THE PERCEPTUAL AND PRODUCTION TRAINING OF /d, ɾ, r/ IN L2 SPANISH:

BEHAVIORAL, PSYCHOLINGUISTIC, AND NEUROLINGUISTIC EVIDENCE

BY

WENDY HERD

Submitted to the graduate degree program in Linguistics
and the Graduate Faculty of the University of Kansas in partial fulfillment
of the requirements for the degree of Doctor of Philosophy

____
Co-Chairperson Dr. Allard Jongman

____
Co-Chairperson Dr. Joan Sereno

____
Dr. Jie Zhang

____
Dr. Robert Fiorentino

____
Dr. Michael Vitevitch

Date Defended: 5 April 2011

The Dissertation Committee for WENDY HERD
certifies that this is the approved version of the following dissertation:

THE PERCEPTUAL AND PRODUCTION TRAINING OF /d, ɾ, r/ IN L2 SPANISH:

BEHAVIORAL, PSYCHOLINGUISTIC, AND NEUROLINGUISTIC EVIDENCE

Co-Chairperson Dr. Allard Jongman

Co-Chairperson Dr. Joan Sereno

Date approved: 5 April 2011

Abstract

When native speakers of American English begin learning Spanish, their acquisition of native-like pronunciation can be hampered by the tap - trill distinction in words like *coro* 'choir' and *corro* 'I run'. The trill proves difficult because it does not exist in English. Although the tap exists as an allophone of /t/ and /d/ in English words like 'writer' and 'rider', students of Spanish must learn to process it as a phoneme rather than an allophone. Similarly, learners have difficulty acquiring the spirantization of voiced stops, where the /d/ in *codo* 'elbow' is produced as a voiced dental fricative or approximant, which is more like the 'th' sound in English.

This study investigates whether American English-speaking learners of Spanish can be trained to perceive and produce the intervocalic tap, trill, and /d/ contrasts in Spanish. Participants were trained using both perceptual and production training methods. Past research has reported that perceptual training alone improves both perception and production and that production training alone improves both as well, but the production training studies have not been limited to production as trainees have been able to listen to the training stimuli.

This study is important because it systematically controls both training modalities so that they can be directly compared and introduces a third training methodology that includes both perception and production to discover whether perceptual training, production training, or a combination of the two is most effective. This study also uses cross-modal priming and ERP data in addition to traditional tasks (identification and production tasks) to evaluate the effect of training, an innovative use of both tasks to determine if trainees not only perceive and produce the trained L2 contrasts but also if they unconsciously process these contrasts and if they have built new phonemic categories for these sounds.

All three training paradigms improved English learners' perception or production. While production trainees did not improve in their overall perception and declined in their perception of one contrast, perception trainees improved in their production and overall perception, indicating that perception training transfers more effectively than production training.

Acknowledgments

I would like to begin by expressing my gratitude to the co-chairs of my dissertation committee. They have given me both invaluable guidance and unconditional support, even when my goals have seemed overly ambitious. First, I am grateful to have had the opportunity to work with Allard Jongman whose Phonetics II course laid the foundation for my interest in phonetics, second language acquisition, and experimental design. I am also thankful to Allard for the support and training I received as a research assistant on the Jordanian Arabic project. Likewise, I am thankful to Joan Sereno for the support and experience I received as a teaching assistant for Language and Mind. As co-chairs, Allard and Joan have always been available to answer questions and offer advice, responding in a timely manner and in detail to even weekend emails. They have been exceptional teachers, models, mentors, and friends.

I have also appreciated the assistance and advice I have received from the other members of my dissertation committee. I appreciate the help and support I have received from Robert Fiorentino. I am especially grateful for the many times Rob has sat down with me to discuss ERP methodology and statistics. Rob always made me feel like a welcome member of the Neurolinguistics and Language Processing Laboratory and has pointed me in the direction of research that will help me to continue EEG work even after I leave KU. I would also like to thank Jie Zhang who, like Allard and Joan, also served as a member of my thesis committee. Jie has always pushed me to consider how my work fits into a larger, theoretical framework and to treat every class project as a potentially publishable work. Even if he has already forgotten, I'll always remember when he asked me, "Where's the revolution?" Finally, Mike Vitevitch from the Psychology Department was more than just an outside observer, offering valuable insights about cross-modal priming research and helping me to work through some of my early questions about statistical analysis.

In addition to the members of my committee, I would also like to thank all of my professors from the Linguistics Department. The collegiality and collaboration apparent in our department gave me a standard to which I can aspire in my future. I have truly enjoyed all of the courses I have taken in our department. In particular, I am thankful to Alison Gabriele and her Second Language Acquisition class, during which I proposed a modest research project that eventually developed into this dissertation.

With the help of co-PI's Allard Jongman and Rob Fiorentino, I was able to procure an NSF Doctoral Dissertation Improvement Grant (ID: 0843653). Given the number of volunteer hours required of each participant, this dissertation would never have been completed in its entirety without the support of NSF. I am also indebted to Corinna Johnson who helped me navigate the red tape necessitated by grant funding.

I would also like to thank my fellow students at the University of Kansas. My classmates from the Empirical Research Seminar sat through more versions of this research than any of us cares to admit, but they were always generous with their praise and suggestions. I am grateful to Alonso Canales-Viquez, Jose Aleman Banon, and Marcela Quintana who volunteered their time and their insights as native Spanish speakers. Kristi Bond, Ella Fund-Reznicek, Jamie Bost, and Stephen Politzer-Ahles from the Neurolinguistics and Language Processing Laboratory also helped me as I learned to administer an EEG and analyze ERP data. I am also thankful to Goun Lee who volunteered to run some of my behavioral experiments during her first semester at KU. Of course, none of this work would have been possible without the KU Spanish 216 students and native Spanish speakers who volunteered hours of their time.

On a more personal note, I am thankful for my wonderful family. I appreciate the support of my parents and siblings, who never doubted that I would complete my PhD, although they were a little surprised at how long it took. I must reserve my most heartfelt thanks for my son, daughter, and husband. Although little Jacob will never remember a time when his mother balanced teaching full-time with completing her dissertation research, Breanna will most likely never forget. She was my strongest supporter even when I missed swim meets or forgot about Girl Scout meetings. Finally, I could never have completed this work without the support of my husband Zac. People always seem amazed that I have been able to be a student, teacher, and mother, but I couldn't have done it without an equally amazing spouse. When I struggled with stress and uncertainty, he never once doubted me. His delusions of my grandeur keep me striving to do my best.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

When second language learners begin learning a new language, they are often faced with phonemes, or contrastive sounds, that do not exist in their native language. Failing to produce these sounds correctly can lead to being perceived as a non-native speaker or being misunderstood. For example, native speakers of American English produce a flap [ɾ] when an underlying /d/ occurs intervocalically preceding an unstressed vowel, so the word ‘pudding’ is produced with medial [ɾ] and not [d]. However, when these speakers learn a language like Spanish, a language that contrasts the flap or tap /ɾ/ with the trill /r/ as well as the dental stop /d/, they have a difficult time learning to contrast the phonemes /ɾ/ and /d/ and learning to distinguish and produce the /ɾ/ - /r/ contrast.

The pronunciation of Spanish by native speakers of American English can be hampered by the tap /ɾ/ - trill /r/ distinction in words like *coro* [ˈko.ɾo] ‘choir’ and *corro* [ˈko.ro] ‘I run’. Native speakers of English have difficulty producing the trill /r/ because the sound does not exist as either a phoneme or an allophone in English (Lord 2005, Face 2006). Although the tap /ɾ/ exists as an allophone of /d/ in American English, learners of Spanish may have difficulty accepting /ɾ/ as a phoneme rather than an allophone /d/ (Face 2006). Similarly, English learners of Spanish have difficulty acquiring the spirantization of voiced stops, where /b,d,g/ are produced as [β,ð,ɣ] or [β̞,ð̞,ɣ̞] intervocalically (Díaz-Campos 2004, Lord 2005, Zampini 1993, 1994, 1997). Recent studies have investigated the acquisition of the /ɾ/ - /r/ distinction and of spirantization by American English-speaking learners of Spanish; however, these two areas of difficulty are usually treated separately. Furthermore, while training methods have been developed to improve second language learners’ abilities to produce non-native contrasts, training English-speaking learners of Spanish to correctly perceive and produce the phonemic /ɾ/ and the allophonic [ð] in Spanish is complicated by the existence of a competing allophone, i.e., the [ɾ] allophone of /d/, and a competing phonological rule, i.e., the flapping of /d/, that occurs in the same environment as spirantization.

1.1 Project Description and Motivation

This dissertation investigates whether native English learners of Spanish can be trained to suppress intervocalic flapping when learning to spirantize intervocalic voiced alveolar stops and to produce phonemic taps and trills in Spanish. In addition, this study investigates the effectiveness of perceptual training methods, production training methods, and combined perceptual and production training methods to find which modality of training proves to be the most effective. In addition to using pre- and post- perception and production tests to evaluate how the trainees improve in comparison to a pre- and post-tested control group of Spanish learners, the current study also uses pre- and post- cross-modal priming experiments and EEG mismatch negativity responses to determine if the trainees exhibit automatic psycholinguistic and neurolinguistic responses more similar to native speakers after training. A change in these psycholinguistic and neurolinguistic responses within the trainees would indicate that, with training, second language learners are able to build new phonemic and allophonic representations for the L2.

This study will advance the knowledge of second language acquisition, particularly with respect to the second language learner's ability to acquire phonemic and allophonic contrasts in the L2. This study will also tease apart the effects of perceptual and production training. Past research has established that perceptual training alone improves both perception and production and that production training alone improves both as well, but the production training studies have not been limited to production as trainees have been able to listen to the training stimuli. This study will systematically control both training modalities so that they can be directly compared and will introduce a third training methodology including both perception and production to discover whether perceptual training, production training, or a combination of the two is most effective. Additionally, while past training studies have focused on the ability of L2 learners to acquire new phonemic contrasts, this study involves the training of a new phonemic contrast, the trill /r/, as well as the reanalysis of an allophonic variant in the L1 as a phoneme in the L2, the tap /ɾ/, and the reanalysis of a phoneme in the L1 as an allophonic variant in the L2, the interdental fricative /ð/. This study also uses cross-modal priming and ERP data to evaluate the effect of training, an innovative use of both tasks to determine how trainees unconsciously process novel sounds in the L2. The project could also result in benefits to second language acquisition and improvements in second language teaching.

1.2 Organization of the Dissertation

The following chapter details the previous work that sets the foundation for this study, including a description of /ð, d, ɾ, r/ in English and Spanish and a review of training, cross-modal priming, and event related potential (ERP) studies. Chapter 2 also presents the hypotheses tested in this dissertation. Chapter 3 describes the methodology used during pretests, training, and posttests. This includes a description of participants, stimuli, and procedures during testing and training sessions. Chapter 4 presents the statistical analyses of the data collected. The section walks through first the posttest and then generalization test of the perception section. Next, the results of the native speaker identification of learner productions are presented, followed by an acoustic analysis of the posttest and generalization stimuli. The cross-modal priming and MMN results are detailed in the final two sections of Chapter 4. Chapter 5 brings together the conclusions drawn based on the above analyses, identifies unanswered questions, and suggests future directions for research.

Chapter 2: Background

2.1 /ð,d,ɾ,r/ American English and Spanish

When L1 speakers of American English acquire Spanish, they must reanalyze two sounds that exist in their native language and learn a new sound in order to acquire the three-way /d, ɾ, r/ contrast in Spanish. The first portion of this chapter describes how the /ð/, /d/ and /ɾ/ are categorized in English. The second portion describes how these phones and /r/ are categorized differently in Spanish. Finally, the third portion reviews recent studies investigating American English learners' L2 acquisition of these sounds in Spanish.

2.1.1 /ð,d,ɾ,r/ in American English

Although the trill /r/ does not exist as an allophone or phoneme in English, the interdental voiced fricative /ð/, alveolar voiced stop /d/, and alveolar flap /ɾ/ are familiar sounds for speakers of American English. Unlike Spanish, where the dental voiced fricative [ð] or approximant [ð̞] only exists as an allophone of /d/, as discussed below, the /ð/ is a phoneme in English which contrasts with /d/ as seen in the minimal pair *though* [ðoʊ] – *dough* [doʊ]. On the other hand, the [ɾ] only surfaces as an allophone of /d/ (and /t/) in American English. For example, the word *rider* can be realized as [ɹaɪdɚ] or [ɹaɪɾɚ], the latter as a result of flapping.

Flapping, a highly productive phonological rule in English, causes /d/ (and /t/) to surface as [ɾ] in post-tonic intervocalic position. Studies have reported that flapping occurs 94% (Patterson and Connine 2001), 96% (Connine 2004), or 99% of the time in the post-tonic intervocalic position (Zue and Laferriere 1979; Byrd 1994; and Herd, Jongman, and Sereno 2010). In a recent study, Boomershine et al. (2008) found that monolingual English speakers rated /d, ɾ/ minimal pairs more similar than native Spanish speakers and displayed a greater latency than Spanish speakers when discriminating the pairs. This suggests that

American English-speaking learners of Spanish will experience perception difficulties also. Since flapping occurs so frequently in an environment where Spanish spirantization also occurs, American English learners may produce intervocalic /d/ as a /ɾ/ in Spanish also, both failing to spirantize /d/ correctly and producing a form that can be confused with another phoneme in Spanish as discussed below.

2.1.2 /ð,d,ɾ,r/ in Spanish

In Spanish, the /d,ɾ,r/ are separate phonemes; however, there are no minimal triplets that distinguish the three because /d/ is spirantized to [ð] or [ð̞] intervocalically, /r/ does not occur word-initially, and /ɾ,r/ are in free variation word-finally. However, the contrast can still be illustrated by looking at a minimal triplet and a minimal pair. The minimal triplet *codo* [koðo] 'elbow' – *coro* [koɾo] 'choir' – *corro* [koro] 'I run' illustrates that the /ɾ,r/ contrast with each other and [ð], the allophonic variant of /d/. Likewise, the minimal pair *dato* [dato] 'fact' – *rato* [rato] 'time' shows the /d, r/ distinction.

As with flapping in English, Spanish spirantization, where voiced stops /b,d,g/ are spirantized to [ß,ð,ɣ], is a highly productive phonological rule in Spanish, with intervocalic spirantization of /d/ occurring 99% of the time (Waltmunson 2005). Since /ð/ contrasts with /d/ in English and since Spanish spirantization occurs in the same environment as English flapping, this difference in how /ð,d,ɾ/ are categorized in the two languages may cause difficulties for English learners of Spanish. The following portion of the chapter details the difficulties reported for English speakers acquiring the /d,ɾ,r/ contrast in Spanish.

2.1.3 SLA difficulties

2.1.3.1 English-speaking learners' acquisition of /ɾ/ and /r/

In order to investigate the difficulties that American English learners of Spanish have acquiring the /ɾ/ - /r/ distinction, Face (2006) recorded 41 native speakers of American English from the Midwest reading a short story in Spanish. The participants were divided into two groups: an intermediate group of 20 students in their fourth semester of college-level Spanish and an advanced group of 21 students who were earning their major or minor in Spanish and who were enrolled in advanced Spanish courses. Additionally, five native speakers of Spanish were also recorded reading the short story. The intermediate and advanced groups' productions of twenty target words from the story, ten targets containing /ɾ/ and ten containing /r/, were then compared to each other and to the productions of the native Spanish-speaking group.

