as evidence that the training was unsuccessful.

Based on the findings of this study, it is important to note that the lack of statistically significant fixed effects (e.g., Time and Time x Training interaction) does not rule out development. Trill development was observed in some students, but the change was inconsistent in the direction and magnitude across learners and therefore may not be fully captured by fixed effect estimates alone. In other words, within the HLM framework, the fixed effects summarize average trends, which can hide learner-specific, non-linear trajectories. This is a recognized limitation of overall trends at the group level in longitudinal SLA research (e.g., Raudenbush and Bryk, 2002). In this context, it becomes clear that pronunciation instruction does not always lead to consistent or measurable improvement in classroom settings. This further motivates the need for approaches that track development over time rather than relying on pre- and post-test comparisons. This concern, which is commonly raised in pedagogical practice and conference discussions, serves as the key motivation for the present study.

To build on this point, not all developmental changes should be interpreted in the same way. At the initial time point (Week 4), some learners produced exaggerated durations as mentioned earlier, which can be interpreted as a case of *hypercorrection* (Labov, 1966b). That is, students over-applied a perceived target-like property due to explicit instructions to focus their attention on the model speaker's articulatory gestures and attempt a more target-like form (in this case, producing longer-than-necessary trill durations). It is crucial to note that some learners later reduced these durations over time. This reduction should not be viewed as regression. Rather, it could be considered a process of adjustment toward the target. However, patterns like hypercorrection followed by a downward shift to adjust or move towards the model speaker's speech are not fully captured by the models' estimates alone and may be mischaracterized as "no change" when trajectories are averaged across learners. For this reason, learner-level visual descriptives are essential for showing what is really going on in the data. This point becomes clearer in Section 7.2, where the discussion turns to learner-level plots.

These plots are not meant to replace the HLM results. Instead, they make the learner-specific trajectories easier to see. This is particularly true when change is heterogeneous or non-linear, since the average fixed-effect trend of Time and Time x Training can mask meaningful individual variation. This pattern is clear in the spaghetti plots in Section 7.2. Some participants increased in duration, others decreased, and others remained relatively stable. When aggregated at the group level, these different trajectories can yield an overall effect that appears minimal. Nonetheless, meaningful changes may still be occurring at the individual level.

Given the heterogeneous and non-linear nature of the developmental trajectories observed in this study, the organization of the semester long format is consistent with work emphasizing the benefits of distributing phonetic training over time (see the phonetic training chapter for a detailed review). For example, Saito and Chen (2025) report advantages for spaced over massed training schedules in L2 phonetic learning. Although their study examines perceptual learning of an English vowel contrast, the broader point supports the motivation for embedding imitation practice across the semester rather than treating it as a single-session lab task.

While the focus in this study is on the Spanish trill, the broader goal is to develop an approach that can be extended to other phonological targets (e.g., the tap, stops) and languages in future work. Rather than treating pronunciation practice as a single session lab activity, this study embedded the audio-visual imitation task into weekly homework across the semester and examined learner change across four time points using acoustic measures (primarily duration). The sections that follow interpret these results by focusing on longitudinal patterns of change and individual-specific trajectories. These sections consider how the imitation-based training may have influenced the outcomes for some learners but not others. In section 7.3, the training implementation is further discussed to clarify what counts as “training” in this classroom-embedded design. Before turning to training in 7.3, Section 7.2 first looks at the learner-level patterns (spaghetti plots and boxplots) to illustrate the underlying patterns of change across time points.

7.2 Development Over Time and Individual Variability

To better understand what is driving the underlying development over time and what is really going on behind the individual variability, this section looks at learner-level descriptives (including spaghetti plots and boxplots) to supplement the HLM results. The goal is to make patterns more transparent than fixed-effect averages alone.

The first question addressed whether trill duration changed over time within individual learners, and the second examined the extent to which learners differ in their rate of development. Overall, the results showed no statistically significant effect of time on trill duration across learners. In the sections that follow, the discussion looks at distributional and trajectory-based plots to clarify how individual patterns of change can be masked in the average fixed-effect trend (i.e., the overall pattern across learners). In other words, it demonstrates how learners can change in different directions, at different rates, and at different points in time, even when the overall average effect looks minimal.

This interpretation is supported by the distributional patterns observed across time points. As illustrated in Figure 7.1 (boxplot), the spread of values varies across assessments. This dispersion is also expected given the inherent variability of trill production: even within a single speaker, repeated productions can differ across tokens and contexts (e.g., Bradley & Willis, 2002). Moreover, there are shifts in both central tendency and dispersion. Although these shifts are not statistically significant differences, they do suggest change for at least a subset of learners. Hence, learners are not static in their production. Rather, they appear to adjust their productions of the target sound in different ways: upward and sometimes downward shifts over time.

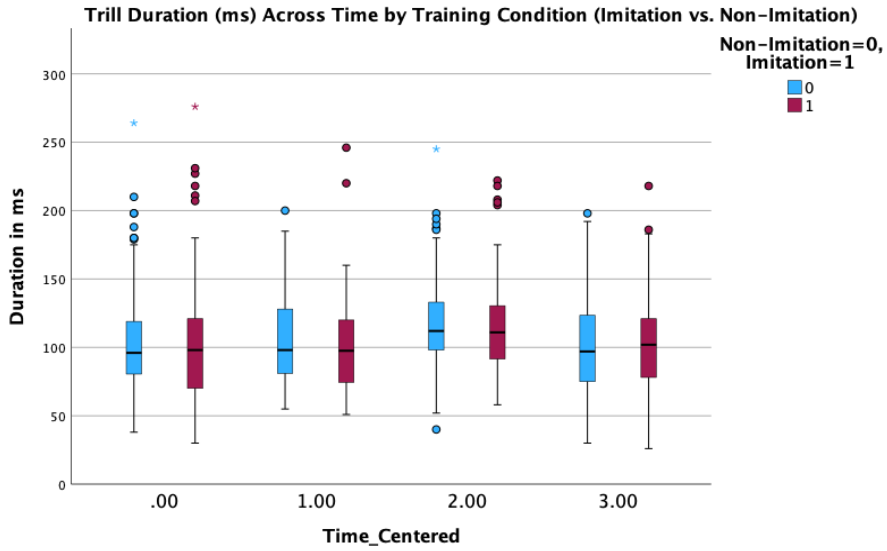
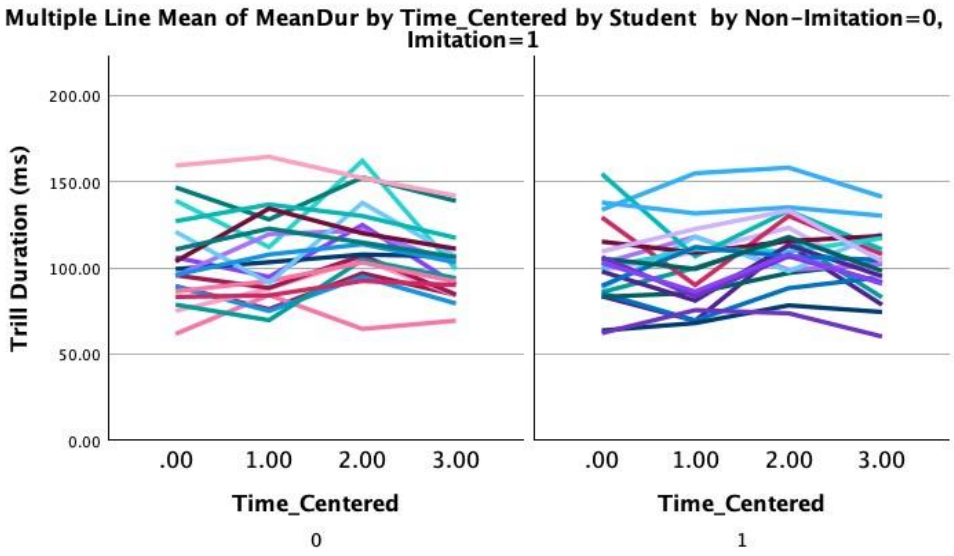


Figure 7.1. *Trill duration (ms) across time by training condition (imitation vs. non-imitation)*

To make these individual patterns more transparent, Figures 7.2 and 7.3 present spaghetti plots of trill duration across the four assessment points. Each line represents a single learner, allowing changes in direction and magnitude to be examined rather than masked by averaging. The plots show considerable variability: some learners display gradual increases in duration, others show decreases or fluctuations, and many do not follow a single linear pattern.



Figures 7.2 & 7.3. *Individual trajectories of trill duration (ms) across time (Centered at Week 4) for the non-imitation (0) and imitation (1) groups.*

When examined at the individual level, even small changes in trill duration may be meaningful. Given the articulatory complexity of the trill and the relatively short instructional period, one could expect gradual change. Therefore, incremental shifts can be interpreted as part of a broader developmental process, which can include ongoing adjustments, fluctuations, and periods of stabilization over time (e.g., Rose, 2010). Overall, development may be better understood as ongoing adjustment rather than a shift toward a fixed target, such as an idealized or prescriptive Spanish trill.

The use of multiple assessment points was also central to this design. By evaluating learners at four different time points, spaced every four weeks, it was possible to capture developmental change over time that is often missed in simple pre- and post-designs (e.g., Raudenbush & Bryk, 2002). This is particularly important given that learners do not develop at the same rate. Some show early gains, while others

require more time to adjust. As a result, examining individual trajectories helps unveil developmental patterns that are often masked when relying solely on group-level averages.

It is also worth noting that variability was observed at the early stages of the study. At baseline (Week 4), learners showed a wide range of trill durations. Some appeared to approximate, or be longer than, the model speaker's trill duration from the outset, while others showed shorter or more variable durations. To further elucidate this variability, an additional conditional growth model including heritage speaker status suggested that heritage speakers may have been somewhat closer to the model speaker at baseline than non-heritage learners, although this effect did not reach the conventional statistical significance threshold of $p < .05$. However, heritage speaker status did not significantly predict developmental trajectories over time. These findings are consistent with the descriptive patterns observed in the learner-level plots, where some learners appeared closer to the reference value at baseline while still exhibiting variable and non-linear trajectories across assessment points. Over time, some learners appeared to exhibit greater consistency in their productions, although developmental trajectories remained heterogeneous across individuals.