Face (2006) found that native speakers produced both the /ɾ/ (92% of the time) and the /r/ (86% of the time) significantly more often than both the intermediate group (/ɾ/ - 49%, /r/ - 5%) and the advanced group (/ɾ/ - 79%, /r/ - 27%). Furthermore, the advanced group used the tap and trill significantly more often than the intermediate group. Although the percentage of tapped targets produced by the advanced group was approaching the production of taps by the native speakers, the percentage of trills produced by both groups and the percentage of taps produced by the intermediate group fell far below the native speaker norm. Face (2006) also analyzed the types of errors made by each group. When mispronouncing the /ɾ/, intermediate speakers replaced it with the English retroflex /ɹ/ 92% of the time and advanced speakers replaced it with /ɹ/ 72% of the time. In the case of /r/ errors, intermediate students produced them as /ɾ/ 31% of the time and as the English /ɹ/ 52% of the time, while advanced students produced them as /ɾ/ 78% of the time and as /ɹ/ 11% of the time. This shows that, regardless of level of acquisition, English-speaking learners of Spanish continue to confuse the tap and trill and to struggle with acquiring the two phonemes.

In an earlier study, Lord (2005) examined whether or not explicit phonetic instruction would improve the production of the Spanish trill /r/ by English-speaking learners of Spanish. Seventeen undergraduate students enrolled in an advanced course in Spanish Phonetics were recorded reading a paragraph from a novel in Spanish. They were recorded reading the paragraph five times throughout the study: once as a pretest during the first week of class, three times throughout the class when it was used for self-analysis, and once as a posttest after the course was complete.

When comparing the percentage of correctly produced /r/ from the pretest to the posttest, the accuracy rate increased significantly from 26% - 39%, leading Lord (2005) to conclude that explicit phonetic instruction can be used to improve a Spanish learner's acquisition of /r/. However, the study contained several methodological flaws. First, the experimental group receiving phonetic instruction was not compared to a comparable control group not receiving instruction, a flaw also pointed out by Lord, so it is unclear whether the phonetics course led to an improvement in /r/ production. Second, the participants read the same passage, a brief paragraph, five times, so any improvement could be due to practice and might not generalize to conversation or to novel reading tasks. Third, the passage only contained six words containing /r/, so the improvement could be due to the acquisition of a few words rather than the phoneme /r/.

2.1.3.2 English-speaking learners' acquisition of spirantization

In several studies, Zampini (1993, 1994, 1997) analyzed the production of intervocalic /b,d,g/ by 32 English learners of Spanish in order to ascertain whether English learners applied the obligatory spirantization rule that causes these phonemes to surface as the allophones [β,ð,ɣ]. The 32 English learners were divided into a low intermediate group of 17 speakers and a high intermediate group of 15 speakers. These participants, as well as a group of 5 native speakers of Spanish, completed an informal conversation task, comprised of responding to questions like 'Where are you from' and 'What do you do in Spanish class' in Spanish, and a formal reading task, during which they read a culture passage in Spanish.

Zampini (1993, 1994, 1997) found that while native speakers spirantized target stops 100% of the time, low intermediate and high intermediate learners produced [β,ð,ɣ] significantly less often than the native speakers. Similarly, high intermediate learners produced [β,ɣ] more frequently than low intermediate learners, but the pattern differed for [ð]. Both low and high intermediate learners produced /d/ as [ð] much less frequently than the other fricatives (low intermediate – 6% in reading, 5% in conversation; high intermediate – 10% in reading, 6% in conversation), and, unlike the other fricatives, there was no significant difference between low and high intermediate learners and no difference between the informal conversation and formal reading conditions.

In her studies, Zampini attributes learners' lag in intervocalic /d/ spirantization compared to other voiced stops to the existence of the phoneme /ð/ in English. Since /ð/ is a phoneme in English, learners are less likely to use it as an allophone for intervocalic /d/, whereas /β,ɣ/ are not phonemes of English, so English speakers are more likely to use them as allophones of /b,g/. It is also possible that English learners are hesitant to produce intervocalic /d/ as /ð/ not because /ð/ is a phoneme, but because /d/ is produced as the allophone [ɾ] in a similar environment in American English. Zampini (1993) does not include a detailed error analysis but notes that spirantization errors always involved producing a voiced oral stop instead of a fricative. Since Zampini's studies did not include acoustic analyses of the targets, it is possible that the English learners produced intervocalic /d/ as a flap [ɾ], not a [d], which would indicate that English learners produce intervocalic /d/

as a fricative or approximant less often than /b,g/ due to the flapping rule in English, not due to the phonemic status of /ð/.

2.2 Previous Training Studies

2.2.1 Perceptual Training Studies

Previous research in perceptual training methods has focused on developing and using systematic training methods to coach second language learners to distinguish, and as a result produce, new contrasts found in the target language. Most of the researchers discussed below test the hypothesis that laboratory perceptual training can improve the second language learner's ability to contrast the L2's phonemes when they are used in a new context, i.e., by new speakers or in previously unknown words, and the learner's ability to transfer his/her perceptual learning to the production domain due to a link between perception and production.

In order to train native French speakers to perceive the distinction between English voiced and voiceless interdental fricatives, Jamieson and Morosan (1986) used a mixture of synthetically produced and naturally produced /θ/-initial and /ð/-initial consonant vowel syllables (CV). After the twenty participants completed pretests, comprised of identification and discrimination tasks, the trainee group completed 90 minutes of training spread out over two or three sessions. The training session began with one tape of identification tasks with only synthesized CV tokens. After participants heard each token, they would press a switch labeled 'teeth' for /θ/ or 'the' for /ð/. When participants had successfully completed the training session on tape one, they could continue to the next tape, where additional tokens were added. By the end of the 12-tape series, the training tapes included background cafeteria noise to increase the difficulty of the task.

After training, Jamieson and Morosan (1986) found that the trainee group improved in their identification of both voiced and voiceless tokens, but the control group showed no improvement. The researchers only mention using the sixteen CVs produced by one male speaker and the eight synthesized CVs, so it appears that the tokens used during the pretest, the training session, and the posttest were all the same. Considering that the trainee group listened to 90 minutes more of the same speaker and tokens than the control group, it is unclear whether the trainee group improved because they had become accustomed to that one speaker and those tokens or because they could distinguish /θ/ and /ð/ more accurately. Additionally, it seems unlikely that training sessions lasting only 90 minutes total and that being trained to distinguish only CV syllables spoken by one native speaker would actually improve a learner's overall ability to distinguish the two phonemes in natural speech. Using a larger variety of speakers and the sounds /θ/ and /ð/ in more contexts might be of more use to learners.

Noting the lack of variety in training sessions, Logan et al. (1991) hypothesized that changing the training stimuli to include the target phonemes in a larger variety of contexts would improve learners' abilities to distinguish the sounds. Logan et al.'s (1991) study focused on native speakers of Japanese who underwent training to perceive the difference between /l/ and /ɹ/ in English. The researchers used 207 English minimal pairs in which /l/ and /ɹ/ contrasted in word-initial, word-final, and intervocalic position and that contrasted by occurring in consonant clusters or as singletons. Also, instead of using only the productions of one native speaker during training, like Strange and Dittmann (1984) and Jamieson and Morosan (1986), Logan et al. (1991) recorded five different native speakers of English, introducing speaker-variety to the training session in order to increase the generalization of the /l/ - /ɹ/ distinction to new words and new speakers.

After the pretests, the six native speakers of Japanese, who had lived in the United States between 6 months and 3 years, were presented with 68 minimal pairs twice during each 40-minute session of training, resulting in 272 minimal pairs per training session. Like in the Jamieson and Morosan (1986) training task, participants would hear a word and then choose /l/ or /ɹ/ by using a response box. If an answer were wrong, the correct response would light up and the token would be repeated, after which the trial would continue. The six participants completed a total of fifteen 40-minute training sessions. As a result of the posttests, Logan et al. (1991) found that all six participants improved significantly in their ability to distinguish /ɹ/ and /l/, improving their mean correct responses from 78.1% to 85.9%. However, they found that the distinction between /ɹ/ and /l/ in word-initial position and consonant cluster position did not improve as much as the distinction between /ɹ/ and /l/ in final position and in intervocalic position. In addition to using a posttest, Logan et al. (1991) administered two generalization tests, one with new words read by one of the five training session speakers and one with new words read by a speaker never heard by the participants before. The researchers found that participants performed numerically better when the new tokens were read by an "old" speaker (83.7% mean correct responses) than when they were read by a "new" speaker (79.5% mean correct responses). Logan et al. (1991) conclude that using a combination of speakers and a larger variety of stimuli as a training method effectively trains Japanese listeners to perceive /ɹ/ and /l/ in isolated English words, improving their ability to distinguish between the two. However, they point out that /ɹ/ and /l/ are still difficult for the trainees to distinguish in certain contexts.

In a similar study about training Japanese speakers to distinguish /l/ and /ɹ/ in English, Bradlow et al. (1997) compared an 11-person trainee group to a 10-person control group. Bradlow et al. (1997) used the stimuli, pretests, training sessions, and posttests discussed above in the Logan et al. (1991) article, but they continued the sessions for a longer period of time, and they followed the participants' progress for a longer period of time. After 45 training sessions, spread out over a 3 to 4 week span, the researchers found that the trainee group improved from 65% on the identification pretest to 81% on the posttest. Likewise, the trainee group performed equally well on the two generalization tests, scoring 83% correct with the "old" speaker and 80% correct with the "new" speaker. Unlike the trainee group, the control group's performance declined after the pretest.

In addition to measuring the participants' improvement in perception, Bradlow et al. (1997) also measured how their production of English /l/ and /ɹ/ changed. As both a pretest and posttest, the Japanese learners completed a speech production task where they repeated a random list of minimal pairs containing /ɹ/ and /l/. The participants both saw the word they were to say and heard it pronounced by a native speaker of English. Based on native American English speakers' preference judgments of the trainee group's productions and on native American English speakers' identification of the Japanese productions, the trainee group's production of /l/ and /ɹ/ improved as a result of training. Although Bradlow et al. (1997) made contributions to this line of study by recognizing and measuring the ability of perceptual training to improve production, the production task may not have elicited natural speech samples. Since the participants heard the word read by a native speaker immediately before they produced it, it is possible that the participants merely mimicked the production of the native speaker. Future studies should include both repetition tasks and other production tasks to see if there is a performance difference between the two.

After contributing to perceptual training by documenting the link between perception and production, Bradlow et al. (1999) made a further contribution when they conducted a three-month follow up on the control and the trainee groups from the 1997 study. They found that after a three-month lapse in training, the trainee group maintained their levels of performance from the pretest. In fact, while their scores on perceptual tests tended to be 1 – 3

% lower than the posttest (a statistically insignificant decline), the trainee group’s mean production, assessed by native speaker judgments, improved about 4% (also statistically insignificant).

Previous researchers having established the effect of perceptual training on segments, Wang, Jongman and Sereno (2003) investigated whether the perceptual training of suprasegmentals, in this case the four Mandarin tones, could improve the learners’ production of these tones. Using the same format as Logan et al. (1991) and Bradlow et al. (1997), Wang, Jongman and Sereno (2003) pretested and posttested 16 native speakers of English, eight of whom underwent training. The pretest consisted of the participants reading a list of 80 monosyllabic Mandarin words, 20 each of the four tones. After the two-week training program, which consisted of only 40 of the 80 recorded stimuli, both the trainee group and the control group repeated the production task with the trainee group performing significantly better than the control group. While the trainee group’s production evaluation scores by 80 native speakers increased 18% on old stimuli (the forty tokens used during training) and 13% on new stimuli (the forty tokens only seen on the pre- and posttest), the control group’s production evaluation increased only 4% on old stimuli and 1% on new stimuli. Based on native speaker judgments and acoustically measured improvement of the trainee group’s productions, it is clear that the training sessions improved their production of the suprasegmental tones.

These studies establish the effectiveness of perceptual training methods to train both segmental and suprasegmental second language phonemic contrasts, but no one has investigated how effective these training methods will be when the contrasts being trained are complicated by their existence as allophones in the first language, as is the case with the tap /ɾ/, or by competing phonological rules that apply in the same context in the first and second language, as is the case with the English flap [ɾ] and the Spanish fricative / approximant [ð/ð̞].

2.2.2 Production Training Studies

In addition to perceptual training, a small body of research has also developed surrounding production training of new phonemes using audio-visual teaching aides. Forty years ago, Bluhme and Burr (1971) hypothesized that using visual images of students' productions of tone, making use of new technology like the pitchmeter and the oscilloscope, would improve native English speakers' production of the four Mandarin tones. In order to test their hypothesis, they had ten native English participants, five with knowledge of Mandarin Chinese and five with no such knowledge, participate in an audio-visual training session. While this study did not use high variability training as discussed above, it did introduce the idea that visual images could improve learners' production. If learners could see a visual image of their speech in real time, it seems like the visualization may help them to improve their production, but it would likely take more than the six repetitions recommended by Bluhme and Burr (1971).

More recently, Hirata (2004) investigated how the use of a computer program allowing students to see a visual image of their productions as they pronounce target Japanese words assists native English speakers in acquiring Japanese pitch and duration and how this training affects their perception of those contrasts. Before the training session, eight native English speakers completed a production pretest of 21 words, only nine of which were also used during training, and a perception pretest of 30 words, none of which were used during training. The native English speakers were separated into a 4-person control group and a 4-person trainee group. The trainee group attended ten 30-minute training sessions, consisting of 33 minimal pairs and 2 triplets of Japanese words contrasting in tone and duration. Participants saw an acoustic visual display of a target form on the computer, and then they tried to reproduce the form. They could listen to the word via headphones if they wished, but they were not required to do so. After the training session, trainee group and control group participants completed the production and perception posttest, which were identical to the pretests. Hirata (2004) found that the trained group improved on both production and perception, improving significantly more than the control group. The trainee group improved from 69.3% on the perception pretest to 83.5% on the posttest, but the control group only improved from 69.3% to 77.1%.

Since the participants in Hirata (2004) could listen to the stimuli, it is unclear whether their improvements were due to the production training alone or due to inadvertent perceptual training. While the participants, who were trained in production, also improved in perception, no clear link between production training and perceptual improvement can be made as the participants were allowed to listen to the stimuli. No one has attempted to manipulate perceptual and production training so that the effects of the two can be compared, a gap which is addressed in the current study.

2.3 Methodological background

2.3.1 Cross-Modal Priming

While production and identification tasks can be used to show whether or not learners are able to produce and perceive L2 contrasts when they must perform overt tasks where they are thinking about the contrasts, these tests do not shed light on whether participants automatically process L2 contrasts in a native-like manner. A cross-modal priming task, where participants hear an auditory stimulus (the prime) and then see a related or unrelated visual stimulus (the target) about which they must make a lexical decision, allows researchers to use participants' reaction times in order to see if the participants recognize the phonetic similarity between the auditory prime and the visual target (Tabossi 1996). Previous research has found that auditory primes followed by identical visual targets were responded to significantly more quickly than primes followed by rhyming targets with different onsets

(Marslen-Wilson and Zwisterlood 1989), followed by competitor targets differing in a final phoneme (Marslen-Wilson 1990), or followed by unrelated targets (Grainger et al. 2001).

One can thus predict that if native speakers of Spanish were presented with the auditory stimulus [kaɾa], they would exhibit the shortest reaction times for the visually identical target *cara* [kaɾa], with longer reaction times for a target differing in one phoneme like *cada* [kaða]. On the other hand, native speakers of American English, who perceive /ɾ/ as an allophone of /d/, presented with the same auditory stimulus [kaɾa] would be predicted to exhibit the opposite pattern, with the shortest reaction times for *cada* [kaða], a stimulus that was not produced, and the longest for *cara* [kaɾa], the auditory prime.

For native Spanish speakers, the same pattern should hold true for other pairs, like /ɾ/ - /r/. If native Spanish speakers hear a word like *caro* [kaɾo], they will exhibit shorter reaction times when this is followed by *caro*, but they should exhibit longer reaction times if the stimulus is followed by *carro*. By contrast, if American English learners are processing /ɾ/ as an allophone of /d/, there should be no reaction time difference between the stimulus above followed by *caro* and *carro*. A pre- and post-training cross-modal priming task could therefore be used as a diagnostic to evaluate whether English-speaking learners of Spanish exhibit reaction times in a pattern more similar to native Spanish-speaking controls after being trained to perceive and/or produce those contrasts.

2.3.2 MMN

Another method that can measure whether or not participants automatically process phonemic differences is the mismatch negativity response (MMN) measured using EEG or the mismatch field (MMF) measured using MEG. Measures of MMN and MMF have been shown to reflect whether or not listeners detect the differences between standard and deviant stimuli regardless of whether listeners are attending to or ignoring the auditory stimuli. The response is elicited 150 – 250 ms after the presentation of a deviant stimulus using the oddball paradigm described in Phillips et al. (2000), where many “standard” stimuli within a phonemic category are followed either by an acoustically unique stimulus within the same category or by an “oddball”, an acoustically unique stimulus from a different phonemic category. The MMN/MMF responses elicited using the oddball paradigm have been shown to reflect whether listeners detect phonemic differences in stimuli.

Näätänen et al. (1997) found that the amplitude of MMN responses increased with the amount that the deviant stimuli differed acoustically from the standard stimuli in a study where the standard stimulus was the vowel /e/ and the deviant stimuli were the vowels /ö/, /õ/, and /o/. They found that for both Estonian and Finnish speakers, the MMN response increased as the deviant vowel differed more from the standard with one exception. The Estonian and Finnish speakers’ MMN responses differed when the deviant stimulus was the vowel /õ/, a phoneme in Estonian but not in Finnish. While the Estonian speakers’ responses showed an increased MMN amplitude as would be expected due to the increased difference between the deviant and the standard, the Finnish did not show an increased response to this vowel, so Näätänen et al. (1997) concluded that the MMN is sensitive to both acoustic differences and to the phonemic inventory of the language.

Phillips et al.’s (2000) oddball paradigm, described above, elicited responses that supported the view that the MMF/MMN response is sensitive to phonemic as well as acoustic differences. In this design, there is no acoustic standard, because the standard stimuli all vary acoustically within a phonemic category, but there is a phonemic standard. If the MMN/MMF response were only sensitive to acoustic differences, one would not expect to get a mismatch response, because no many-to-one acoustic ratio is created. However, Phillips et al. (2000) found that when the standard included various acoustic versions of ‘da’, a deviant ‘ta’ elicited an MMF response, but an acoustically unique deviant ‘da’ did not. This supports the view that the MMN/MMF is sensitive to phonemic differences.

While previous studies investigated whether a contrast was important depending on whether or not the sound existed in a given language, Kazanina et al. (2006) looked at two different languages, Russian and Korean, in which the difference between /d/ and /t/ is important for different reasons. The difference is phonemic in Russian but allophonic in Korean. While the acoustic difference between /t/ and /d/ is not important for meaning in Korean, it changes systematically based on environment, /t/ occurring word-initially and /d/ word-medially, so the acoustic difference is important to the language. Kazanina et al. (2006) found that when a standard /da/ was followed by a deviant /ta/, a strong MMF response was elicited in the Russian speakers but not in the Korean speakers.

Kazanina et al. (2006), Näätänen et al. (1997), and Phillips et al. (2000) provide evidence that the MMN/MMF responses of monolingual speakers are sensitive to phonemic contrasts in their languages, but not allophonic differences; however, they do not address how MMN/MMF responses to allophonic and phonemic differences might vary in bilinguals. Winkler et al. (1999) explores this question by measuring the MMN response of native Finnish speakers, native Hungarian speakers fluent in Finnish, and native Hungarian speakers naïve of Finnish to standard /e/ followed by deviant /æ/ or /y/. All three groups exhibited significant MMN responses to the /e/-/y/ contrast, a contrast that exists in both Hungarian and Finnish. For the /e/-/æ/ pair, unique phonemes in Finnish but allophones in Hungarian, the native Finnish speakers and native Hungarian speakers fluent in Finnish exhibited a significant MMN, but no response was found in Hungarian speakers with no knowledge of Finnish. This indicates that learning a second language changes the way vowels are perceived and that second language learners are capable of building new phonemic vowel categories in their L2.

Contrastively, in a follow-up study, Winkler et al. (2003) found that native speakers of Hungarian who were fluent in Finnish could not “unhear” the /e/-/æ/ contrast when the vowels were presented in a Hungarian context. Twelve native Hungarian speakers fluent in Finnish were presented with standard /pæti/ followed by deviant /peti/. While these two lexical items have different meanings in Finnish, ‘was qualified’ and ‘bed’ respectively, they are the same word, a nickname for ‘Peter’, in Hungarian. The stimuli were presented first in a Hungarian-context, created by giving directions in Hungarian and telling subjects that they would hear Hungarian words, followed by the stimuli in a Finnish-context. The Hungarian-Finnish bilingual participants exhibited a significant MMN response to the /pæti/-/peti/ pairs in both contexts, and the contexts did not differ from each other. Differences between monolingual and bilingual MMN responses have also been reported in Peltola et al. (2003, 2005, 2007), yet these studies have only probed MMN responses to vowels and they have not investigated the effects of training on MMN responses.

Tremblay et al. (1997) and Tremblay, Kraus, and McGee (1998) used MMN measurements to pre- and post-test monolingual English trainees and controls. The trainees were trained to identify a three-way contrast between /mba/ (-50 ms to -20 ms VOT), /ba/ (-10 ms to 20 ms VOT), and /pa/ (30 ms to 50 ms VOT). While the /ba – pa/ contrast exists in English, /mba/ was a new phone for the trainees. After training, the trainees exhibited increased MMN amplitudes when a /ba/ standard was followed by a /mba/ deviant while the controls showed no change in MMN responses. Tremblay et al. (1997) and Tremblay, Kraus, and McGee (1998) were able to conclude that adult learners can reclassify existing phonemic categories as a result of training. In addition, both studies found the identification of the bilabial /mba – ba/ contrast transferred to the corresponding alveolar contrast /nda – da/ and that the MMN amplitude to deviant /nda/ preceded by standard /da/ also increased significantly. Although these studies show that identification training can affect the MMN response of learners, the contrast trained in these studies was arbitrary.

A logical extension of previous research is to use MMN measures as pre- and post-tests to evaluate the effectiveness of training phonemic categories that exist in a second language. For example, using the oddball paradigm from Phillips et al. (2000), if native Spanish speakers are presented with standard versions of *ere* [eɾe], a MMN response should occur if the deviant *ede* [ede] is introduced since /ɾ/ and /d/ represent different phonemes in Spanish. For native English speakers, on the other hand, /ɾ/ is an allophonic variant of /d/, so no MMN response is predicted. It will be informative to see if English-speaking learners of Spanish will exhibit a MMN response like native Spanish speakers after training. As with cross-modal priming tasks, pre- and post-MMN responses could be used to evaluate whether training an L2 contrast will affect the way the trainees automatically process auditory stimuli and reclassify existing phonemic categories. It differs from cross-modal priming in that MMN measures an earlier, pre-lexical stage of processing. It is possible that trainees will have built new phonemic categories for the trained contrasts, resulting in an increased MMN response, but will not have encoded those new contrasts in their lexicons, resulting in no measurable change in cross-modal priming reaction times.

2.4 Theoretical Background

Linguists and psychologists have developed a variety of theories to account for the increased difficulty exhibited when acquiring a second language as compared to acquiring a first language. The Critical Period Hypothesis, the strong version of which claims that learners must be exposed to L2 stimuli before a critical age to attain native-like proficiency, has received much attention from the public as well as linguists. Flege's (1995) Speech Learning Model asserts that L2 learners' ultimate attainment of L2 pronunciation depends on their ability to perceive the distinctions between L2 phonemes, with sounds unique to the second language being easier to acquire than sounds that are confusable with phonemes found in the learners' L1. Similarly, Best's (1991) Perceptual Assimilation Model contends that learners' ability to discriminate L2 phonemes depends on how closely those phonemes assimilate to L1 phonemic categories. Brief descriptions of these theoretical schemas follow, and predictions based on these theories are presented.

2.4.1 Critical Period Hypothesis

Lenneberg's (1967) Critical Period Hypothesis identified puberty as the critical age by which a human must be exposed to language input in order to attain functional language proficiency. Although the work was originally based on incomplete language acquisition of neglected children who received impoverished input or on incomplete language recovery by adult aphasics, Lenneberg also postulated that adult second language learners would be unable to attain native-like pronunciation of an L2, primarily due to language processes being left-laterialized and crystallized by that age. Since adult learners' brains lack the neural plasticity found in children from birth to puberty, they would be unable to acquire new phonological categories, resulting in foreign-accented speech. While Lenneberg's (1967) original hypotheses focused on the acquisition of the L2 phonological system, the Critical Period Hypothesis was later expanded to include other aspects of language, including syntax, morphology, and semantics (Snow and Hoefnagel-Höhle 1978, Johnson and Newport 1989, Birdsong and Molis 2001). However, as the present study focuses on the perception and production of L2 phonemic contrasts, this review is limited to studies that investigate how the Critical Period Hypothesis relates to acquiring an L2 phonological system.

In order to find if age of acquisition, the age at which learners first began learning or first came into contact with the L2, is correlated with degree of foreign accent in the L2, Seliger, Krashen, and Ladefoged (1975) surveyed 243 learners of English and 121 learners of Hebrew. Almost all learners who came into contact with their second language prior to age

10 self-reported having no foreign accent while most learners who came into contact with the L2 after age 16 reported having an accent. There were exceptions, with 7.6% of post-pubescent learners reporting no accent and 8.3% of pre-pubescent learners reporting an accent. Similarly, Oyama (1976) recorded and scored the productions of 60 Italian learners of English, finding that degree of foreign accent was significantly correlated with age of arrival in the United States. Basically, like Seliger, Krashen, and Ladefoged (1975), Oyama (1976) found that the later a learner began acquiring English, the more likely that individual would produce Italian-accented English, with immigrants who arrived after age 12 being identified as the most strongly accented.

While this strong correlation between age of acquisition and accented L2 productions has been repeatedly verified both anecdotally and in empirical studies like those described above, researchers have also identified a small number of post-pubescent L2 learners who attained native-like productions, like the 7.6% of learners reported in Seliger, Krashen, and Ladefoged (1975). Ioup et al. (1994) presented a case study of two English learners of Egyptian Arabic who were consistently identified as native speakers in spite of coming into contact with the language after puberty. Bongaerts et al. (1997) examined whether the pronunciations of highly proficient Dutch learners of English would be identified more closely with native English speakers or less proficient Dutch learners of English. Of the 11 highly proficient learners recorded, a small number (5 in the first study and 3 in the second) were consistently identified as native-English speakers. Moyer (1999) investigated the perceived accent of 24 English learners of German, finding, like previous studies, that age of acquisition was significantly correlated with accented speech. However, one learner, who first came into contact with German at age 22, was consistently identified as a native speaker. Likewise, in their investigation of Mandarin learners of English, Munro and Mann (2005) reported that late L2 learners were more likely to be identified as accented than early learners, yet a small number of late learners were identified as native speakers while a small number of early learners were identified as accented. Together, these studies indicate that a later age of acquisition is usually correlated with a more foreign accent in the L2, but native-like attainment of the L2 phonology is possible and has been documented in a small number of post-pubescent learners.

While these studies make apparent the possibility of acquiring native-like pronunciation after a critical age, they do not measure the extent to which this acquisition is related to brain plasticity. In more recent training studies, researchers have attempted to measure whether acquiring new L2 contrasts in behavioral tasks results in changes in neural responses. If high variability training or intensive input can result in trainees exhibiting more native-like brain responses, this would indicate that post-pubescent learners' neural pathways are malleable when provided with sufficient input, contradicting the predictions and neural basis of the Critical Period Hypothesis. Wang et al. (2003) reported that six English learners of Mandarin exhibited increased neural activation as measured by fMRI in Brodmann's area and Wernicke's area as a result of high variability training of the four Mandarin tones. Using MEG, Zhang et al. (2009) elicited a posttest MMF response from seven Japanese learners of English who were successfully trained to discriminate English /ɹ/ and /l/. While both of these studies were based on a small number of trainees, they present convincing evidence that neural plasticity still exists in adult language learners.

Given the success of the high variability training studies discussed previously and the enhanced neural sensitivity documented in the two studies discussed above, it is likely that the training paradigms used in the present study will also lead to improved perceptual identification of the trained contrasts. If measured improvement in perception occurs, one would also expect the trainees to exhibit heightened sensitivity to the contrasts, resulting in increased MMN responses as a result of training. These results would indicate that, while learning L2 contrasts may be more difficult after a critical age or sensitive period, such learning is not impossible nor even limited to exceptional language learners.

2.4.2 Speech Learning Model

While Flege, Munro, and MacKay (1995) reported similar findings to those reviewed in the previous section, they did not attribute the inability of most late language learners to acquire native-like pronunciation to a critical age. While age of acquisition strongly influenced the ability of Italian learners of English to produce L2 contrasts, Flege, Munro, and MacKay (1995) asserted that the results are predicted by the Speech Learning Model (Flege 1995) than the Critical Period Hypothesis.

According to the Speech Learning Model, L2 learners only form new phonemic categories for L2 sounds when those sounds are distinguished from existing L1 contrasts. In most cases, like the difference in VOT between English and Italian /b, d, g/ (MacKay et al. 2001), the L2 sounds are merged into existing L1 categories. This merger results in bilinguals differing in their pronunciation of contrasts from both L1 and L2 monolinguals. For example, MacKay et al. (2001) found that Italian learners of English produced pre-voiced /b, d, g/ in Italian less often than Italian monolinguals, indicating that L2 English /b, d, g/ merged with existing Italian /b, d, g/ categories. On the other hand, when L2 contrasts are easily discriminated from existing L1 categories, new category formation may result.

Based on this model, learners are able to build new phonemic categories only after perceiving the difference between L2 sounds and existing L1 categories. Since English learners of Spanish, for whom the tap /ɾ/ is an allophone of /d/, will need to perceive the difference between the /ɾ/ and /d/ in Spanish in order to build a new phonemic category, it is predicted that they will be unable to distinguish this pair at pretest. On the other hand, as the difference between the tap /ɾ/ and trill /r/ is more salient, it is predicted that English learners of Spanish will more easily distinguish this pair and more quickly form a new phonemic category for the trill /r/. Unlike the Critical Period Hypothesis, the Speech Learning Model predicts that learning L2 contrasts as an adult is possible, and it offers an explanation as to why building new phonemic categories might be more difficult in the L2 than the L1.

2.4.3 Perceptual Assimilation Model

Best's (1991, 2001) Perceptual Assimilation Model also offers an alternative to the Critical Period Hypothesis. According to the Perceptual Assimilation Model, learners will assimilate L2 contrasts into existing L1 categories depending on how similar the sounds are to those existing categories. Crucially, unlike Flege's Speech Learning Model, the Perceptual Assimilation Model does not predict that learners will build new phonemic categories.

The learner's ability to discriminate between any two L2 contrasts is dependent upon how the contrasts assimilate to the L1. For example, L2 sounds will be easily distinguishable when they assimilate well to two different categories. In the present study, the Perceptual Assimilation Model predicts that learners will easily distinguish [ð] from [ɾ], because these exist as two separate categories in the L1, as seen in the minimal pair *other* [ʌðɚ]– *udder* [ʌɾɚ]. When two sounds are assimilated to the same category with equal degrees of goodness, learners would exhibit poor discrimination, but if one sound were perceived as a better representative of that category, learners would be able to distinguish the two sounds. Finally, if one sound were "non-assimilable," not able to assimilate to an existing L1 category, learners would be able to easily distinguish that sound from another as it would be considered a "non-speech" sound. This would be the case for distinguishing the [ɾ] – [r] and [ð] – [r] contrasts for the learners in the present study, because the trill [r] would not assimilate to an English phonemic category. In short, Best's Perceptual Assimilation Model predicts that learners will be able to easily discriminate all three paired contrasts in the present study at pretest. Further, as the Perceptual Assimilation Model does not predict learners' ability to build a new phonemic category, it also predicts that learners will not exhibit MMN response changes as a result of training.