To further illustrate these trajectory patterns, the spaghetti plot in Figure 7.4 presents learners' distance from the model speaker across the same four assessment points. Learners are grouped according to whether they reduce their distance from the model speaker over time. While some learners move closer to the model speaker over time, others show fluctuation or minimal change. This pattern further confirms that development is not uniform across individuals.

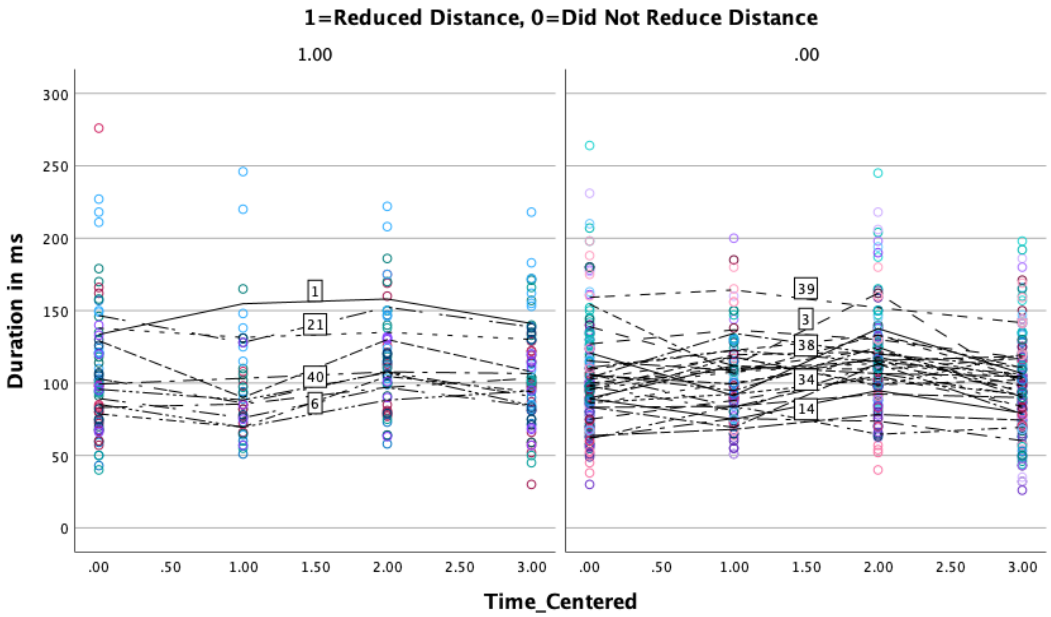


Figure 7.4. *Individual learners' trill duration trajectories across four assessment points*

Note. Learners are grouped according to whether they reduce their distance from the model speaker over time. Each line represents one learner.

In sum, the longitudinal patterns in the Spanish trill duration show that development is variable across learners and often non-linear (e.g., different learners showing increases, decreases, or stabilization at different points in time), emphasizing the value of multiple time points and individual-trajectory visualizations alongside overall group trends.

7.3 Training and Imitation

Expanding on the overview presented in Section 7.1, the third question addressed whether the training condition influenced the rate of change over time. Specifically, it asked whether learners in the imitation group showed greater development than those

in the non-training group. Overall, the results showed no statistically significant difference between groups. This section therefore clarifies what counts as “training” in this classroom-embedded design and considers how that implementation may account for the null group effect, that is, the lack of statistically significant differences between the training and non-training groups.

In the training condition, learners completed a weekly audio-visual homework task in which they listened to and watched a model speaker produce the target trill and then recorded their own productions. They were explicitly instructed to imitate the model as closely as possible, focusing on articulatory gestures (which may include features such as timing and overall quality of the trill) across repeated attempts over the semester. By contrast, learners in the non-training condition completed the same course assessments but did not participate in this imitation-based practice as part of the homework.

Moreover, this type of practice was not designed to be a one-time lab task. The imitation activity was repeated throughout the semester and integrated into weekly homework, making the audio-visual imitation task the only systematic and repeated intervention specifically targeting trill production. Learner productions were collected at multiple time points (Weeks 4, 8, 12, and 16). Because recordings were completed as part of a structured homework context within the course (including distractor items alongside target tokens), the task aimed to capture development in a classroom-based setting while still eliciting productions that could be compared across learners. Students’ trill development was then tracked longitudinally using acoustic measures. Duration was used to evaluate how learners adjusted trill duration over time. Although

students practiced oral skills and vocabulary as part of their regular coursework, these activities did not systematically target the trill production.

Since the weekly homework was integrated with chapter topics, the lexical items varied across the semester, and the 10 target tokens used for analysis were distributed across four assessment points (with distractors included). This design aligned the training materials with the course content, but it may also have introduced variability in performance and reduced the likelihood of a uniform, group-level training effect. It is also important to note that although all students practiced the target vocabulary, only the training group engaged in explicit imitation. In the present study, repetition and imitation did not function in the same way. While repetition involved general practice, as is typically done in foreign language courses, imitation required a more focused effort to match the articulatory properties of the model speaker. In addition, the training was completed outside of a laboratory setting. Although learners were instructed to work in a quiet environment and use headsets to minimize noise, recording conditions at home were not fully controlled. This may have introduced additional variability in the acoustic signal and reduced the likelihood of detecting uniform between-group effects.

At first glance, this may suggest that the imitation-based training did not produce measurable gains. However, such an interpretation would be incomplete, as it would overlook important patterns in the data. A closer examination of within learner (Level-1) trajectories reveals a clearer picture.

Although statistically significant group-level effects were not observed, visual inspection suggests that some learners did exhibit changes in trill duration over time.

As shown in Figure 7.1 (boxplot), the distribution of values shifts across time points. Moreover, Figures 7.2 and 7.3 (spaghetti plots) show that some learners adjusted their productions over time. However, these changes were not uniform across the group. Hence, the effects of training appear to be present, but unevenly distributed.

One explanation for this pattern is the variability observed at baseline. Some learners began the study with trill durations close to the reference value (e.g., model speaker's duration). Therefore, they had limited room for improvement. Others began further from the target and showed greater fluctuation. As a result, the magnitude and direction of change varied across learners. Additionally, some learners produced durations above the reference value (possible hypercorrection) and later adjusted downward. Hence, negative values do not necessarily reflect poorer performance, but rather movement relative to the reference.

This interpretation is further supported by the Full Growth Model Including Baseline Duration (Model 4B) discussed in Chapter 6. The results indicated that learners' starting points significantly influenced developmental trajectories across the semester. Learners with lower initial trill duration values exhibited greater increases over time, whereas learners who were already closer to the target model tended to show more limited changes in their trajectories. These findings further suggest that the effects of imitation training are not uniform across learners but may depend on learners' initial production levels and distance from the target model.

Regarding imitation, the findings suggest a more gradual and individualized process. Although learners were instructed to imitate a model, their ability to do so likely varied.

In addition, engagement with the imitation task appeared to strengthen over time. This was observed by the twelfth week, where it appeared that many students showed signs of refining their productions. This pattern is consistent with the possibility that imitation benefits from task familiarity and continuous practice before it becomes a reliably effective learning mechanism. Factors such as perception, motor control, and prior experience may have influenced outcomes. Therefore, imitation may function less as an immediate mechanism and more as a skill that develops over time (e.g., Saito & Chen, 2025). This further supports the view that imitation is not necessarily an automatic ability, but one that may develop with continuous practice over time. Adult learners, in particular, may require this type of practice to hone this imitation skill.

Overall, the absence of statistically significant findings does not indicate that the training was ineffective. Rather, it suggests that any benefits may be conditional and may not be fully captured by group-level models when learners respond to the training in different ways. Some learners appear to adjust their productions over time, while others do not show measurable change within the timeframe of the study. Hence, phonological development should be viewed as gradual and individualized.

In sum, the training results suggest that imitation may serve as a useful training tool but not a uniform component of development. It highlights the need to interpret outcomes through individual trajectories as well as group-level estimates.

7.4 Nature of the Trill and Rhotic Variability

I interpret the findings of this study in light of the nature of the target sound. As discussed in Chapter 4, Spanish rhotics exhibit considerable variability across speakers and contexts (e.g., Bradley & Willis, 2012). Hence, trill production does not

reflect a single, invariant form. Rather, it includes a range of realizations that differ in duration and articulation.

Because of this variability, development cannot be defined as convergence toward a single canonical target. Instead, it may involve movement within a range of acceptable forms. Therefore, fluctuations in production do not necessarily indicate instability. Rather, they reflect engagement with a complex phonetic category. In the present study, this is reflected in the use of a reference value based on the model speaker and heritage speakers.

Moreover, prior research shows that duration plays a central role in maintaining the tap–trill contrast (Bradley & Willis, 2012). Even when articulatory features vary, duration remains a key cue. Hence, changes in duration, even small ones, may be meaningful, and can be interpreted as evidence of development toward the reference value. Overall, the use of duration as the primary measure is justified and consistent with the phonetic properties of the trill.

7.5 Longitudinal Contribution

A key contribution of the present study lies in its longitudinal design and the use of HLM to examine individual trajectories over time. This approach remains less common than cross-sectional or pre- and post-test designs in L2 phonological research. Such designs provide only snapshots of performance, but do not capture change over time.

By contrast, this study follows the same learners across multiple time points.

Hence, it reveals patterns of fluctuation, stability, and gradual change. The results show that development varies both across and within learners. Moreover, trajectories are often non-linear.

Overall, these findings support a more flexible view of how learners develop their L2 phonology. Rather than progressing in clear-cut stages, learners appear to reorganize their productions gradually over time, often in variable and non-linear ways.

7.6 Pedagogical Implications

One of the goals of this study was to develop a training approach that could be realistically implemented in a foreign language classroom. I suggest that by embedding imitation tasks into regular homework, pronunciation practice becomes part of the normal learning routine rather than an isolated activity. This makes the approach more sustainable and accessible for both instructors and students.