2.5 Hypotheses

While the link between perception and production has been established by previous perceptual training studies, it is unclear if production training will also lead to improved perception and if a combination of perceptual and production training will be more effective than either training method alone. Also, most of the segmental evidence focuses on the /ɹ/ - /l/ contrast in English and its difficulty for Japanese speakers learning English, so it will be useful to look at a contrast in a language other than English that is difficult for non-native speakers.

The remainder of this dissertation presents a study that evaluates the effectiveness of using perceptual training and production training methods separately and concurrently to train native English speakers to perceive and produce the Spanish /ɾ, r, d/ contrasts. The research investigates whether native English-speakers can be trained to produce the allophonic American English tap as a separate phoneme in Spanish and if they can acquire Spanish intervocalic spirantization in spite of the existence of a competing intervocalic flapping rule in their native phonology. Furthermore this study will measure whether perceptual training has a greater effect on production or production training has a greater effect on perception.

It is likely that the combined perceptual and production training group will outperform the perception training only, production training only, and control groups during the posttest, demonstrating that combined perceptual and production training is the most effective training method. Comparing the perception-only and the production-only groups should help to determine which has a greater effect on the other: production or perception. Considering that participants' production in the above-mentioned studies improved less than their perception, whether they were trained exclusively on perception or production, a tentative hypothesis can be made that production training will have a larger effect on perception due perhaps to the resistance of production to training.

Finally, the addition of cross-modal priming and MMN tasks should shed light on how participants process these contrasts after training. Using cross-modal priming will demonstrate whether learners can immediately encode these new contrasts in their lexicon. If improvement is exhibited in the perception and / or production of these contrasts, one can hypothesize that participants will have built new phonemic categories for the tap /ɾ/ and /r/, leading to native-like MMN responses.

Chapter 3: Methods and Experimental Design

3.1 Participants

The productions of nine native Spanish speakers (5M) with an average age of 26, eight from Peru and one from Spain, were recorded to create pretest, training, and posttest stimuli. Forty-two native speakers of American English (9M) with a mean age of 20 who were enrolled in an intermediate college-level Spanish course (Spanish 216) at the University of Kansas also participated as controls and trainees. These students had completed three to four years of high school Spanish and were enrolled in their second or third semester of college Spanish. They were randomly assigned to four groups: perception training, production training, combination training, and control. A second group of twelve native Spanish speakers from Peru, Costa Rica, Honduras, Ecuador, and Mexico (3M) and with a mean age of 29 also participated as native controls. A third group of eight native Spanish speakers from Chile (3M) with an average age of 23, none of whom had ever traveled in an English-speaking country, participated as judges for the identification and rating tasks. It is important to note here that many varieties of Spanish exist; however, the phonemes /ɾ, r/ and the intervocalic allophone [ð / ð̞] investigated in this study are present in all varieties of Spanish. Additionally, every attempt was made to recruit speaking participants who use the same variety of Spanish.

All participants completed a human consent form, a dialect questionnaire (see Appendices A and B), and the Edinburgh Handedness Inventory (Oldfield 1971, Appendix C) before completing any pretests, training, or posttest sessions. The Edinburgh Handedness Inventory confirmed that all participants were right-handed. All participants were paid $10 per hour for their participation, and the learners of Spanish, who were required to visit the lab from two to twelve times depending on group, were paid an additional $20 completion bonus upon the completion of all sessions.

3.2 Materials

Nine native Spanish-speaking participants [8 from Peru (4 M) and 1 (1M) from Spain] read 252 minimal pairs: 84 that contrast /ɾ/ and /r/ like *coro* ‘choir’ and *corro* ‘I run’, 84 that contrast /d/ and /r/ like *moda* ‘fashion’ and *morra* ‘chic’, and 84 that contrast /ɾ/ and /d/ like *loro* ‘parrot’ and *lodo* ‘mud’. For each of these contrasts, 45 of the minimal pairs were word – nonword pairs while the other 39 were word – word pairs.

Of the 252 minimal pairs, 30 minimal word – word pairs and 30 minimal word – nonword pairs were used for the pretests, posttests, and generalization tests. A paired samples t-test verified that the word – word pairs used on the pretests, posttests, and generalization tests did not significantly differ from each other in word frequency as determined by the Corpus del Español (Davies 2002) [$t(59) = 0.343$, $p = 0.733$]. Furthermore, paired samples t-tests confirmed that the /ɾ/-/r/ pairs [$t(19) = 0.994$, $p = 0.333$], /r/ - /d/ pairs [$t(19) = 1.284$, $p = 0.214$], and /ɾ/-/d/ pairs [$t(19) = 0.568$, $p = 0.577$] did not differ significantly from one another in word frequency. The pretest and posttest stimuli are listed in Appendix D and the generalization stimuli in Appendix E.

Thirty of the minimal pairs described above (15 word – word and 15 word – nonword) and produced by F1 (a female speaker from Peru) were used to create the perception pretest and posttest. These thirty minimal pairs read by M1 (a male speaker from Peru) and M5 (a male speaker from Spain) were also used for the old words – new speaker and old words – new dialect generalization tests. Note that the generalization tests, like the perception tests, contained both words and nonwords; however, they are referred to as “old word – new speaker” or “new word – new speaker” to avoid unnecessary wordiness and confusion when using acronyms (i.e., new stimuli – new speaker would be abbreviated NSNS). The same list of tokens was then used as the pretest and posttest production stimuli.

Twenty-four additional tokens containing intervocalic taps in the form CVɾV and produced by F1 were used as the auditory primes (and visual targets in one-third of the cases) during the pre- and post-cross-modal priming tasks. The cross-modal priming stimuli were balanced for word frequency to ensure that prime – target word pairs [$t(11) = 0.100$, $p = 0.922$] and different contrasts amongst the visual targets [$F(2, 21) = 1.149$, $p = 0.336$] did not differ in frequency. Appendix F lists the auditory primes and the visual targets used during the cross-modal priming task.

An additional ninety minimal pairs read in equal parts by speakers F1, M1, and M5 were used to create new word – familiar speaker, new word – new speaker, and new word – new dialect generalization tests. As mentioned above, the generalization stimuli are listed in Appendix E. The remaining 132 minimal pairs read by three female (F2, F3, and F4) and three male speakers (M2, M3, and M4) were used during training sessions. The training stimuli can be found in Appendix F. The pretest, training, posttest, and generalization tasks are discussed in more detail below. Additionally, speaker F1 who read the pretest and posttest stimuli also read twenty repetitions of two bisyllabic tokens: *ere* [eɾe] and *ede* [ede], which were used to create the pretest and posttest mismatch negativity (MMN) task.

3.3 Pretest, Training, and Posttest Designs

Prior to and following training, English-speaking participants completed perception, production, cross-modal priming, and MMN tasks that were used to evaluate their acquisition of the /ɾ/, /r/, and /d/ contrasts in Spanish. The control group of 10 native Spanish speakers also completed the cross-modal priming and MMN tasks. The tasks were presented in the same order to all participants.

The perception pretest and posttest were identical forced-choice perceptual identification tasks presented via Paradigm (Tagliaferri 2008). The task included thirty minimal pairs read by a native Spanish speaker from Peru (1NF) for a total of 60 tokens. Participants first heard an auditory stimulus that contained [ɾ], [r], or [ð] intervocalically, and then they saw two words on the computer screen, the orthographic representation of the word they heard and the other word in the minimal pair. For example, participants might hear *cara* 'face' [ˈka.ɾa] and then see *cara* 'face' and *cada* 'each'. Their task was then to mouse-click the word they heard. The stimuli were presented in random order.

In addition to the perception posttest, participants completed a five-part perceptual generalization task, identical in presentation to the perception pretest and posttest. In parts one and two, participants identified the thirty minimal pairs previously used in the pre/posttest as read by a male speaker from Peru (new speaker) and a male speaker from Spain (new dialect). In the final three parts, participants identified ninety new minimal pairs, thirty read by the female speaker from Peru (new words, familiar speaker), thirty read by the male speaker from Peru (new words, new speaker), and thirty read by the male speaker from Spain (new words, new dialect). The perception and generalization stimuli were presented as one experiment, but the stimuli were blocked by speaker. The speakers were always presented in the following order: familiar speaker, new speaker, new dialect. The combined perception and generalization identification task included a total of 60 minimal pairs read by three speakers for a total of 360 stimuli.

During production pretest and posttest tasks, participants read the thirty minimal pairs used in the perception pretest and posttest in a randomized list including 40 additional words as fillers. During the posttest, participants also read an additional 30 minimal pairs taken equally from the generalization perception tests.

The pre and post cross-modal priming tasks were identical. Participants heard a series of 24 Spanish words and nonwords, the auditory primes, that contained an intervocalic flap /ɾ/, like *cara* 'face' [ˈka.ɾa] or *nura* (nonword) [ˈnu.ɾa]. Next, participants saw the orthographic representations of either identical words, in these cases *cara* or *nura*, or different words, like *cada* ‘cada’ or *nuda* ‘nude’. These were the visual targets. The participants’ task was to decide whether or not the items they saw were words.

For native speakers of Spanish, when they hear *cara* [ˈka.ɾa] and then see c*ara*, a visual target identical to the auditory prime, their response times should be faster than when they hear *cara* [ˈka.ɾa] and then see *cada,* a different target than the auditory prime. On the other hand, English-speaking learners of Spanish, who may identify the tap in *cara* [ˈka.ɾa] as an underlying /d/, may exhibit faster response times for *cada* than *cara.* Using the cross-modal priming task as a pre and post assessment will help determine if training causes learners to recode phonemic categories for nonnative contrasts in the lexicon, which would be demonstrated by response time patterns more similar to those of native speakers. The cross-modal priming task is also a less overt measurement than the identification and production tasks, and it evaluates whether learners have developed automaticity when perceiving and producing the /ɾ/, /r/, and /d/ contrasts in Spanish.

Like the cross-modal priming task, using the MMN task as a pre and posttest will help to determine how learners unconsciously process nonnative contrasts and whether training affects their perception of these contrasts. Whereas participants are instructed to listen to the auditory prime but to respond to the visual target during the cross-modal priming task, participants are instructed to ignore the auditory stimuli during the MMN task. Since the MMN response can be elicited even when participants are not attending to the auditory stimuli, it measures whether or not L2 trainees are able to automatically detect the phonemes of Spanish. During the MMN task, participants watched a silent movie while listening to syllables over headphones. The syllables were arranged in a 7-to-1 oddball paradigm, where an average of seven acoustically different tokens within the same phonemic category, seven *ere* [eɾe] syllables for example, were followed by an acoustically and phonemically different token in Spanish, an *ede* [ede] syllable.

It should be noted that *ede* would usually be pronounced as [eðe], not [ede], in Spanish; however, [ede] is included in the MMN stimuli to test the findings of Kazanina et al. (2006). Kazanina et al. (2006) found that although Korean speakers never produce /t/ intervocalically, they did not exhibit a MMF response when intervocalic /d/ was followed by intervocalic /t/, so Korean speakers did not detect the difference between intervocalic /t/, which they never produce, and intervocalic /d/, which they do produce, because they are acoustically different variants of the same phoneme in Korean. Hence similar predictions can be made concerning Spanish.

In this paradigm, native Spanish speakers are expected to exhibit a MMN effect following the oddball syllable *ede* [ede], but no such MMN effect would be exhibited following another *ere* [eɾe] syllable. Native English speakers on the other hand, who assign [d] and [ɾ] to the same /d/ phonemic category, are more likely to exhibit no MMN effect for either *ere* [eɾe] or *ede* [ede], because the "oddball" stimulus is still within the same phonemic category. If the training sessions improve learners' perception of the /ɾ/, /r/, and /d/ contrasts and lead learners to build new phonemic categories for Spanish contrasts, the posttest MMN results should look more similar to those of native Spanish speakers, with MMN effects between [eɾe] and [ede].

3.3.1 Perception Training

One group of ten participants underwent perception training following the procedure described in Logan et al. (1991) and refined in Bradlow et al. (1997). The participants were trained using 120 minimal pairs recorded by six different speakers. During the training sessions, which lasted between 20 – 30 minutes, the participants completed forced choice identification tasks similar to the perception pretests and posttests. After hearing a word that contained either [ɾ], [r], or [ð] over a pair of headphones, participants saw two orthographic choices on a computer screen. Participants then chose the item they heard by mouse-clicking their response. After choosing an item, participants either saw the message, "Right! That was token. Let's hear token again," or "Oops! That was token. Let's hear token again," at which point the auditory stimulus was replayed. Participants attended six training sessions during a period of two to three weeks, practicing one pair of sounds, i.e., [ɾ] vs. [ð], [ɾ] vs. [r], or [ð] vs. [r], read by two different speakers each day. Two sessions were spent on each contrast, and the contrasts and speakers were never repeated in consecutive sessions.

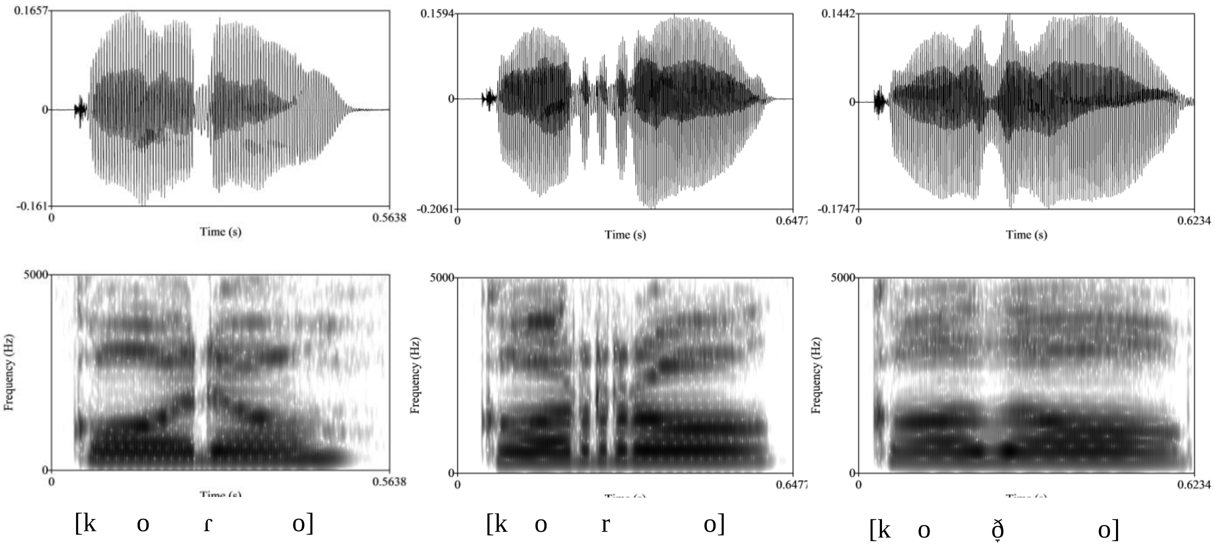
3.3.2 Production Training

Ten participants underwent production training following a procedure based on that described in Hirata (2004). As was the case with perception training, production trainees were presented with 120 minimal pairs read by six different speakers. Trainees practiced one contrast per session for a total of six sessions completed within two – three weeks. Neither the same contrast nor the same speaker was trained on consecutive days.

During training, participants were presented with the waveform and spectrogram of a native speaker's production of a word via Praat (Boersma & Weenink 2005). Each participant was prompted to inspect the native speaker's production, and then to click 'continue' when ready to record a version of the word. The program would record the participant for 1.5 s, and then the participant's waveform and spectrogram would appear. The participant would next be prompted to compare the two versions of the stimulus, and then to press 'continue' when ready to see a new word. Participants were instructed to attempt different pronunciations in order to match their waveforms and spectrograms to those of the native speakers and to continue using a pronunciation once the waveforms and spectrograms matched. Production trainees were never allowed to hear the native speakers' stimuli.