The findings also have implications for pronunciation instruction. Overall, the results suggest that a single instructional approach may not yield uniform outcomes. Therefore, instructors may need to incorporate varied forms of input and practice.

Additionally, expectations that students will improve rapidly need to be reconsidered. Given the complexity of the trill, development is likely to be gradual. Hence, instruction should emphasize continuous practice and attention to phonetic detail. Moreover, recognizing that small improvements are meaningful may help both instructors and learners adopt a more realistic perspective. Rather than focusing solely on accuracy, instruction may benefit from emphasizing development over time.

In addition, future work will expand this longitudinal, imitation-based approach to other phonological features beyond the trill, such as the tap and stops, and to other languages. Finally, the use of visual representations, such as trajectory plots, plays an important role in making individual progress visible. These representations help highlight that learners develop in different ways, even when group-level patterns appear minimal. Also, that even small changes should be considered meaningful rather than relying solely on statistical significance to determine whether development has occurred.

7.7 Conclusion

Overall, the findings in this study provide a clearer understanding of imitation in phonetic development and Spanish trill acquisition in adult learners. I interpret the absence of statistically significant findings in some of the models not as indicating that imitation is absent in adult learners, but instead as suggesting that learners appear to use imitation in different ways and to different degrees across time. The HLM analyses showed variability in individual trajectories and patterns of change that may not be fully captured through conventional pre- and post-test comparisons alone. Hence, phonetic development should be viewed as dynamic, variable, and influenced by multiple, perhaps interacting, factors across learners and across assessment points.

Moreover, these findings suggest that pronunciation development may depend on other factors that were not examined in the present study. These factors may include previous exposure to Spanish, study schedules, motivation, language use in the home, or interactions with parents and caregivers during childhood. Future

research should continue to explore these variables longitudinally and aat the individual level.

Beginning with the pilot study and preliminary stages of this research, many of the audio-visual training tasks were developed based on personal classroom observations and feedback from colleagues regarding the ways pronunciation tends to be approached in foreign language courses. In many classrooms, greater emphasis is often placed on grammar, vocabulary, reading, and writing; meanwhile, oral communication and pronunciation may receive less attention. This may be especially true for the Spanish trill, since many students perceive it as one of the most difficult sounds to produce and often want explicit guidance on how to approach it. Students may repeat words and phrases during classroom activities; however, repetition alone does not necessarily mean that they are paying close attention to how sounds are being articulated. In many cases, it seems that learners appear to rely primarily on the auditory component instead of attending to the articulatory gestures involved in production unless they are explicitly directed to do so. So, these observations suggest that when students focused on mouth movements, lips and tongue positioning, and imitation of the speaker, their productions often reached another level of awareness and control.

Ultimately, this study aimed to highlight the significance of focused attention, imitation, and pronunciation awareness in adult language acquisition. This is not because having an accent is wrong or unacceptable. An accent is part of a speaker's

linguistic identity. Instead, being cognizant of phonological features may aid in communication, comprehensibility, and confidence during oral interaction.

Bibliography

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *The psychology of learning and motivation,* Vol. 2 K. W. Spence & J.T. Spence (eds.). London: Academic Press.

Au, T. K.-F., & Oh, J. S. (2003). Late second language learners: What predicts good outcomes? The 4th International Symposium on Bilingualism. Arizona State University. Presenters.

Au, T. K., Knightly, L. M., Jun, S.-A., & Oh, J. S. (2002). Overhearing a language during childhood. *Psychological Science*, 13, 238-243.

Babel, M. (2010). Dialect convergence and divergence in New Zealand English. *Language in Society*, 39 (4), 437-456.

Barry, W. (1989). Perception and production of English vowels by German learners: instrumental – phonetic support in language teaching. *Phonetica,* 46, 155-168.

Best, C.T., & Tyler, M.D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M.J. Munro and O, -S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production,* 13-34. Amsterdam: John Benjamins.

Best, C. T., McRoberts, G. W., & Goodel, E. (2001). Discrimination of nonnative consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, 109, 775-794.

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language,* 44, 235-249.

Birdsong, D. (Ed.). (1999). Second language acquisition and the critical period hypothesis. Mahwah: Lawrence Erlbaum Associates.

Blandin, Y., & L. Proteau. (2000). On the cognitive basis of observation learning: Development of mechanisms for the detection and correction of errors. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 53(3), 846-867.

Boersma, P. & Weenick, D. (2014). Praat: doing phonetics by computer. Version 5.3.76: Retrieved from http://www.praat.org/.

Bosch, L. (1983). El desarrollo fonológico infantil: Una prueba para su evaluación. *Anuario de sicología*, 28, 87-114.

Bourhis, R. Y. and Giles, H. (1977). The language of intergroup distinctiveness. In H. Giles (Ed.), *Language, ethnicity, and intergroup relations,* 119 -136. London: Academic Press.

Bradley, T.G. & E.W.Willis. (2012). Rhotic variation and contrast in Veracruz Mexican Spanish. *Estudios de Fonética Experimental,* 21, 43-74.

Branigan, H. P., Pickering M. J., & Cleland, A. A. (2000). Syntactic coordination in dialogue. *Cognition, 75,* B13-B15.

Broadbent, D. E. (2013). *Perception and Communication*. Oxford: Pergamon Press, 1958. *Internet Archive: American Libraries*. Web. Aug. 2013.

Brown, C. (2000). The interrelation between speech perception and phonological acquisition from infant to adult. In J. Archibald (Ed.), *Second language acquisition and linguistic theory*, 4-63. Malden, MA: Blackwell.

Brown, H. D. (2001). *Teaching by principles* (2nd ed.). White Plains, NY: Pearson.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: The MIT Press.

Curtiss, S. (1977). *Genie: A psycholinguistic study of a modern-day 'Wild Child'. Perspectives in Neurolinguistics and Psycholinguistics.* New York: Academic Press.

Curtiss, S. (1985). The development of human cerebral lateralization. In D. Frank Benson and E. Zaidel (Eds.) The Dual Brain. *The Dual Brain: Hemispheric Specialization in Humans, UCLA Forum in Medical Sciences,* 26, 97-116.

DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition,* 22, 499-533.

Delvaux, V., & Soquet, A. (2007). The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica, 64,* 145-173.

Díaz-Campos, M., & Morgan, T. A. (2002). On the production and perception of Spanish palatal obstruents: An acoustic study with implications for phonology, dialectology, and pedagogy. In J. F. Lee, K. L. Geeslin, & J. C. Clements (Eds.), Structure, meaning, and acquisition in Spanish: Papers from the 4th Hispanic Linguistics Symposium (pp. 244–268). Somerville, MA: Cascadilla.

Elliot, R. A. (1995). Foreign language phonology: Field independence, attitude, and the success of formal instruction in Spanish pronunciation. *The Modern Language Journal,* 79(iv), 530-542.

Ellis, N. C., & Beaton, A. (1993). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. Quarterly Journal of Experimental Psychology, 46A, 533-558.

Fadiga, L., et al. (1995). Motor Facilitation during Action Observation: A Magnetic

Stimulation Study. *Journal of Neurophysiology,* 73, 2608-2611.

Flege, J. E. (1987). The instrumental study of L2 speech production: Some methodological considerations. *Language Learning*, 37(2), 285-296.

Flege, J. E. (1993). Production and perception of a novel, second-language phonetic contrast. *Journal of the Acoustical Society of America*, 93, 1589-1608.

Flege, J. E., Mackay, I. R. A., & Meador, D. (1999). Native italian speaker's production and perception of English vowels. *Journal of the Acoustical Society of America*, 106 (5), 2973-2987.

Flege, J. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed) Timonium, MD: York Press. 229 - 273.

Flege, J.E. (2009). Give Input a Chance! In T. Piske and Young-Scholten, M. (Eds) *Input matters in SLA.* Bristol: *Multilingual Matters,* 175-190.

Flege, J. E. (1987). A critical period for learning to pronounce foreign languages? *Applied Linguistics,* 8, 162-177.

Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, 23, 527-552.

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustic Society of America,* 89 (1), 395-411.

Flege, J. E. (1992a). The intelligibility of English vowels spoken by British and Dutch talkers. In R. D. Kent (Ed.), *Intelligibility in speech disorders: Theory, measurement, and management,* 157-232. Philadelphia: Benjamins.

Flege, J. E. (1992b). Speech learning in a second language. In C. A. Ferguson & L. Menn & C. Stoel-Gammon (Eds.), *Phonological development: models, research, implications,* 565-604. Timonium, MD: York.

Genesee, F. (1978). Is there an optimal age for starting second language instruction? *McGill Journal of Education*, 13(2), 145-154.

Giles, H. (1973). Accent mobility: A model and some data. *Anthropological Linguistics,* 15, 181-215.

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: communication, context and consequences, in *Contexts of Accommodation: Development in Applied Sociolinguistics*. Giles, J. Coupland, and N. Coupland (Eds). 1-68. Cambridge, Cambridge University Press.

Glisky, E. L. (2007). Changes in cognitive function in human aging. In D. R. Riddle (ed.), 3-20. Boca Raton: CRC Press. Print. Brain Aging: Models, Methods, and Mechanisms.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105, 251-259.

Goldstein, H. (1995). *Multilevel Statistical Models.* London, Edward Arnold; New York, Halstead Press.

González-Bueno, M., & Quintana-Lara, M. (2011). Students' awareness of Spanish spirantization allophonic rule. In J. Levis and K. LeVelle (Eds.), *Proceedings of the 2nd Annual Pronunciation in Second Language Learning and Teaching Conference, Sept. 2010,* 180-198.

Hatch, E.M. (1983). Psycholinguistics: A second language perspective. Rowley, MA:

Newbury House.

Hauert, C. A. (1986). The relationship between motor function and cognition in the developmental perspective. *Italian Journal of Neurological Science*, 5, 101-107. Haven, CT: Yale University Press.

Heddeker, D., & Gibbons, R. D. (2006). *Longitudinal data analysis.* New York: Wiley.