The first training session lasted about 60 - 75 minutes, half of that time devoted to training how to distinguish [ɾ], [r], and [ð] using waveforms and spectrograms. As seen in Figure 3.1below, a tap [ɾ] consists of one short closure while a trill [r] consists of a series of 2 to 10 taps (Quilis 1993). The [ð] was visually distinguished from the [r] and [ɾ] by the

Figure 3.1. Waveforms (above) and spectrograms (below) of 'coro' (left), 'corro' (middle), and 'codo' (right).



presence of frication or the approximation of frication instead of one complete closure. After completion of the first session, the other five production training sessions lasted 35 – 45 minutes each.

3.3.3 Combination Training

The third group included eleven participants, all of whom completed both perceptual and production training. This group completed three perceptual training sessions and 3 production training sessions within two - three weeks, rotating each modality from session to session. As was the case with perception and production trainees, these participants practiced each paired contrast twice, once through perception training and once through production training.

3.3.4 Controls

A fourth group of eleven American English speakers learning Spanish completed the pretests and posttests but did not undergo training. These participants, called the Spanish-learning controls, completed the posttests two to three weeks after the pretests. A group of ten native Spanish speakers were also part of a control group called Spanish-speaking controls. These participants completed the cross-modal priming and the MMN tasks.

Chapter 4: Results

4.1 Identification Task

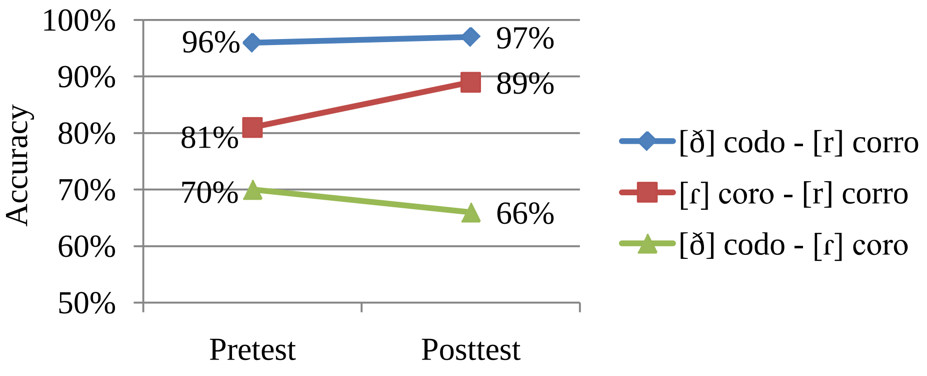
4.1.1 Perception Results

A 2 × 3 (Test × Contrast) repeated measures Analysis of Variance was conducted on the percent of correct responses with Group as a between subjects factor. The two levels of Test refer to the perception pretest and posttest, and the three levels of Contrast include /d/ vs. /ɾ/, /ɾ/ vs. /r/, and /d/ vs. /r/. Group refers to the four different experimental groups of Spanish learners: perception trainees, production trainees, combination trainees, and controls.

First, a main effect of Test neared significance [$F (1, 38) = 4.173$, $p = 0.058$], and a main effect of Contrast [$F (2, 76) = 390.466$, $p < 0.0001$] reached significance. The main effect of Test indicates that average posttest scores across the four groups (84%) were significantly higher than average pretest scores (82%). The main effect of Contrast means that the three different contrasts were not equally difficult for participants to identify. The /d/ – /r/ contrast was the easiest for participants to distinguish with an average 97% accuracy, while the /ɾ/ – /r/ contrast, with an 85% accuracy rate, proved more difficult, and the /ɾ/ – /d/ contrast proved the most difficult with a 68% accuracy rate. Pairwise comparisons indicated that all three contrasts differed significantly from each other at the $p < 0.0001$ level.

As can be seen in Figure 4.1 below, an interaction between Test and Contrast also reached significance [$F (1, 38) = 10.116$, $p = 0.003$]. This interaction indicates that while accuracy rates for distinguishing /d/-/r/ remained at ceiling levels from pretest to posttest, the /ɾ/-/r/ and /d/-/ɾ/ accuracy rates changed. Since /d/-/r/ accuracy rates were at ceiling levels and since a paired samples t-test comparing /d/-/r/ pretest and posttest accuracy did not exhibit significant change [$t (41) = 0.961$, $p = 0.342$], this contrast will not be discussed

Figure 4.1. Accuracy of identification responses from pretest to posttest by contrast.



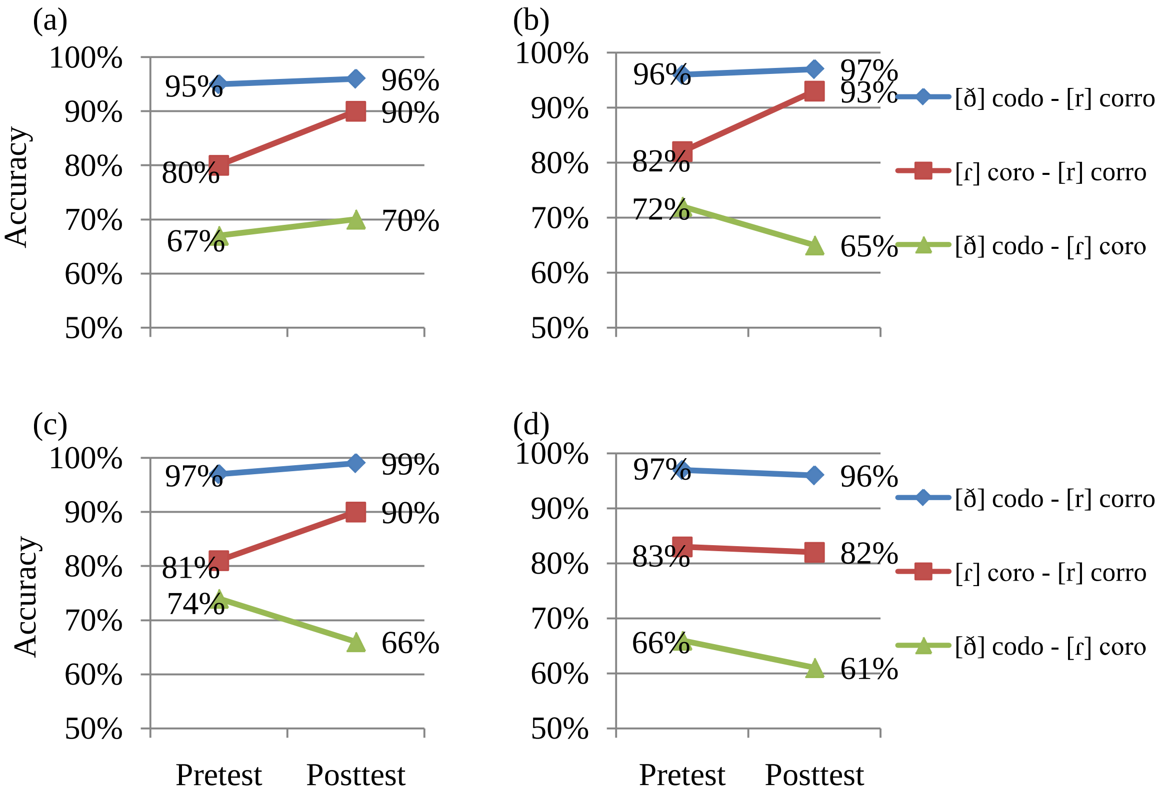
further. On the other hand, a paired samples t-test found that the /ɾ/-/r/ accuracy rate improved significantly from pretest to posttest [t (41) = 3.824, p < 0.0001], while the decline in /d/-/ɾ/ accuracy also reached significance [t (41) = 2.876, p = 0.006]. In order to determine how training affected these contrasts, one must look at interactions between Test, Contrast and Group.

The interaction between Contrast and Group proved nonsignificant [F (6, 76) = 1.957, p = 0.137], but the Test and Group interaction reached significance [F (3, 38) = 3.860, p = 0.017]. This interaction indicates that, depending on training, groups performed differently from pretest to posttest across the three different contrasts. While the accuracy of all three training groups improved 2 – 5% between pretest to posttest, the controls' accuracy declined 2%. Pairwise comparisons found that the Perception Trainees' overall improvement from 81% to 85% proved significant [p = 0.032]; however, the Production Trainees' improvement from 83% to 85% [p = 0.234] and the Combination Trainees' improvement from 84% to 85% [p = 0.503] did not reach significance. Meanwhile, the Controls' decline from 82% to 80% [p = 0.026] proved significant. While this Test by Group interaction shows that perception training proved the most effective training method overall for training perception, the three-way interaction between Test, Contrast, and Group best illustrates the effectiveness of training.

As illustrated in Figure 4.2 below, the three-way interaction between Test, Contrast, and Group was nearing significance [$F (6, 76) = 2.671$, $p = 0.061$]. This interaction indicates that the participants' posttest improvement (or decline) in the identification of the three contrasts differed depending on which training the participants received. As predicted and shown in Figure 4.2d, the controls did not improve in the identification of any contrast. Paired samples t-tests comparing pretest to posttest perception accuracy found that the controls' ability to distinguish the /ɾ/-/r/ [$t (10) = .265$, $p = 0.796$] and the /d/-/ɾ/ contrasts [$t (10) = 1.437$, $p = 0.181$] did not evidence significant change. By comparison, all three training groups improved an average of 9% to 11% in their perception of the /ɾ/-/r/ contrast. This improvement proved significant for the perception [$t (9) = 2.366$, $p = 0.045$], production [$t (9) = 2.538$, $p = 0.032$], and combination trainees [$t (10) = 2.654$, $p = 0.024$]. Although none of the training groups improved significantly in the perception of the /d/-/ɾ/ contrast, the perception trainees evidenced an average 3% numerical increase. By contrast, both the production trainees [$t (9) = 3.545$, $p = 0.006$] and the combination trainees [$t (10) = 2.924$, $p = 0.015$] declined significantly in the perception of this contrast.

The results of the perception pretest and posttest reveal that speakers of American English have difficulty distinguishing both the /ɾ/-/r/ (81% accuracy at pretest) and the /d/-/ɾ/ (70% accuracy at pretest) contrasts, a result that has not been reported in previous research. As a result of training, regardless of the modality, all trainees were able to improve their ability to differentiate between the /ɾ/ and /r/ while controls without training showed no improvement. Furthermore, given that the production and combination trainees declined significantly in their perception of the /d/-/ɾ/ contrast while perception trainees improved numerically, perception training appears to be the most effective training modality for improving the perception of nonnative contrasts.

Figure 4.2. Identification accuracy of (a) perception trainees, (b) production trainees, (c) combination trainees, and (d) controls from pretest to posttest by contrast.



Of the two contrasts that presented difficulty for American English speakers, the /d/-/ɾ/ contrast proved the most resistant to training. Since the pretest accuracy of this contrast was significantly lower than the /ɾ/-/r/ accuracy, one could have predicted that it would show the most, rather than the least, improvement. This clearly demonstrates that the /d/-/ɾ/ differs in some way from the /ɾ/-/r/ contrast. To distinguish the /ɾ/-/r/ contrast, American English speakers only need to acquire the /r/ as a new phonemic category, because the [ɾ] already exists as part of the allophonic inventory of English. When distinguishing the /d/-/ɾ/ however, American English speakers must acquire the [ð] as an allophone of /d/ and reassign the [ɾ], an allophone of /d/ in American English, to a separate phonemic category. The resistance of the /ɾ/-/r/ contrast to training indicates that reassigning an allophonic variant from the L1 as a separate phonemic category in the L2 is more difficult than creating a phonemic category for a new contrast in the L2.

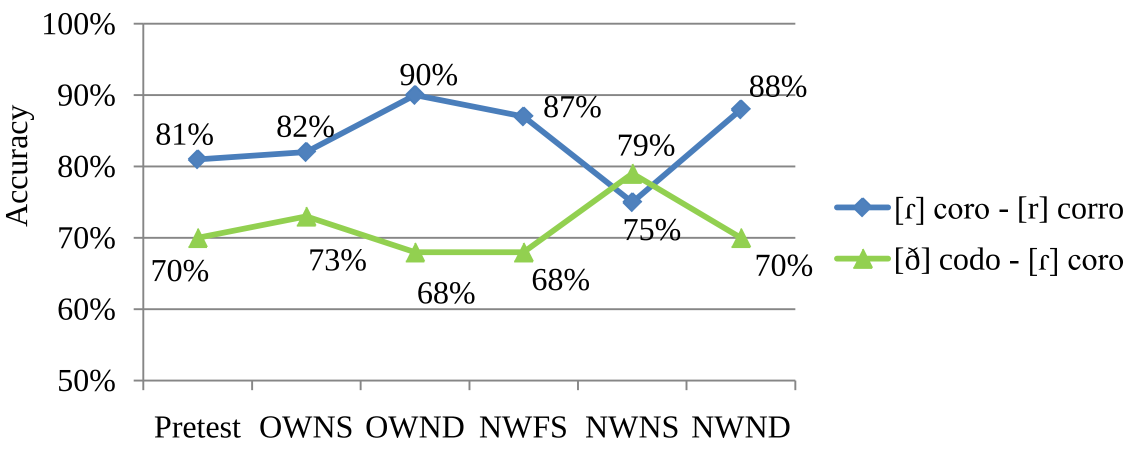
This difficulty is also reflected in the significantly declining accuracy rates of the production and combination trainees. Since the production trainees never heard the /d/-/ɾ/ contrast and the combination trainees only heard it during one training session, they appeared to recognize that distinguishing the /d/ from the /ɾ/ was difficult, but they did not hear enough input to identify the correct cues used to distinguish the two sounds. The improvement of the perception trainees, who heard the /d/-/ɾ/ during two of the training sessions, may have reached significance if they had spent three days training each of the two difficult contrasts (i.e., /ɾ/-/r/ and /d/-/ɾ/) rather than two days on each of those contrasts and two days on the /d/-/r/ contrast. Regardless, these results indicate that perception training is necessary to improve the identification of a contrast that involves reassigning an existing sound in the L1 to a new category in the L2.

4.1.2 Generalization Results

In order to see how effectively training transferred to new speakers, new dialects, and new words, a 6 $\times$ 2 (Test $\times$ Contrast) repeated measures Analysis of Variance was conducted on the percent of correct responses with Group as a between subjects factor. The six levels of Test refer to the identification pretest and the five separate generalization tests: old words – new speaker; old words – new dialect; new words – familiar speaker; new words – new speake; and new words – new dialect. The two levels of Contrast refer to the /ɾ/ – /r/ and the /ɾ/ – /d/ contrasts; the /d/ – /r/ contrast was not included in these analyses due to ceiling effects found when comparing identification pretest to posttest. Group again refers to Perception Trainees, Production Trainees, Combination Trainees, and Controls.

First, a significant main effect of Test [$F(5, 190) = 2.349$, $p = 0.043$] indicates that learners performed differently on the pretest than the generalization tests. Pairwise comparisons show a general increase from pretest (75%) to the generalization tests (78% to 79%) with the difference between the pretest and the new words – new dialect (79%) nearing significance [$p = 0.055$]. A main effect of Contrast also reached significance [$F(1, 38) = 54.164$, $p < 0.0001$], meaning that participants performed differently in their identification of the /ɾ/ – /r/ and the /ɾ/ – /d/ contrasts. Pairwise comparisons show that participants perceived the /ɾ/ – /r/ contrast correctly (84%) significantly more often that the /ɾ/ – /d/ contrast (71%) [$p < 0.0001$]. An interaction between Test and Contrast also reached significance [$F(5, 190) = 36.642$, $p < 0.0001$], because, as can be seen in Figure 4.3 below, participants distinguished

Figure 4.3. Accuracy of identification responses from pretest to generalization tests (old words – new speaker [OWNS], old words – new dialect [OWND], new words – familiar speaker [NWFS], new words – new speaker [NWNS], new words – new dialect [NWND]) by Contrast.

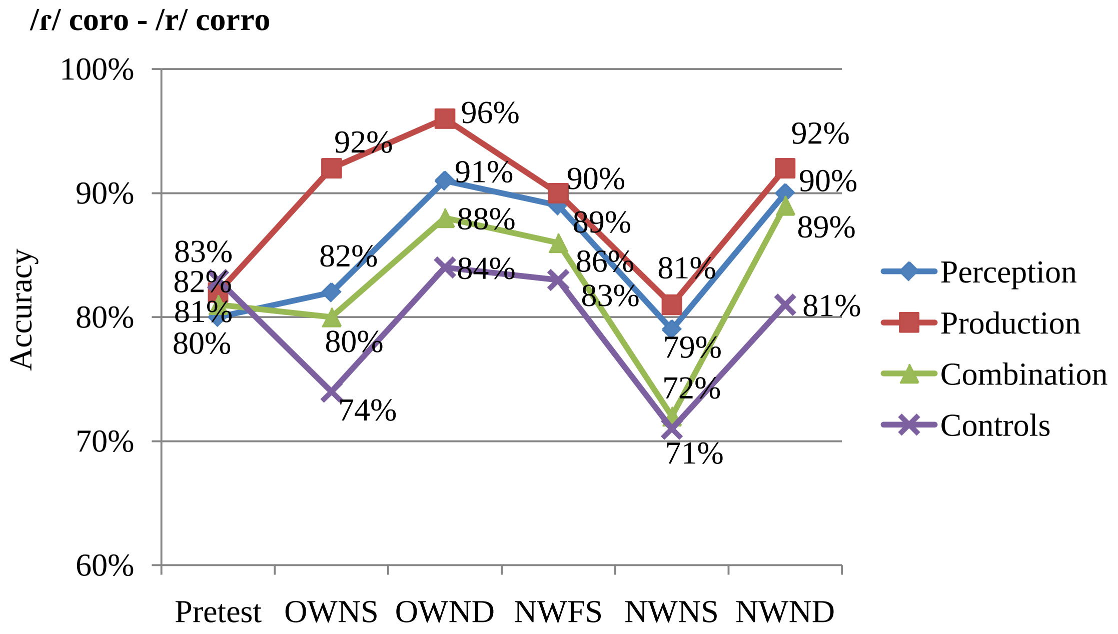


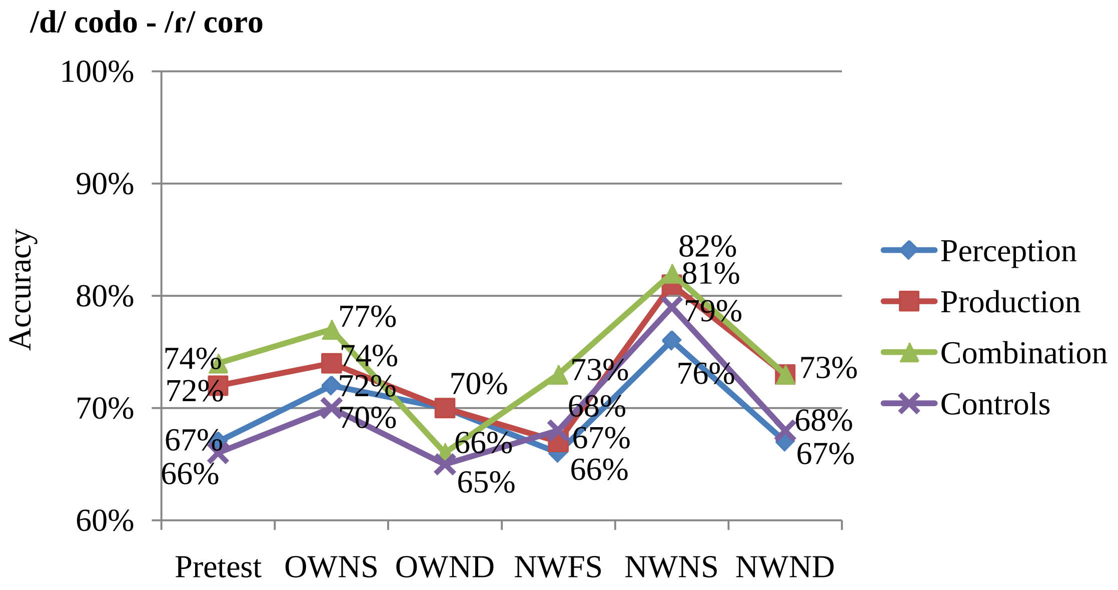
the /ɾ/ – /r/contrast significantly more accurately than the /ɾ/ – /d/ contrast [all p-values below 0.0001]; however, the opposite was true of the new words produced by a new speaker, but this difference did not reach significance [$p = 0.103$]. While interesting, these effects do not shed any light on the effectiveness of training.

The three-way interaction between Test, Contrast, and Group was nearing significance [$F(3, 38) = 2.681$, $p = 0.060$], meaning that participants performed differently from pretest to generalization test depending on whether they received training and which contrasts were tested. Figure 4.4 below illustrates the change in accuracy from pretest to generalization tests by the different groups. With respect to the /ɾ/ – /r/ contrast illustrated in the top panel of Figure 4.4, a paired samples t-test comparing the pretest to different generalization tests found that Perception Trainees significantly improved their perception of the /ɾ/ – /r/ contrast on the old words – new dialect [$t(9) = 2.436$, $p = 0.038$] and the new words – new dialect [$t(9) = 2.535$, $p = 0.032$] generalization tests. Similarly, the Production Trainees exhibited significant improvement on the old words – new dialect generalization task [$t(9) = 3.151$, $p = 0.012$] as well as near-significant improvement on the old words – new speaker [$t(9) = 2.141$, $p = 0.061$] and new words – new dialect [$t(9) = 2.204$, $p = 0.055$] tasks. Likewise, the Combination Trainees improved significantly on the new words – new dialect [$t(10) = 2.789$, $p = 0.019$] and nearly significantly on the old words – new dialect [$t(10) = 1.955$, $p = 0.079$] generalization tasks. By contrast, the Controls made no significant improvements, instead evidencing a significant decline on the old words – new speaker [$t(10) = 2.666$, $p = .024$] and the new words – new speaker [$t(10) = 6.294$, $p < .0001$] tasks. In short, all three training groups transferred their improvement in the perception of the /ɾ/ – /r/ contrast to new words and a new dialect while the Controls made no improvements.

Looking at the /ɾ/ – /d/ contrast in the bottom panel of Figure 4.4, one first notes that all participants, trainees and controls, made significant gains from the pretest to the new words – new speaker task [$p < 0.05$]. Since this is the only task that exhibited improvement

Figure 4.4. Accuracy of identification responses from pretest to generalization tests (old words – new speaker [OWNS], old words – new dialect [OWND], new words – familiar speaker [NWFS], new words – new speaker [NWNS], new words – new dialect [NWND]) by Contrast and Group.





in the /ɾ/ – /d/ contrast and since that improvement is exhibited in similar amounts across all groups, it cannot be due to training. Other than this overall improvement on the new words –

new speaker task, the perception of the /ɾ/ – /d/ on the generalization tasks is very similar to that on the posttest, with no groups showing significant improvement.

Based on the perception pretest, posttest, and generalization data, all three training modalities effectively improve the perception of the /ɾ/ – /r/ contrast, and this improvement transfers to new words and new dialects. Nevertheless, perceptual training most effectively improves perception, because the Perception Trainees also exhibited significant improvement overall while the Production and Combination Trainees did not and the Perception Trainees maintained their ability to distinguish the /ɾ/ – /d/ contrast while the Production and Combination Trainees actually declined significantly in their perception of this contrast from pretest to posttest. Concerning the generalization tasks, all three trainees improved significantly or near significantly in their perception of the /ɾ/ – /r/ contrast during the old words – new dialect and new words – new speaker tasks, but the Controls did not exhibit improvement. Notably, only the Perception Trainees demonstrated significant improvement on both of these tasks, reiterating that perception training most effectively trains perception. Before determining whether perception training transfers better to production or production training transfers better to perception, the results of the production posttest must first be analyzed.

4.2 Production Task

4.2.1 Native Speaker Identification and Rating Results

In order to evaluate the trainees' improvement in production from pretest to posttest, identification and rating data were collected from eight native Spanish speakers from Chile who had never visited an English-speaking country. The purpose of the identification and rating tasks was to find if native Spanish speakers could correctly identify the phoneme intended by the Spanish learner and if intelligibility and pronunciation improved as a result of training. Native Spanish speakers were presented with the pretest and posttest productions of one minimal pair from each of the three paired contrasts read by the forty-two learners of Spanish, resulting in 504 tokens (2 tests × 2 words × 3 contrasts × 42 speakers = 504 tokens). The 504 tokens were heard in two different experimental blocks: the identification task followed by the rating task.

During the identification task, native Spanish speakers heard the word produced by the Spanish learner, and then chose the word they thought they heard from three choices presented orthographically on the computer screen. For example, if the Spanish learner intended *moda* 'fashion', the native speaker would choose from *moda* 'fashion', *mora* 'blackberry', or *morra* 'crown (of the head)'. During the rating task, the native Spanish speaker would see the word intended by the Spanish learner, hear the word pronounced by the learner, and then rate the pronunciation from '1 – Very Accented' to '5 – Unaccented or Near Native'. The results of the identification task are discussed first, followed by the analysis of the rating task.

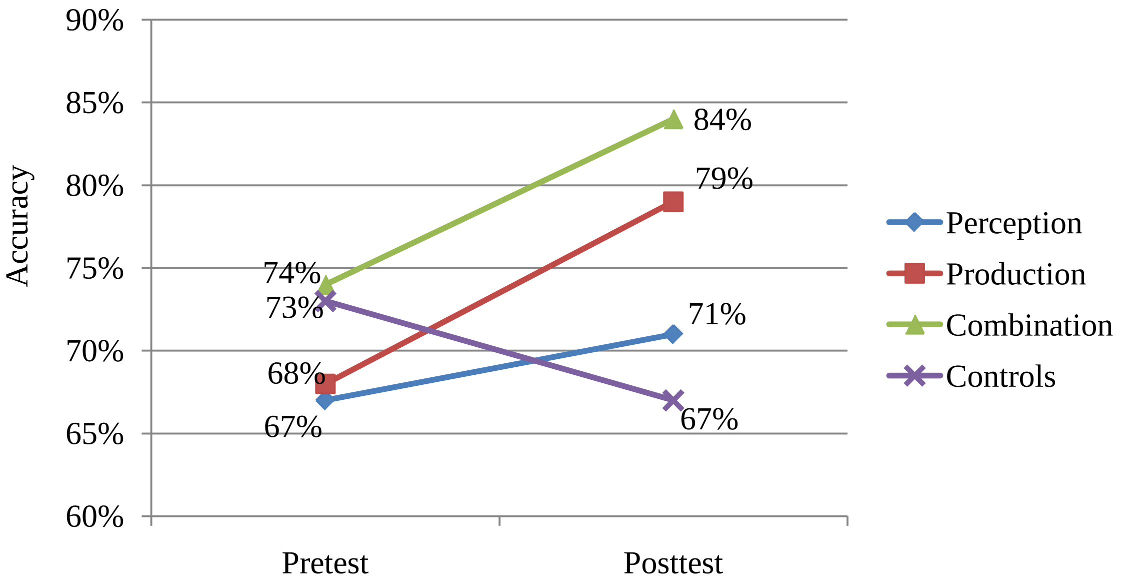
4.2.1.1 Native Speaker Identification Results

A 2 × 3 (Test × Contrast) repeated measures Analysis of Variance was conducted on the percent of correctly identified words with Group as a between subjects factor. As with previous analyses, the variable Test refers to pretest versus posttest productions of the target stimuli, Contrast refers to three different target contrasts /ɾ/, /r/, and /d/. A main effect of Test [$F(1, 38) = 4.188$, $p = .048$] reached significance, because overall intelligibility improved from 70% for the pretest stimuli to 75% for the posttest stimuli. There was also a main effect of Contrast [$F(2, 76) = 12.723$, $p < .0001$] due to the large difference in the perceived identity of the /r/, /ɾ/, and /d/. Pairwise comparisons found that the /r/ was perceived correctly at 57%, significantly less often than both the /ɾ/ at 80% [$p = .004$] and the /d/ at 81% [p <

.0001]. In addition to these main effects, which shed very little light on the effectiveness of training, two interactions were also nearing significance.

The near significant Test by Group interaction is illustrated in Figure 4.5 below [F (3, 38) = 2.772, p = .055]. As can be seen in the figure, the intelligibility of the three training

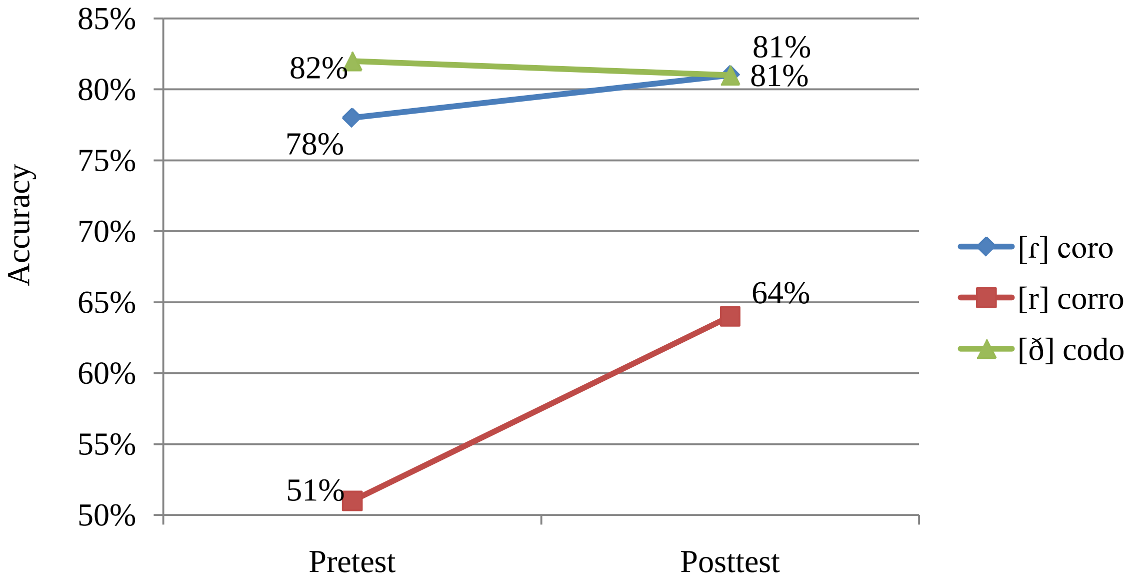
Figure 4.5. Percent of pretest and posttest Spanish learner productions perceived correctly by native Spanish speakers by Group.



groups improves while that of the Controls declines. Paired samples t- tests found that only the Combination Trainees improved significantly in their overall intelligibility [t (10) = 2.641, p = .025] while the Production Trainees were nearing significance [t (9) = 2.391, p = .06]. The changes in the intelligibility of the Perception Trainees and the Controls did not reach significance. A Test by Contrast interaction also neared significance [F (2, 76) = 2.781, p = .068] as shown in Figure 4.6 below. This interaction is due to the drastic improvement in the pronunciation of /r/ as compared to the /ɾ/ and /d/. Paired samples t-tests found that for /r/, the Perception Trainees' improvement from 46% to 61% [t (9) = 2.639, p = .042] and the Production Trainees' improvement from 44% to 70% [t (9) = 2.391, p = .040] reached significance while the Combination Trainees' 58% to 79% gain neared significance [t (10) =

1.906, p = .086]. The Controls, on the other hand, exhibited no such improvement, declining numerically from 56% to 45%.