Heimann, M., Ullstadius, E., & Swerlander, A. (1998). Imitation in eight young infants with downs-syndrome. *Pediatric Research,* 44, 780-784.

Herschensohn, J. (2007). *Language development and age.* New York: Cambridge University Press.

Heyes, C. M., & Foster, C. L. (2002). Motor learning by observation: Evidence from a serial reaction time task. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 55 (2), 593-607.

Iacoboni, M. (2008). The role of premotor cortex in speech perception: evidence from fMRI and rTMS. *Journal of Physiology*, 102(1-3), 31-34.

Jarvelainen, J., et al. (2001). Stronger reactivity of the human primary motor cortex during observation of live rather than video motor acts. *Neuroreport*, 12, 3493-3495.

Johnson, J.S., and Newport, E.L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology,* 21(1), 60-99.

Kappes, J., Baumgaertner, A., Peschke, C., & Ziegler, W. (2009). Unintended imitation in nonword repetition. *Brain Language, 111,* 140-151.

Kim, J.-Y., & Repiso-Puigdelliura, G. (2020). Deconstructing heritage language dominance: Effects of proficiency, use, and input on heritage speakers' production of the Spanish alveolar tap. Phonetica, 77, 55-80.

Klein, W. (1986). *Second language acquisition*. Cambridge: Cambridge University Press.

Klempin, F., & Kempermann, G. (2007). Adult hippocampal neurogenesis and aging. *PubMed - indexed for MEDLINE*. N.p., Aug. 2007. Web. 4 Mar. 2015, 257(5), 271-80.

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *National Review Neuroscience, 5,* 831- 843.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277,* 684-686.

Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: vocal imitation and developmental change. J. Acoust. Soc. Am., 100, 2425- 2438

Kumaravadivelu, B. (2003). Beyond methods: Macro strategies for language teaching. New Haven, CT: Yale University Press.

Kumaravadivelu, B. (2005). Understanding language teaching: From method to post method. Mahwah, NJ: Lawrence Erlbaum.

Ladefoged, P., & Maddieson, I. (1996). The sounds of the world's languages. Oxford: Blackwell.

Laird, N. M., & Ware, J. H. (1982). Random effects models for longitudinal data. *Biometrics,* 38, 963-974.

Lavoie, L. (2001). Consonant strength: phonological patterns and phonetic manifestations. New York: Garland.

Lenneberg, E. H. (1967). Biological foundations of language. New York: Wiley & Son.

Levis, J.M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly,* 39(3), 369-377.

Lindau, M. (1985). The story of /r/; in Fromkin, Phonetic linguistics: essays in honor of P. Ladefoged, 157.167. Orlando: Academic Press.

Long, M. H. (1980). Inside the black box: Methodological issues in classroom research on language learning. *Language Learning*, 30(1), 1-42. Print. Major,

Major, R. C., & Kim, E. (1996). The similarity differential rate hypothesis. *Language Learning,* 46, 465-496.

Major, Roy C. (2001). Foreign accent: The ontogeny and phylogeny of second language phonology. Mahwah: Routledge.

McDonough, J. and Johnson, K. (1997). Tamil liquids: an investigation into the basis of the contrast among five liquids in a dialect of Tamil. Journal of the international Phonetics Association, 27, 1-26.

Meléndez-Ballesteros, N. (2014). Why teaching pronunciation to Spanish L2 learners matters. *Voices,* 2(1). ucla_spanport_voices_22797. Retrieved from: https://escholarship.org/uc/item/51q604qx

Meltzoff, A. N. & Moore, M. K. (1999). A new foundation for cognitive development in infancy: The birth of the representational infant. In E. K. Scholnick, K. Nelson, S. Gelman, & P. H. Miller (Eds.), *Conceptual development: Piaget's legacy*, 53-78. Mahwah, NJ: Erlbaum Press.

Meltzoff, A. N. (1999). Origins of theory of mind, cognition and communication. *Journal of Communication Disorders,* 32, 251-269.

Mirochnic, S., Wolf, S., & Staufenbiel, M. (2009). Age effects on the regulation of adult hippocampal neurogenesis by physical activity and environmental enrichment in the APP23 mouse model of Alzheimer disease. *Hippocampus,* 19, 1008- 1018.

Morgan, J.L. (2002). Word recognition and phonetic structure acquisition: Some possible relations. Paper Presented at the Meeting of Acoustical Society of America*:* Pittsburgh, PA. New York: Garland.

Nagle, C. L. (2019). A longitudinal study of voice onset time development in L2 Spanish stops. Applied Linguistics, 40(1), 86–107. https://doi.org/10.1093/applin/amx011

Nishitani, N., & Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences of the United States of America,* 97, 913-918.

Oyama, S. (1978). The sensitive period and comprehension of speech. *Working Papers on Bilingualism,* 16, 1-17.

Parodi, C. (2010). El otro México: español chicano, koineización y diglosia en Los Ángeles, California. Historia de la sociolingüística en México. El Colegio de México.

Parodi, C. (1999b). Koineización e historia: la sincronía, ventana de la diacronía. Boletín de Filología (Chile), 37, 915-931.

Parodi, C. (2004). Contacto de dialectos en Los Ángeles: español chicano y español salvadoreño. Sonora.

Patkowski, M. (1980). The sensitive period for the acquisition of syntax in a second language.” *Language Learning,* 30, 449-472.

Penfield, Wilder, & Roberts, L. (1959). *Speech and Brain Mechanisms*. Princeton: Princeton UP.

Pickering, M. J. & Garrod, S. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*, 169-226.

Pinker, S. (1994). The language instinct: The new science of language and mind. New York: William Morrow and Company.

Quilis, A. (1993). *Tratado de fonología y fonética españolas*. Madrid: Gredos. Ratey, M.D., John J. (2002). *A user's guide to the brain*. New York: Vintage Books, a Division of Random House, Inc.

Raudenbush, S. (2001). Comparing personal trajectories and drawing causal inferences from longitudinal data. *Annual Review of Psychology*, 52, 501-525.

Recasens, D. (1991). On the production characteristics of apicoalveolar taps and trills. *Journal of Phonetics, 19,* 267-280.

Recasens, D. 1986: Estudis de fonètica experimental del català oriental central. Barcelona: Publicacions de l’Abadia de Montserrat.

Repiso-Puigdelliura, G., & Kim, J.-Y. (2021). The missing link in Spanish heritage trill production. Bilingualism: Language and Cognition, 24(3), 454-466.

Sato, M., Grabski, K., Garnier, M., Granjon, L., Schwartz, J., & Nguyen, N. (2013). Converging toward a common speech code: imitative and perceptuo-motor recalibration processes in speech production. *Frontiers in Psychology,* 2013, 4:600. *Epub 2013 Sep 11.*

Schwartz, M.D., Jeffrey M., & Begley, S. (2003). *The Mind & the Brain*. New York: Harper Perennial.

Scovel, T. (2000). A critical review of the critical period research. *Annual Review of Applied Linguistics,* 20, 213-223.

Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech.* Rowley: Newbury House.

Shaw, D. J., & Czekóová, K. (2013). Exploring the development of the mirror neuron system: Finding the right paradigm. *Developmental Neuropsychology,* 38(4), 256-271.

Singer, J. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics,* 24(4), 323-355.

Snow, C.E. (1983). Age differences in second language acquisition: Research findings and folk psychology. In K.M. Bailey, M.H. Long and S. Peck (eds.), *Second Language Acquisition Studies.* Rowley, MA: Newbury House Publishers, 141-150.

Solé, M. -J. (2002). Aerodynamic characteristics of trills and phonological patterning. *Journal of Phonetics,* 30, 655-688.

Tees R.C., Werker, J.F. (1984). Perceptual Flexibility: Maintenance or Recovery of the Ability to Discriminate Non-native Speech Sounds. *Canadian Journal of Psychology*, 38, 579-590.

Trubetzkoy, N. S. (1939/1969). Principles of phonology (C. A. M. Baltaxe, Trans.). Berkeley: University of California Press. (Original work published as Grundzüge der Phonologie, Travaux du Cercle Linguistique de Prague, 7).

Wohlschlager, W., & Bekkering, H. (2002). Is human imitation based on a mirror neuron system? Some behavioral evidence. *Experimental Brain Research,* 143, 335-341.

Wulf, G, C. Shea, & Park, J. H. (2001). Attention and motor performance: Preferences for and advantages of an external focus. *Research Quarterly for Exercise and Sport*, 72, 335-344.

Wulf, G. McNevin, N., Shea, C.H., & Wright, D.L. (1999). Learning phenomena: Future Challenges for the Dynamical Systems Approach to Understanding the Learning of Complex Motor Skills. *International Journal of Sport Psychology,* 30, 531-557.

Yrttiaho, Santeri, et al. (2010). Temporal integration of vowel periodicity in the auditory cortex. *Acoustical Society of America,* 128 (1), 224- 234.

Appendix A

Participant Information

Table A1

Participant Information

Participant (Student) ID	*Heritage Status*	*Training Condition (Trained vs. Non-trained)*	*Spanish Spoken by Parents*	*Least comfortable skill*	*Most comfortable skill*
001	1	1	Dad only	Speaking	Listening
002	1	0	None	Writing	Speaking
003	1	0	None	Writing	Speaking
004	1	1	None	Speaking	Listening
005	0	1	Both parents	Speaking	Listening
006	0	1	Mother only	Speaking	Listening
007	0	1	Dad only	Speaking	Listening
008	0	1	Both parents	Speaking	Listening
009	1	0	Both parents	Writing	Listening
010	1	1	None	Speaking	Reading
011	1	1	None	Speaking	Reading
012	1	1	None	Speaking	Listening
013	1	1	None	Writing	Listening
014	1	1	None	Writing	Listening

015	0	1	Both parents	Speaking	Listening
016	0	1	Both parents	Speaking	Reading
017	0	0	Both parents	Writing	Speaking
018	0	0	Both parents	Speaking	Listening
019	0	0	Both parents	Speaking	Reading
020	0	0	Dad only	Speaking	Reading
021	1	1	None	Listening	Reading
022	1	0	None	Speaking	Writing
023	0	1	Both parents	Speaking	Reading
024	0	1	Both parents	Writing	Listening
025	1	0			
026	0	1	Both parents	Writing	Listening
027	0	1	Mother only	Writing	Speaking
028	1	1	None	Writing	Listening
029	1	0	None	Speaking	Reading
030	1	0	Mother only	Writing	Reading
031	1	0	None	Listening	Speaking

032	0	0	Mother only	Listening	Reading
033	1	0	Both parents	Speaking	Listening
034	0	1	Both parents	Writing	Listening
035	1	0	None	Writing	Reading
036	0	0	Mother only	Writing	Listening
037	0	0	Both parents	Writing	Speaking
038	0	0	Both parents	Writing	Speaking
039	0	0	Both parents	Writing	Speaking
040	0	1	Both parents	Writing	Speaking

Note. Heritage status was coded as a binary variable (1 = non-heritage, 0 = heritage), and training condition was coded as 1 = trained and 0 = non-trained. Additional questionnaire-based background variables are included for descriptive purposes only and were not entered as predictors in the HLM models.