Figure 4.6. Percent of pretest and posttest Spanish learner productions perceived correctly by native Spanish speakers by Contrast.



The results of the native speaker identification task confirm that all three training types result in improved intelligibility, with both Perception and Production Trainees improving significantly in their pronunciation of the /r/ and with Combination Trainees improving significantly overall. The results of the identification task indicate that trainees did not improve in their pronunciation of the /ɾ/ or /d/, but this could be due to a ceiling effect since pretest /ɾ/ and /d/ identification were near 80%. Since the identification rate of these two contrasts was so high, native speaker ratings might offer a more fine-grained measure of improvement.

4.2.1.2 Native Speaker Rating Results

To determine whether improvement from pretest to posttest for each contrast could be measured using native speaker ratings, paired samples t-tests were run for each group and each contrast. Only the Production Trainees exhibited significant improvement, improving the rating of their /r/ productions from 3.1 to 3.7 on a 5-point scale [$t(9) = 2.547$, $p = .031$]. This confirms that Production Trainees improved their production of the trill, but it offers no evidence of improvement not gained already through the native speaker identification task.

Since native speakers completed the identification task prior to the rating task, with each task containing 504 stimuli and lasting approximately 20 minutes, it is possible that the native speakers became fatigued. This may have resulted in the rating task only being sensitive to very large improvements in the Spanish learners' productions, like the improved pronunciation of 'coro' from [koɹo] to [koro].

4.2.2 Acoustic Analyses

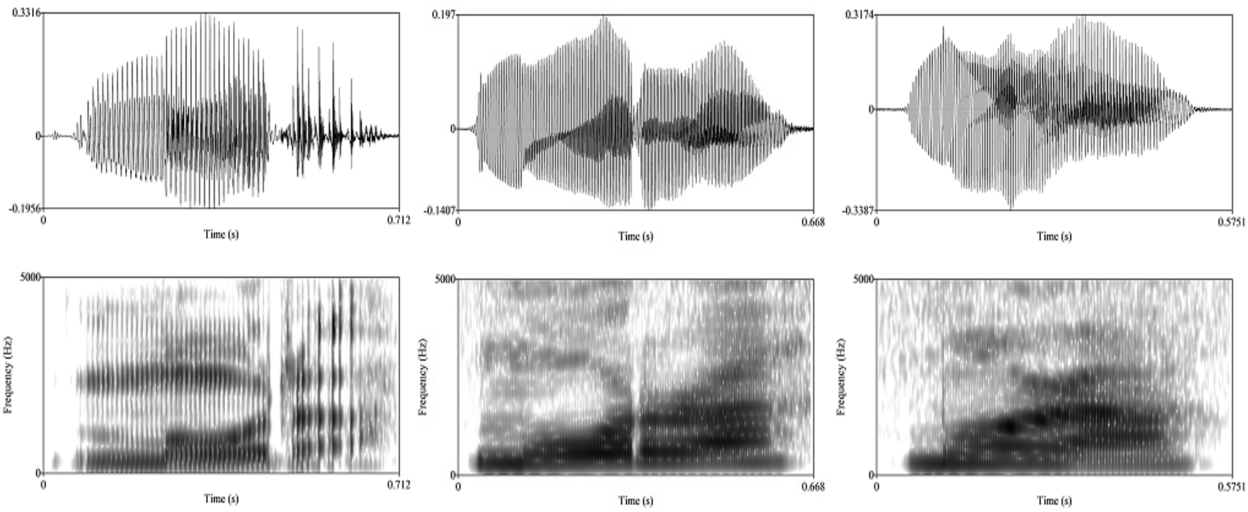
In addition to a portion of stimuli being presented to native Spanish speakers for identification and ratings, all of the pretest, posttest, and generalization productions were analyzed via waveform and spectrogram. Each Spanish learner produced 30 minimal pairs at pretest, the same 30 minimal pairs at posttest, and 30 new minimal pairs as a generalization test, resulting in 180 stimuli. Each stimulus was then analyzed and scored using Praat based on visual inspection of the waveform and spectrogram (Boersma & Weenink 2005). Stimuli received a '0' if the target Spanish phoneme was replaced by an English one, a '0.5' if the production approached the intended target, and a '1' if the intended target was pronounced correctly. A more detailed explanation of how each contrast was scored is presented below. This gradient scoring scale should be able to capture the improvement of Spanish learners who produce the trained contrasts in a more target-like manner without reaching native pronunciation. These scores were then used to conduct the statistical analyses that follow.

If a Spanish learner intended to produce a [ɾ], they received a ‘1’ if the waveform contained one brief and complete closure as in the example of *mora* ‘blackberry’ in Figure 4.7a below. A combination of a Spanish [ɾ] and an American English [ɹ] as seen in Figure 4.7b resulted in a ‘0.5’. Substituting a Spanish [r] for a [ɾ] was also scored a ‘0.5’. Finally, the use of an American English [ɹ] as seen in 4.7c was scored ‘0’. This scoring system reflects that producing a “retroflex tap” [ɾɹ] is better than producing a retroflex [ɹ] without a closure and that substituting a [r] for a [ɾ] is a “native-like” error.

A similar scale was also used to evaluate /d/ and /r/ productions as seen in Figures 4.8 and 4.10 respectively. The waveforms and spectrograms in Figure 4.8 illustrate how productions were scored when the target was *moda* ‘fashion’. If the /d/ in *moda* was produced as a voiced dental fricative [ð] or a voiced dental approximant [ð̞], as in Figure 4.8a, the production was scored as ‘1’. If the Spanish learner produced a [d] as in 4.8b, it was scored 0.5. Because native Spanish speakers very

Figure 4.7. Waveforms and spectrograms of (a) *mora* with a native-like [ɾ], (b) a combination of [ɾ] and [ɹ], and (c) a non-native [ɹ].

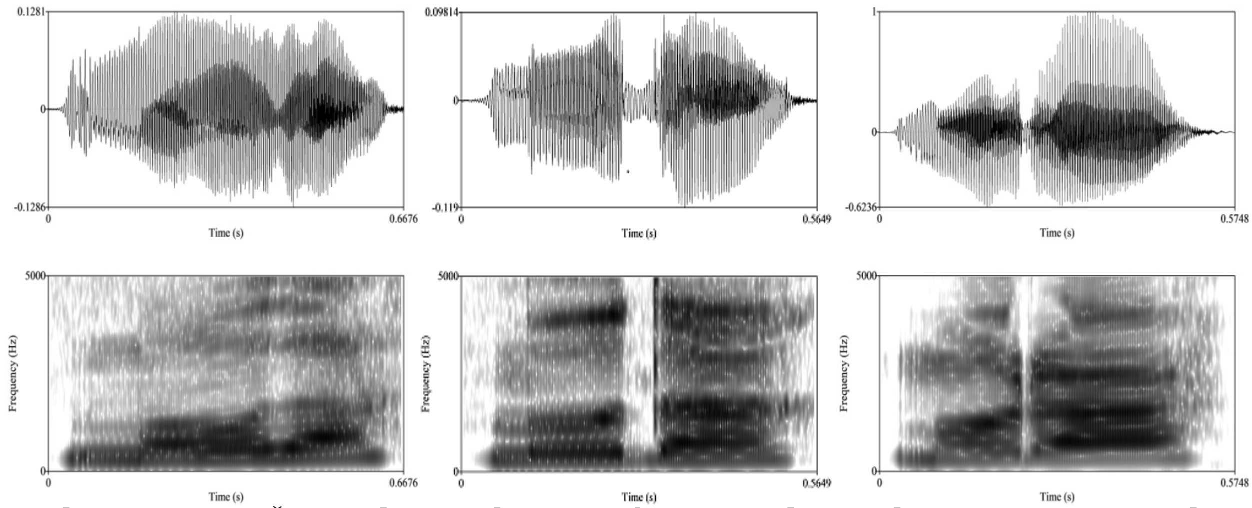
(a) **(b)** **(c)**



[m o ɾ a] [m o ɾ ɹ a] [m o ɹ a]

Figure 4.8. Waveforms and spectrograms of (a) *moda* with a native-like [ð], (b) a near-native [d], (c) and a non-native [ɾ].

(a) **(b)** **(c)**

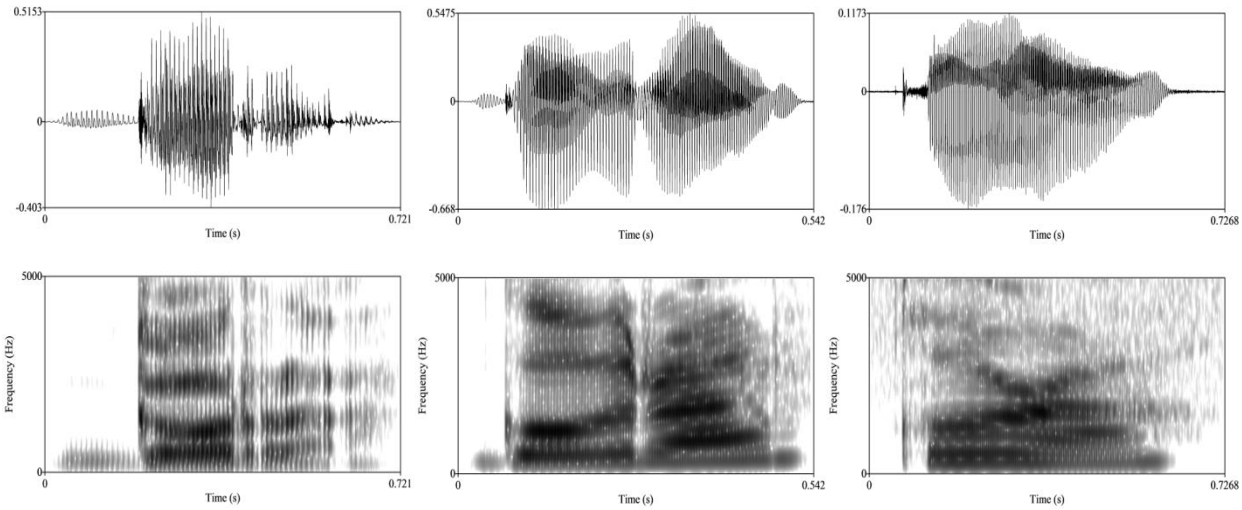


[m o ð̞ a] [m o d a] [m o ɾ a]

rarely produce intervocalic /d/ as a stop and a pronunciation like [moda] would sound very unnatural, this production was scored ‘0.5’. However, [d] cannot be confused with any other Spanish phonemes. On the other hand, producing the intervocalic /d/ as [ɾ], as in Figure 4.8c,

Figure 4.9. Waveforms and spectrograms of *gorra* (a) with a native-like [r], (b) a near-native [ɾ], (c) and a non-native [ɹ].

(a) **(b)** **(c)**



[g o r a] [g o ɾ a] [g o ɹ a]

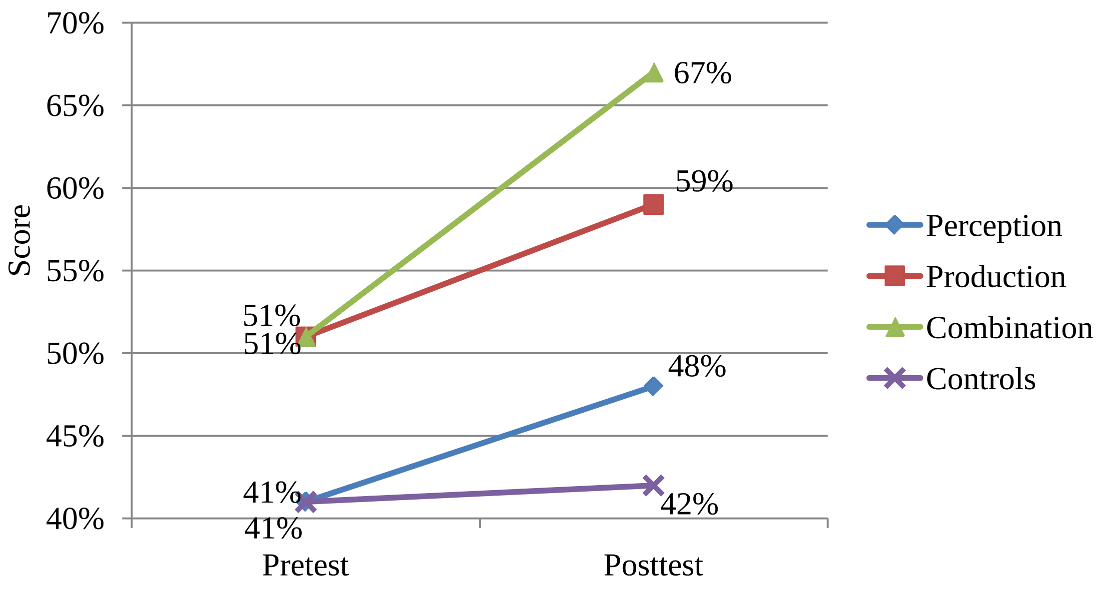
resulted in a ‘0’, because it involved replacing /d/ with another phoneme in Spanish and producing a different word.

The examples in Figure 4.9 above illustrate how productions were scored when the target was *gorra* ‘ball cap’. Productions of *gorra* with a /r/ consisting of two or more complete occlusions, as shown in Figure 4.9a, received a score of ‘1’. Replacing the /r/ with a [ɾ] as seen in Figure 4.9b, an error occasionally reported in the speech of native Spanish speakers, resulted in a ‘0.5’. However, producing the Spanish phoneme /r/ as the English [ɹ] was scored ‘0’.

4.2.2.1 Production Results

A 2 × 3 (Test × Contrast) repeated measures Analysis of Variance was conducted on the scores participants received based on the scoring system described above. Since the scores ranged from 0 – 1.0, they are described as percentages, with a higher percentage indicating a pronunciation nearer that of a native Spanish speaker. Group was used as a between-subjects factor. This analysis revealed a main effect of Test and an interaction between Test and Group. The main effect of Test indicates that participants’ productions of the three contrasts improved from pretest (46%) to posttest (54%) [$F(1, 38) = 37.290$, $p < 0.0001$]. More interestingly, the interaction between Test and Group, shown in Figure 4.10 below, reached significance because the training groups improved while the Controls did not [$F(1, 38) = 5.105$, $p = 0.005$]. Paired samples t-tests found that Perception Trainees made near-significant improvement overall [$t(9) = 2.043$, $p = 0.071$] and both Production [$t(9) = 3.157$, $p = 0.012$] and Combination Trainees [$t(10) = 6.640$, $p < 0.0001$] improved significantly while the Controls made no such improvement [$t(10) = 0.551$, $p = 0.594$].