Appendix B

Additional Articulatory Illustrations of Rhotic Variation

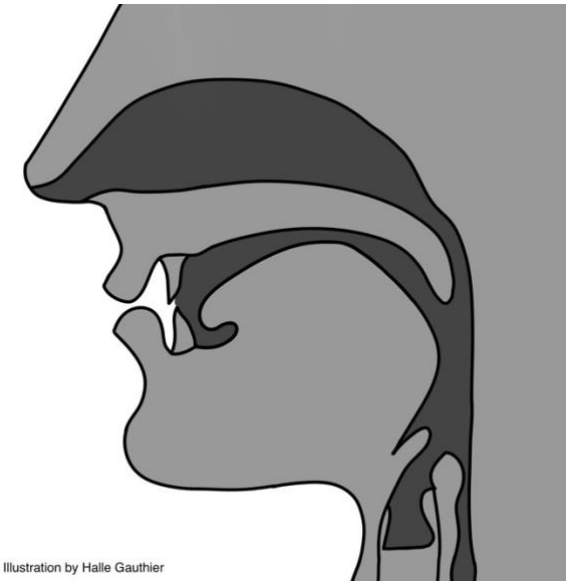


Figure B1. *Bunched rhotic realization*

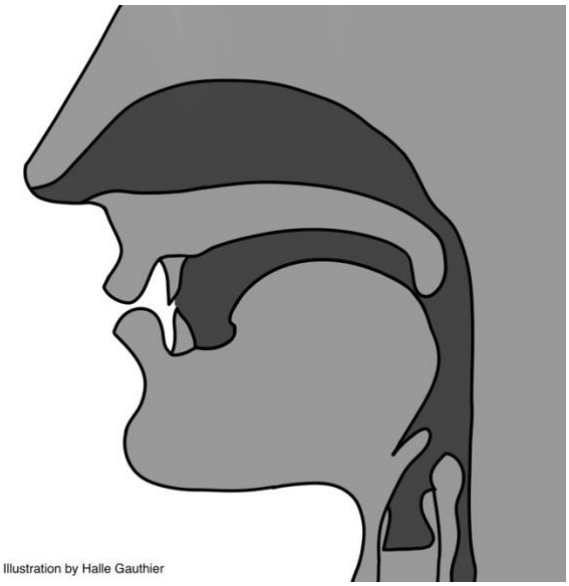


Figure B2. *Retracted rhotic realization observed in some Caribbean Spanish varieties*

Appendix C

Training Materials and Instructions

This appendix contains examples of the training materials and oral assessment tasks embedded in weekly or bi-weekly homework assignments during the preliminary phase of the study, when different activities were explored to determine which tasks elicited imitation more effectively.

Language Background Questionnaire

1. Have you ever taken Spanish formally?
2. If you have taken Spanish, for how long?
3. How well did you do?
4. Do any of your parents speak Spanish?
5. Do you speak Spanish at home?
6. As a child, did anyone ever speak Spanish to you on a regular basis?
7. With regards to the last question, did you respond in Spanish?
8. Do you ever find yourself using Spanglish- half Spanish, the other English?
9. On average, how often do you listen to Spanish-speaking radio stations or watch tv shows in Spanish?
10. Which skill are you most uncomfortable with or fearful of when learning a new language?
11. Out of the four basic learning skills, which one are you most comfortable with?
12. What has been your experience with Spanish?
13. When learning a new language, do you consider yourself...
14. When dealing with technology, such as the internet or other, you are:
15. How much outside class time will you put into learning the material?
16. What kind of learner do you consider yourself to be?
17. When learning, I tend to...
18. When learning, it helps me to ...

Figure C1. Language Background Questionnaire

Note. Students were given the questionnaire at the beginning of the semester to collect background information related to language experience and exposure.

Training Tasks Used During the Preliminary Phase of the Study

Figure C2 illustrates one of the oral homework activities used during the preliminary phase of the study. Students first listened to the instructor pronounce a target word or phrase and then repeated the item before recording their own production through the Voice Authoring applet.

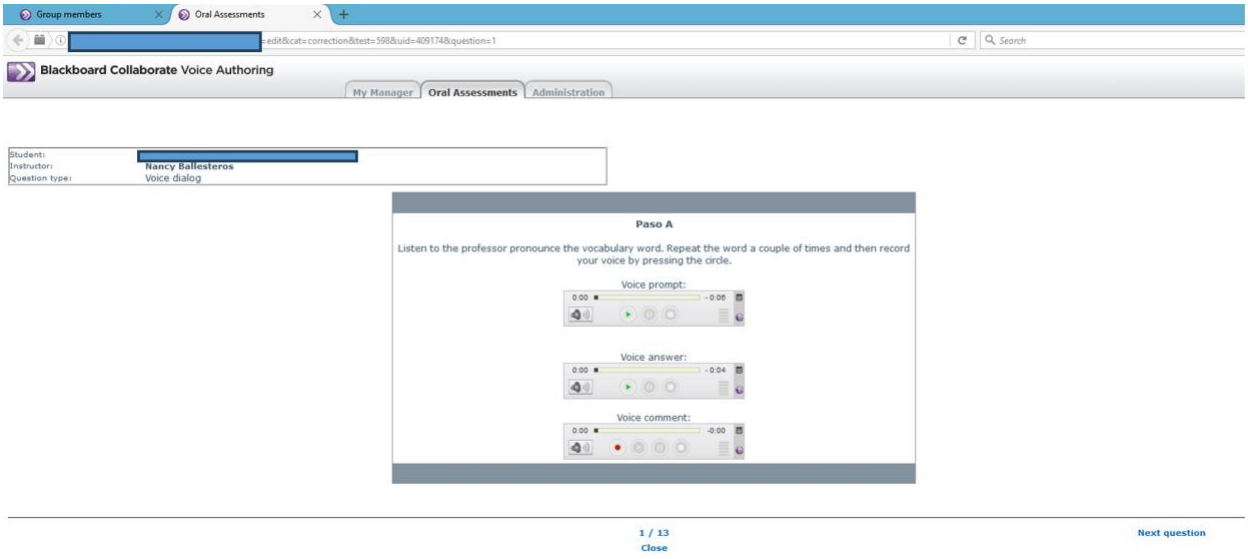
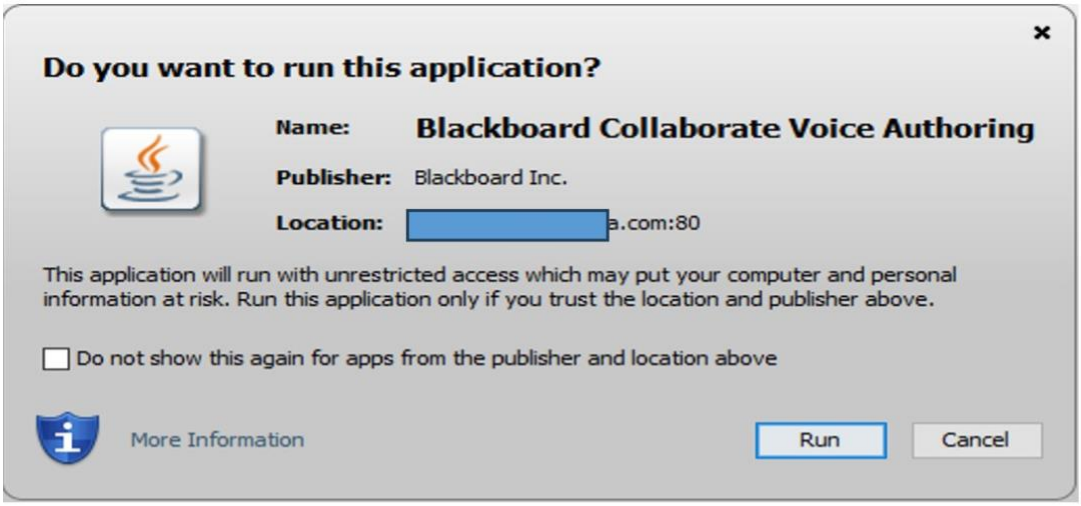


Figure C2. *Example of an oral production task in Horizon Wimba using Voice Authoring.*

Figure C3 illustrates the Blackboard Collaborate Voice Board used for pronunciation practice activities. Students listened to recorded target words and then recorded their own productions within each section of the applet.



Always click 'run' when using Horizon Wimba.

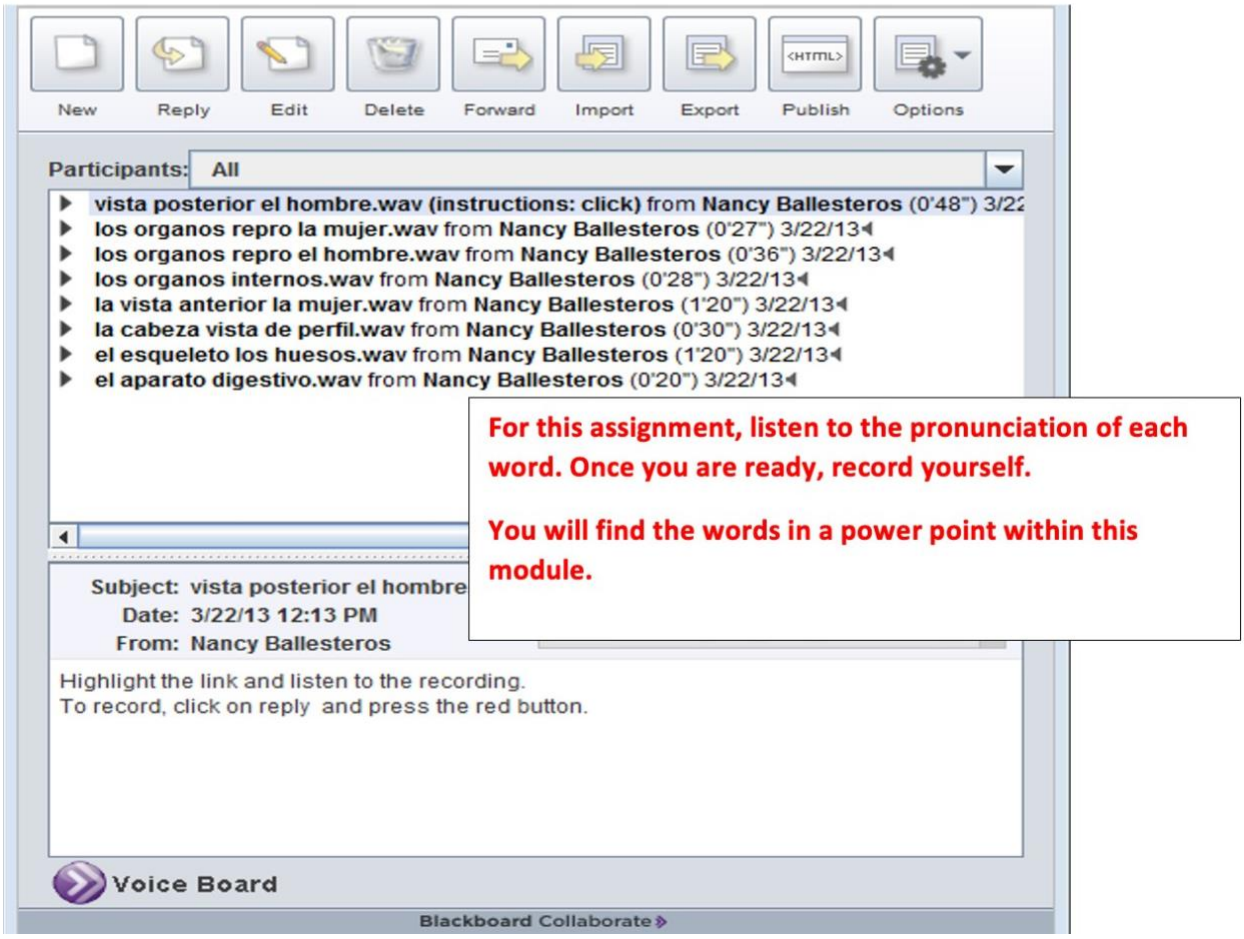


Figure C3. *Blackboard Collaborate Voice Board used for pronunciation recording activities.*

Figure C4 illustrates a listening and oral response task used. Students listened to a narrated fairy tale while following the written text on the screen and later responded orally to comprehension questions using the past tense.

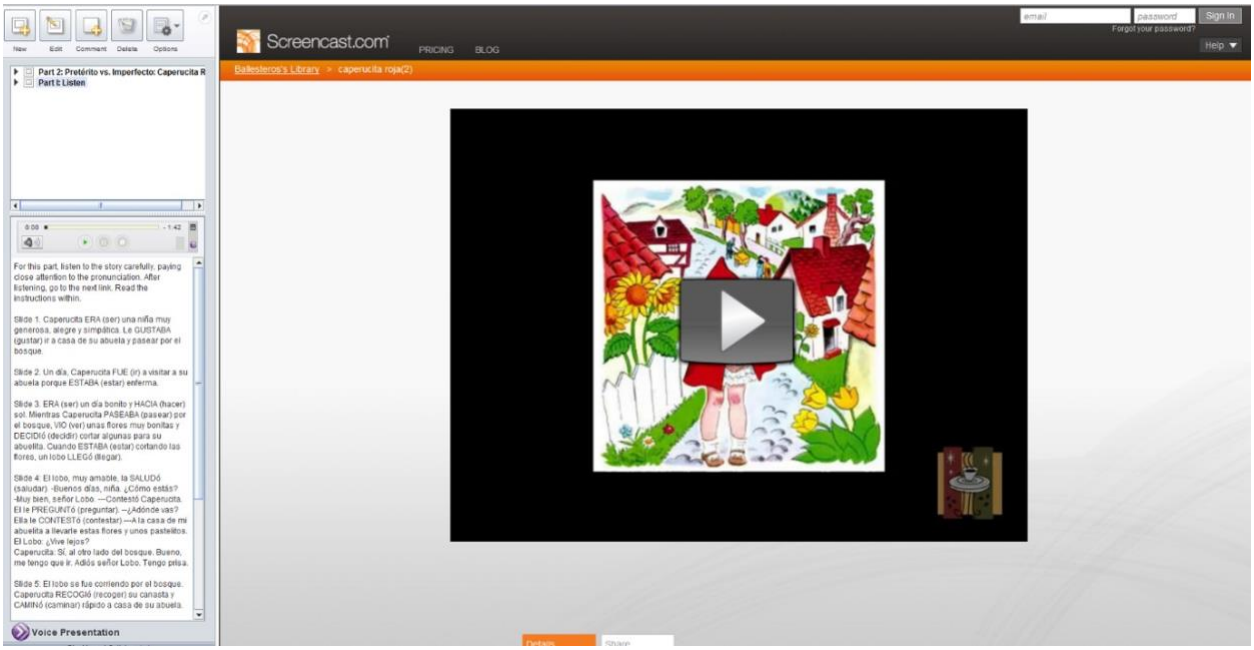


Figure C4. *Example of a listening and oral response task in Horizon Wimba using Voice Presentation.*

Figure C5 illustrates a stress perception task. Students listened to recorded words and identified the stressed syllable by selecting the corresponding orthographic option.

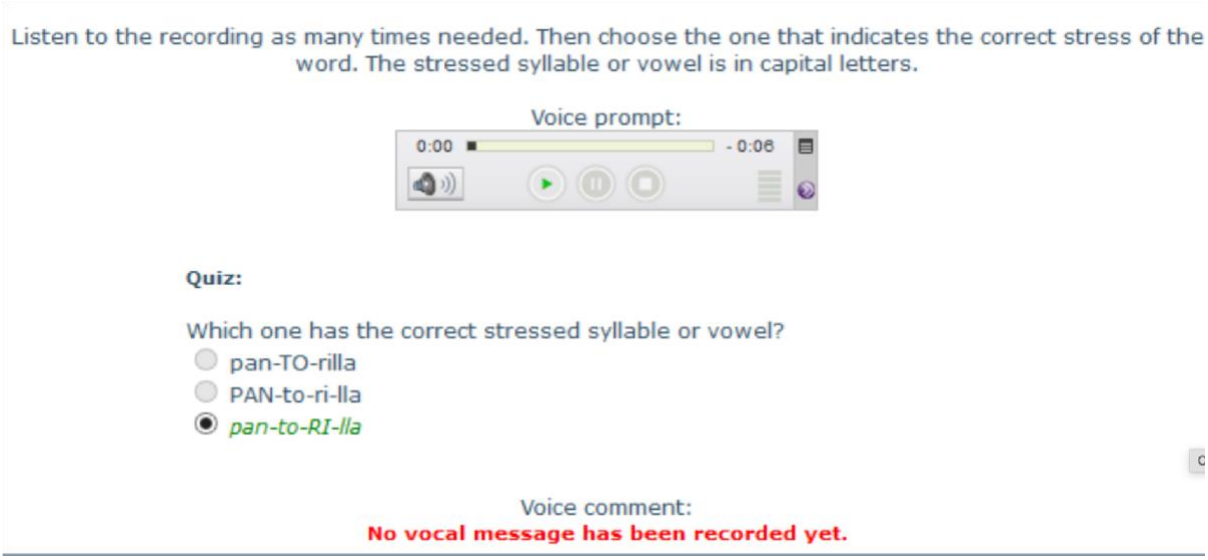


Figure C5. *Example of a stress perception activity administered through Horizon Wimba*

Figure C6 illustrates an audio-visual oral activity in Horizon Wimba Voice Board. Students were instructed to observe and listen to the speaker while recording themselves repeating the pronunciation.

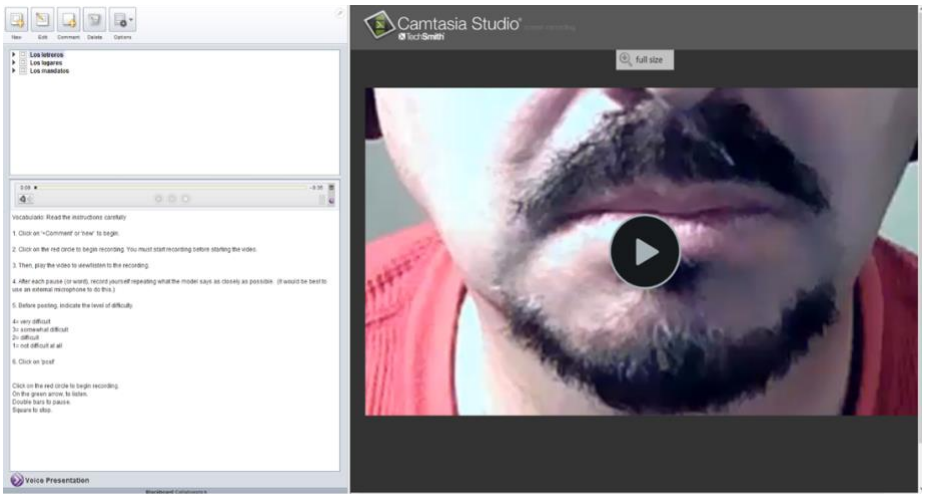


Figure C6. *Screenshot of an audio-visual oral activity in Horizon Wimba*

Appendix D

Preliminary Pilot Study Graphs and Spectrographic Representations

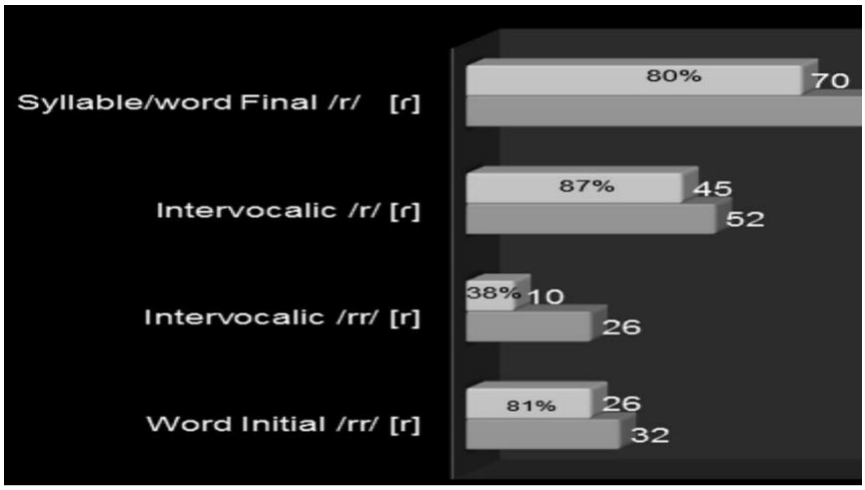


Figure D1. *Preliminary descriptive graph of rhotic productions from the pilot study*

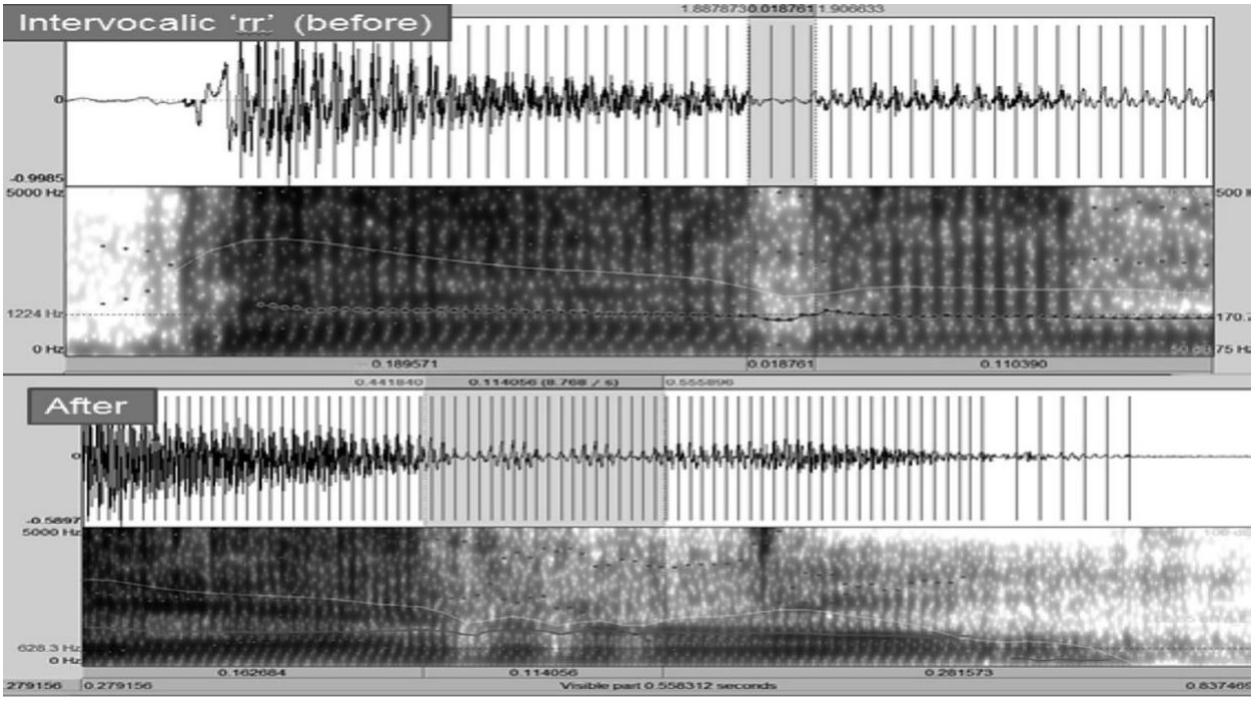


Figure D2. *Spectrographic representation of a learner's trill production before and after training*

Appendix E

Consent Form

The consent form included the following statement:

During the course of the semester(s), you were given listening and speaking exercises that focused on honing two skills that have not been emphasized enough in foreign language instruction, namely listening and speaking. It has been attested that when these are enhanced, oral communication becomes more comprehensible and effective. For this reason, I, (Instructor's Name), created listening and speaking exercises that aimed at helping you, Spanish 1 or 2 student, improve these skills. Your aural work is of utmost importance to studies dealing with the acquisition of phonology or the study of sounds in the Linguistic and Second Language Acquisition fields. It is for this reason that I, (Instructor's Name), would like to use your aural expression or speaking work as part of my research in this field. Your work may be presented at conferences or published. It will NOT reveal your identity in any way or form (your voice will not be heard), unless you consent to it. Instead of your name, there will be a random number assigned. Below you will find different ways that you can consent to the use of your work.

- I agree to the following:
    - o I grant permission for my recordings to be heard and analyzed. o I grant permission for my recordings to be heard but not analyzed. o I grant permission for my recordings to be analyzed but not heard.
    - o I DO NOT want my recordings to be used in any way or form.

Appendix F

Token List

Table F1

Target Tokens and Distribution Across Assessment Points

Word	***Assessment Points***
carrera	T1 - T4
carretera	T4
carro	T1 - T4
perro	T1 - T4
recámara	T1
rizado	T1, T3
regar	T1 - T4
rubio	T1, T3
corrida	T4
torre	T4

Appendix G

Data Structure and Variable Coding

This appendix presents a sample of the dataset used in the analysis. Each row represents a token-level observation nested within participants across time. Duration is measured in milliseconds and compared to a model speaker. Difference scores reflect deviation from the model.

The dataset was organized in long format, where each row represents one observation (i.e., a word produced) for a participant at a given time point. This organization method allows for the analysis of differences within and across learners over time using multilevel modeling. The dataset was initially organized in Excel in long format and coded to ensure compatibility with SPSS, where it was prepared for analysis and then analyzed using HLM.

Table G1

Variable Coding and Descriptions

Variable	***Description***
ID	Participant identifier
Heritage Status	Binary variable (1 = Nonheritage, 0 = Heritage)
Training Condition	Trained vs. non-trained
TIME_CEN	Time variable centered at Week 4

Word	Lexical item produced
Position	Word-initial or medial
Following Vowel	Vowel following the trill
Stress	Stress condition
Duration (ms)	Acoustic duration
Model Duration	Reference speaker duration
Duration Difference	Participant − model
Abs Diff	Absolute difference
Occlusions	Number of lingual contacts
Model Occlusions	Reference occlusions
Occlusion Diff	Difference from model

Table G2

Participant and Linguistic Context Variables

Student	***Time***	***Training***	***Token***	***Word***	***Position***	***Following Vowel***	***Stress***
001	1	1	1	carrera	2	1	1
001	1	1	2	carrera	2	1	1
001	1	1	4	carro	2	3	0
001	1	1	1	perro	2	3	0
001	1	1	1	recámara	1	1	0
002	1	0	1	carrera	2	1	1
002	1	0	2	carrera	2	1	1
002	1	0	4	carro	2	3	0
002	1	0	1	perro	2	3	0
002	1	0	1	recámara	1	1	0

Note. Heritage status was coded as a binary variable, where non-heritage speakers were coded as 1 (NON.HS = 1) and heritage speakers were coded as 0 (HS = 0). This coding was used consistently throughout the analyses. Position was coded as initial = 1 and medial = 2, and following vowel was coded as e = 1, i = 2, and o = 3). Stress was coded as a binary variable, with 1 indicating stressed position and 0 indicating nonstressed position. Lastly, since time was centered at Week 4, the time variable was recoded as follows:

Table G3

Time-Centered Coding

Time	***Time centered***
Time 1 (week 4)	*0.00*
Time 2 (week 8)	*1.00*
Time 3 (week 12)	*2.00*
Time 4 (week 16)	*3.00*

Table G4

Acoustic and Outcome Variables

	Duration (ms)	Model Duration	Duration Diff	Duration Diff Absolute	Occlusions	Model Occlusios	Occl Diff
001	67	146	-79	79	2	3	-1
001	64	146	-82	82	2	3	-1
001	119	175	-56	56	2	4	-2
001	227	170	57	57	3	4	-1
001	110	104	6	6	3	3	0
002	83	146	-63	63	2	3	-1
002	77	146	-69	69	2	3	-1
002	59	175	-116	116	2	4	-2
002	67	170	-103	103	2	4	-2
002	74	104	-30	30	0	3	-3

Appendix H

Unconditional Means Model (Model 1)

Outcome Variable: DURATION

Model Specification Level-1 Model

$DURATION_{mj} = \psi_{0j} + e_{mj}$ **Level-2 Model**

$\psi_{0j} = \gamma_{00} + u_{0j}$

Mixed Model

$DURATION_{mj} = \gamma_{00} + u_{0j} + e_{mj}$

Model Results

Variance Components

- $\sigma^2_{\varepsilon} = 0.00101$
- $\tau_{00} = 0.00035$

Reliability estimate

- $\psi_{0:}$ 0.878

Log-Likelihood

- 1.640550E+03

Table H1

Fixed Effects (standard)

Effect	Coefficient	SE	t	df	p
γ_{00} *(Intercept)*	0.104624	0.003176	32.941	.39	< .001

Table H2

Robust Standard Errors

Effect	Coefficient	SE	t	df	p
γ_{00} *(Intercept)*	0.104624	0.003136	33.363	39	< .001

Table H3

Random Effects

Effect	SD	Variance	df	χ^2	p
u_0	0.01882	0.00035	39	322.79	< .001
e	0.03177	0.00101	—	—	—

Model Fit

- Deviance = -3281.099365
- Number of parameters = 2

Appendix I

Unconditional Growth Model: No Level-2 Predictors(Model 2)

Outcome Variable: DURATION

Model Specification

Level-1 Model

$$DURATION_{mj} = \psi_{0j} + \psi_{1j}*(TIME_CEN_{mj}) + e_{mj}$$

Level-2 Model

$$\psi_{0j} = \gamma_{00} + u_{0j}$$

$$\psi_{1j} = \gamma_{10}$$

Mixed Model

$$DURATION_{mj} = \gamma_{00} + \gamma_{10}*TIME_CEN_{mj} + u_{0j} + e_{mj}$$

Model Results

Variance Components

- $\sigma^2_{\varepsilon} = 0.00101$
- $\tau_{00} = 0.00035$

Reliability estimate

- $_{\psi}0:$ 0.878

Log-Likelihood

- 1.635423E+03

Table I1

Fixed Effects (standard)

Fixed Effect	*Coefficient*	*Standard Error*	*t-ratio*	*Approx. d.f.*	*p-value*
For INTRCPT1, ψ0					
INTRCPT2, γ_{00}	0.104251	0.003469	30.056	39	<0.001
For TIME_CEN slope, ψ_1					
INTRCPT2, γ_{10}	0.000243	0.000912	0.267	790	0.790

Table I2

Robust Standard Errors

Fixed Effect	*Coefficient*	*Standard Error*	*t-ratio*	*Approx. d.f.*	*p-value*
For INTRCPT1, ψ_0					
INTRCPT2, γ_{00}	0.104251	0.003687	28.274	39	<0.001

For TIME_CEN *slope, ψ_1*					
INTRCPT2, γ_{10}	0.000243	0.000635	0.383	790	0.702

Table I3

Random Effects

Random Effect	Standard Deviation	Variance Component	*d.f.*	χ^2	*p*-value
INTRCPT1, u_0	0.01882	0.00035	39	322.33258	<0.001
level-1, *e*	0.03179	0.00101			

Model Fit

- Deviance = -3270.846075
- Number of estimated parameters = 2

Appendix J

Training Predicting Initial Status (Model 3)

Outcome: DURATION

Model Specification

Level-1 Model

$DURATION_{mj} = \psi_{0j} + \psi_{1j}*(TIME_CEN_{mj}) + e_{mj}$

Level-2 Model $\psi_{0j} = \gamma_{00} + \gamma_{01}*(TRAINING_j) + u_{0j}$ $\psi_{1j} = \gamma_{10}$

Mixed Model

$DURATION_{mj} = \gamma_{00} + \gamma_{01}*TRAINING_j + \gamma_{10}*TIME_CEN_{mj} + u_{0j} + e_{mj}$

Model Results

Variance Components

- $\sigma^2_\varepsilon = 0.00101$
- $\tau_{00} = 0.00036$

Reliability estimate

- $\psi_0 = 0.880$

Log-Likelihood

- 1.630421E+03

Table J1

Fixed Effects (standard)

Fixed Effect	Coefficient	Standard Error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, ψ0					
INTRCPT2, γ00	0.105312	0.004861	21.666	38	<0.001
TRAINING, γ01	-0.002028	0.006434	-0.315	38	0.754
For TIME_CEN slope, ψ1					
INTRCPT2, γ10	0.000243	0.000912	0.266	790	0.790

Table J2

Robust Standard Errors

Fixed Effect	Coefficient	Standard Error	t-ratio	d.f.	p-value
For INTRCPT1, ψ0					
INTRCPT2, γ00	0.105312	0.005253	20.048	38	<0.001
TRAINING, γ01	-0.002028	0.006304	-0.322	38	0.749
For TIME_CEN slope, ψ1					
INTRCPT2, γ10	0.000243	0.000635	0.382	790	0.702

Table J3

Random Effects

Random Effect	Standard Deviation	Variance Component	*d.f.*	χ^2	*p*-value
INTRCPT1, u_0	0.01907	0.00036	38	321.80274	<0.001
level-1, e	0.03179	0.00101			

Model Fit

- Deviance = -3260.842199
- Number of estimated parameters = 2

Appendix K

Full Conditional Model: Training and Predictors of Growth (Model 4A)

Outcome Variable: DURATION

Model Specification

Level-1 Model

$$DURATION_{mj} = \psi_{0j} + \psi_{1j}*(TIME_CEN_{mj}) + e_{mj}$$

Level-2 Model

$$\psi_{0j} = \gamma_{00} + \gamma_{01}*(TRAINING_j) + u_{0j} \qquad \psi_{1j} = \gamma_{10} + \gamma_{11}*(TRAINING_j)$$

Mixed Model

$$DURATION_{mj} = \gamma_{00} + \gamma_{01}*TRAINING_j + \gamma_{10}*TIME_CEN_{mj} + \gamma_{11}*TRAINING_j*TIME_CEN_{mj} + u_{0j} + e_{mj}$$

Model Results

Variance Components

- $\sigma 2_{\varepsilon} = 0.00101$ (within-learner variance)
- $\tau_{00} = 0.00036$ (between-learner intercept variance)

Log-likelihood

- 1.626016E+03

Table K1

Fixed Effects (standard)

Fixed Effect	***Coefficient***	***Standard Error***	***t-ratio***	***d.f.***	***p-value***
For INTRCPT1, ψ_0					
INTRCPT2, γ_{00}	0.105827	0.005075	20.854	38	<0.001
TRAINING, γ_{01}	-0.003018	0.007015	-0.430	38	0.669
For TIME_CEN slope, ψ_1					
INTRCPT2, γ_{10}	-0.000094	0.001317	-0.071	789	0.943
TRAINING, γ_{11}	0.000648	0.001826	0.355	789	0.723

Table K2

Robust Standard Errors

Fixed Effect	***Coefficient***	***Standard Error***	***t-ratio***	***d.f.***	***p-value***
For INTRCPT1, ψ_0					
INTRCPT2, γ_{00}	0.105827	0.005693	18.588	38	<0.001
TRAINING, γ_{01}	-0.003018	0.007415	-0.407	38	0.686
For TIME_CEN slope, ψ_1					
INTRCPT2, γ_{10}	-0.000094	0.000879	-0.107	789	0.915

TRAINING, γ_{11}	0.000648	0.001264	0.512	789	0.608

Table K3

Random Effects

Random Effect	Standard Deviation	Variance Component	*d.f.*	χ^2	*p*-value
INTRCPT1, u_0	0.01907	0.00036	38	321.55441	<0.001
level-1, *e*	0.03181	0.00101			

Model Fit

- Deviance = -3252.032418
- Number of estimated parameters = 2

Appendix L

Full Growth Model Including Baseline Duration (Model 4B)

Outcome Variable: DURATION

Model Specification

Level-1 Model

$$DURATION_{mj} = \psi_{0j} + \psi_{1j}*(TIME_CEN_{mj}) + e_{mj}$$

Level-2 Model

$$\psi_{0j} = \gamma_{00} + \gamma_{01}*(DURATION_j) + \gamma_{02}*(TRAINING_j) + u_{0j} \quad \psi 1j = \gamma 10 + \gamma 11*(DURATIONj) + \gamma 12*(TRAININGj)$$

Mixed Model

$$DURATION_{mj} = \gamma_{00} + \gamma_{01}*DURATION_j + \gamma_{02}*TRAINING_j + \gamma_{10}*TIME_CEN_{mj} + \gamma_{11}*DURATION_j*TIME_CEN_{mj} + \gamma_{12}*TRAINING_j*TIME_CEN_{mj} + u_{0j} + e_{mj}$$

Model Results

Variance Components

- $\sigma^2_\varepsilon = 0.00096$ (within-learner variance)

- τ_{00} = 0.00000 (between-learner intercept variance)

Log-likelihood

- 1.670044E+03

Table L1

Fixed Effects (standard)

Fixed Effect	*Coefficient*	*Standard Error*	*t-ratio*	*d.f.*	*p-value*
For INTRCPT1, ψ_0					
INTRCPT2, γ_{00}	-0.014625	0.009571	-1.528	37	0.135
DURATION, γ_{01}	0.001140	0.000087	13.059	37	<0.001
TRAINING, γ_{02}	-0.000797	0.003473	-0.229	37	0.820
For TIME_CEN slope, ψ_1					
INTRCPT2, γ_{10}	0.009692	0.004958	1.955	788	0.051
DURATION, γ_{11}	-0.000093	0.000045	-2.050	788	0.041
TRAINING, γ_{12}	0.000497	0.001783	0.279	788	0.780

Table L2

Robust Standard Errors

Fixed Effect	*Coefficient*	*Standard Error*	*t-ratio*	*d.f.*	*p-value*
For INTRCPT1, ψ_0					
INTRCPT2, γ_{00}	-0.014625	0.005234	-2.794	37	0.008
DURATION, γ_{01}	0.001140	0.000049	23.065	37	<0.001
TRAINING, γ_{02}	-0.000797	0.001727	-0.461	37	0.647
For TIME_CEN slope, ψ_1					
INTRCPT2, γ_{10}	0.009692	0.003470	2.793	788	0.005
DURATION, γ_{11}	-0.000093	0.000033	-2.833	788	0.005
TRAINING, γ_{12}	0.000497	0.001132	0.439	788	0.661

Table L3

Random Effects

Random Effect	*Standard Deviation*	*Variance Component*	*d.f.*	*χ^2*	*p-value*
INTRCPT1, u0	0.00005	0.00000	37	0.00915	>0.500
level-1, e	0.03103	0.00096	—	—	—

Model Fit

- Deviance = -3340.087375
- Number of estimated parameters = 2