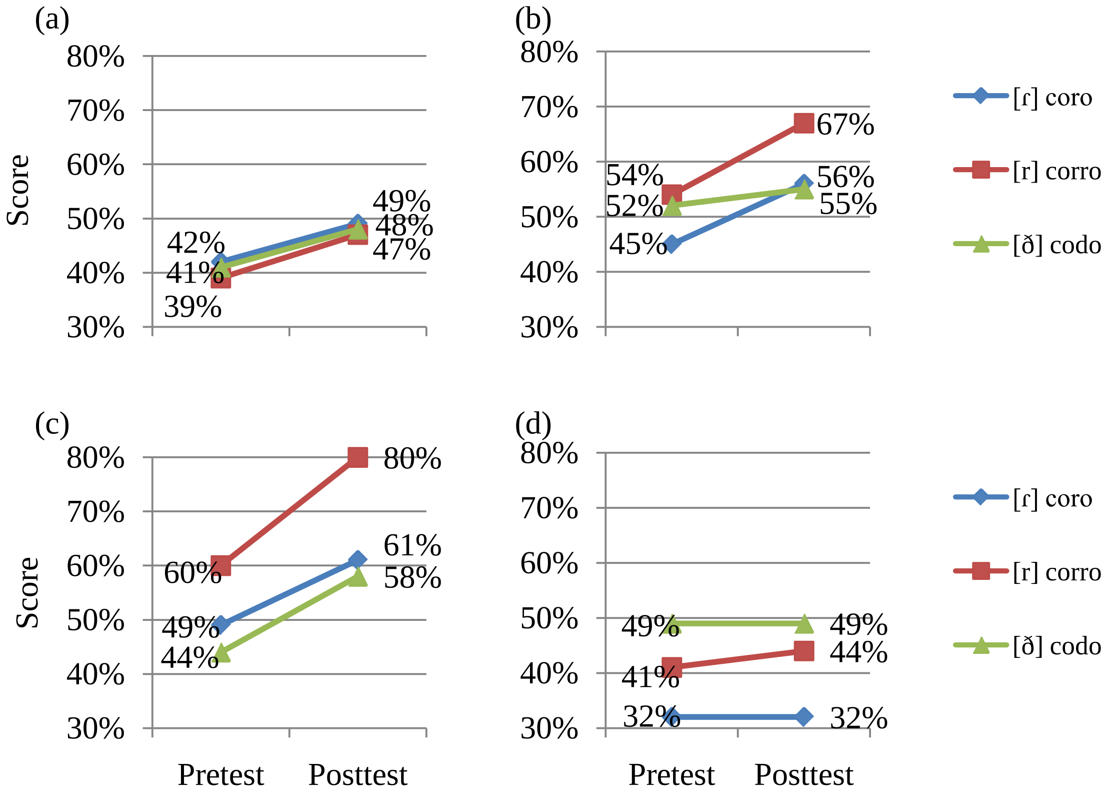
Figure 4.10. Acoustic analysis score of pretest and posttest Spanish learner productions by Group.



Although there was neither a main effect of Contrast [F (2, 76) = 2.343, p = 0.103] nor an interaction between Contrast and Group [F (6, 76) = 1.654, p = 0.144], paired samples t-tests comparing pretest and posttest productions by Group revealed that the training groups made significant or near-significant improvement on the production of some of the contrasts while the Controls did not. Specifically, Perception Trainees exhibited near-significant improvement in the pronunciation of the trill as illustrated below in Figure 4.11a [t (9) = 2.107, p = 0.064]. Production Trainees, whose pretest and posttest results are graphed in Figure 4.11b evidenced significant improvement in the pronunciation of the trill [t (9) = 2.421, p = 0.039]. The Combination Trainees, shown in Figure 4.11c, significantly improved their pronunciation of the tap [t (10) = 3.574, p = 0.005], the trill [t (10) = 3.287, p = 0.008], and the spirantized-/d/ [t (10) = 3.178, p = 0.010]. By contrast, the Controls did not improve in the pronunciation of any of the contrasts [t (10) < 1.0 for all contrasts].

In summary, based on visual inspection of the waveforms and spectrograms of all tokens produced by all speakers, the Production and Combination Trainees significantly

Figure 4.11. Acoustic analysis score of Spanish learner productions of (a) perception trainees, (b) production trainees, (c) combination trainees, and (d) controls from pretest to posttest by contrast.



improved their overall production, and the Perception Trainees made near-significant improvement overall. The Production and Combination Trainees both significantly improved their pronunciation of the trill /r/, and the Perception Trainees made near-significant improvement in the pronunciation of the trill /r/. Additionally, the Combination Trainees significantly improved their pronunciation of the tap /ɾ/ and the spirantized-/d/. However, the Controls exhibited no improvement from pretest to posttest overall or by contrast.

These results closely parallel those reported based on the native speaker identification task. Just as the Production and Perception Trainees made significant improvement and the

Combination Trainees made near-significant improvement according to the native speaker identification task, the Production and Combination Trainees made significant improvement and the Perception trainees made near-significant improvement based on the acoustic analysis. Similarly, according to the native speaker identification task, the Perception and Production Trainees significantly improved their trill productions and the Combination Trainees made near-significant improvement while, according to the acoustic analysis, the trill improvement of Combination and Production Trainees reached significance and that of the Perception Trainees neared significance.

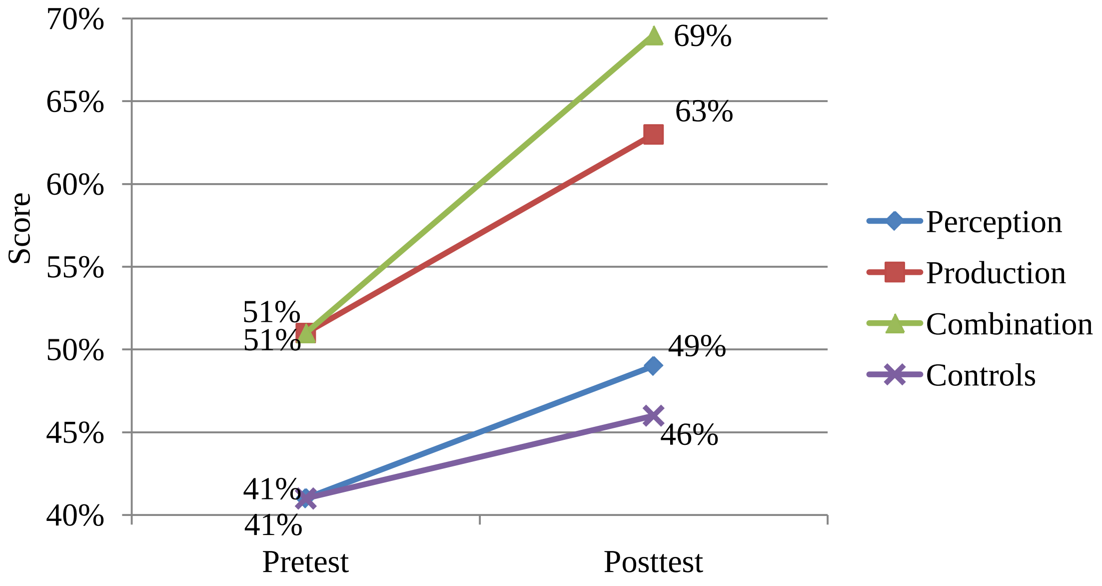
The largest discrepancy between the native speaker identification task results and the acoustic analysis results is the Combination Trainees' significant improvement in the pronunciation of /ɾ/ and /ð/. This discrepancy could be due to a ceiling effect, because the Spanish learners' pretest /ɾ/ and /ð/ productions were identified correctly 80% and 82%, respectively, in the native speaker identification task. Alternately, if the Spanish learners produced the /ɾ/ as [ɹ] and the /ð/ as [d], the two sounds would be identified correctly even if they were not produced in a native-like manner. In short, the acoustic analysis provides a more fine-grained measure of improvement of the /ɾ/ and /ð/ than the identification task.

4.2.2.2 Generalization Results

In order to see if the improvement in production reported above generalized to novel stimuli, a 2 × 3 (Test × Contrast) repeated measures Analysis of Variance was conducted on the acoustic analysis scores. Group was used as a between-subjects factor. The two levels of Test refer to the pretest and generalization test, and the three levels of Contrast refer to the /ɾ/, /r/, and /d/ contrasts. A main effect of Test reached significance, reflecting that overall pronunciation scores improved significantly from pretest (46%) to posttest (57%) [F (1, 38) =

82.313, p < 0.0001]. A Test by Group interaction also reached significance [F (3, 38) = 4.976, p = 0.005], indicating that the Combination Trainees improved more than the

Figure 4.12. Acoustic analysis score of pretest and generalization test Spanish learner productions by Group.



Production Trainees and that the Perception Trainees improved more than the Controls as illustrated above in Figure 4.12. However, paired samples t-tests found that all groups, including the Controls, improved significantly overall. The t-test results are detailed in Table 4.1.

Table 4.1. Paired samples t-test results showing Spanish learners' overall improvement from pretest to generalization test by Group.

Group	T-test Results
Perception Trainees	t (9) = 2.657, p = 0.026
Production Trainees	t (9) = 5.179, p = 0.001
Combination Trainees	t (10) = 9.837, p < 0.0001
Controls	t (10) = 3.085, p = 0.012

A Test by Contrast interaction also reached significance, meaning that the tap and trill productions improved more than the /ð/ productions as seen in Figure 4.13 below. According to paired samples t-tests, the tap [t (41) = 5.257, p < 0.0001], trill [t (41) = 5.926, p < 0.0001], and spirantized-/d/ [t (41) = 2.333, p = 0.025] all exhibited significant improvement. When the improvement of each contrast was analyzed by group, paired samples t-tests found that

Figure 4.13. Acoustic analysis score of Spanish learner productions from pretest to generalization test by Contrast.

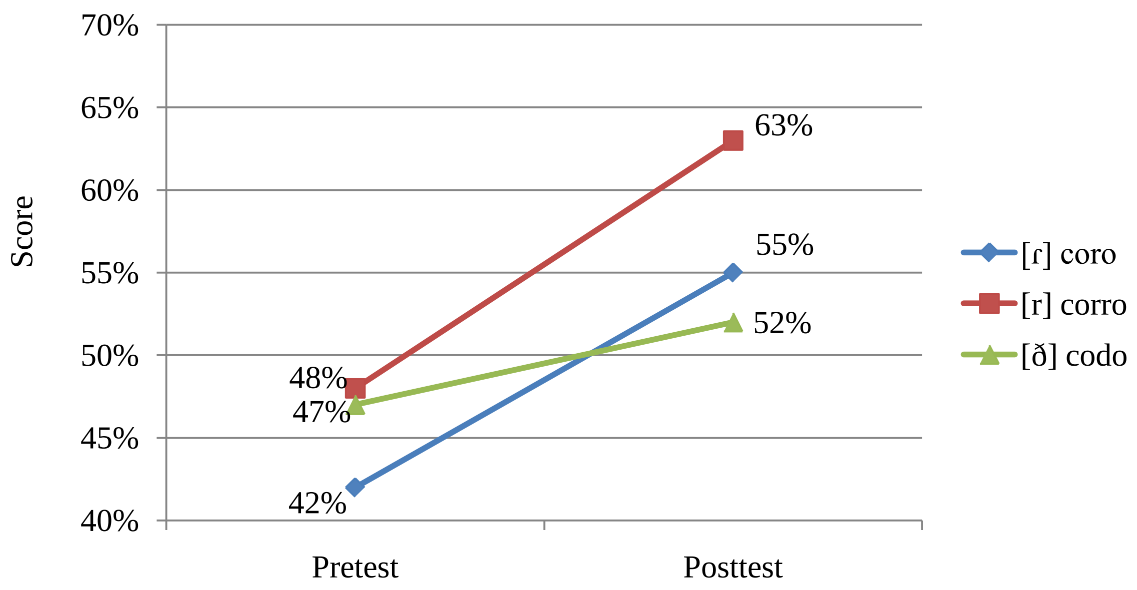
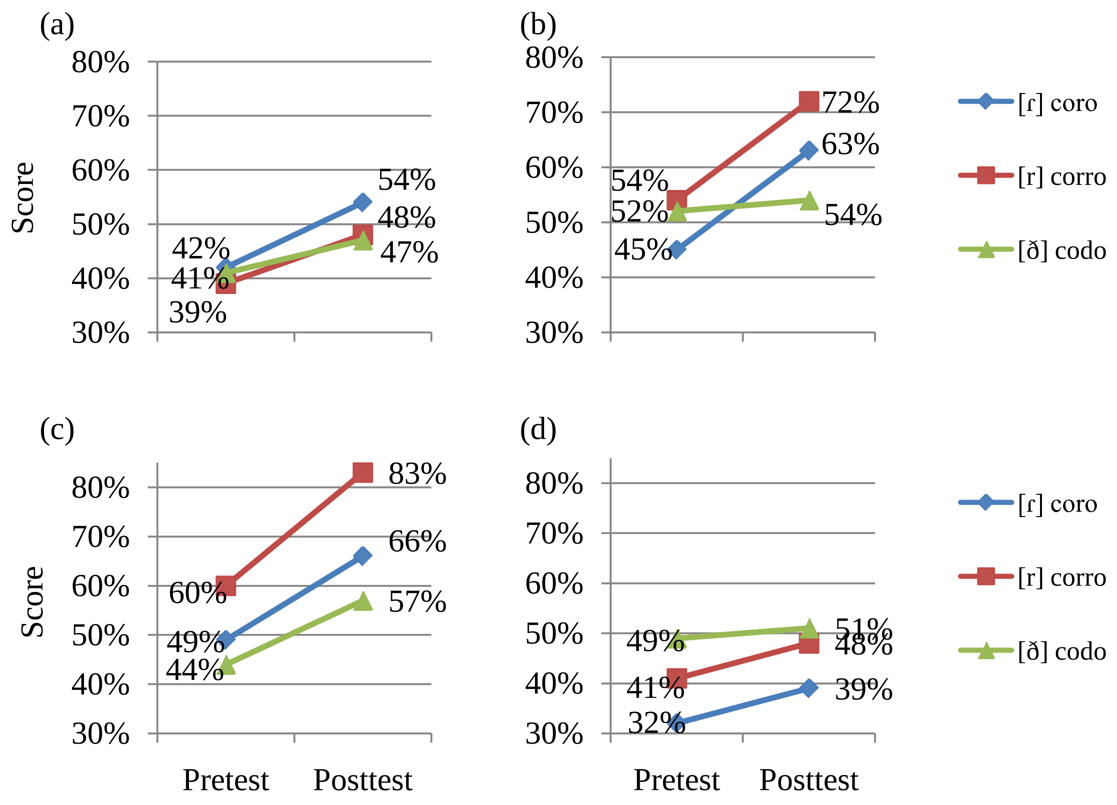


Table 4.2. Paired samples t-test results showing Spanish learners' improvement of the tap from pretest to generalization test by Group.

Group	/ɾ/ T-test Results
Perception Trainees	t (9) = 2.951, p = 0.016
Production Trainees	t (9) = 2.116p = 0.063
Combination Trainees	t (10) = 4.853, p = 0.001
Controls	t (10) = 3.085,p = 0.007

the improvement of the tap for all groups, including the Controls, reached or neared significance. Table 4.2 above summarizes the statistical findings. However, as can be seen below in Figure 4.14, only the Perception [t (9) = 3.632, p = 0.005], Production [t (9) = 3.299, p = 0.009], and Combination Trainees [t (10) = 4.115, p = 0.002] significantly improved their production of the trill. Likewise, only the Combination Trainees exhibited significant improvement in the pronunciation of the spirantized-/d/ [t (10) = 2.277, p = 0.046].

Figure 4.14. Acoustic analysis score of Spanish learner productions of (a) perception trainees, (b) production trainees, (c) combination trainees, and (d) controls from pretest to generalization test by contrast.



With respect to the Combination Trainees' significant improvement in the pronunciation of spirantized-/d/, as addressed above, these results and those of the acoustic analysis of the production results may differ from the results of the native speaker identification task due to a ceiling effect. The Spanish learners' productions of /d/ were identified correctly 82% at pretest.

Although the Controls significantly improved their pronunciation overall and their pronunciation of the /ɾ/, it is important to note that their improvement did not reach the same magnitude as that exhibited by the training groups. Also, while the trainees significantly improved their pronunciation of the trill, a result mirroring that reported above in the native speaker identification task and the acoustic analysis of the production posttest, the Controls exhibited no improvement in the pronunciation of the trill.

4.3 Cross-Modal Priming Task

While the perception and production tests discussed above measure whether the English learners of Spanish could perceive and produce the three-way contrast /ɾ, r, d/, they do not shed light on whether the trainees automatically processed the sounds differently. The cross-modal priming task uses reaction times in order to show whether trainees have encoded the new contrasts in their lexicons. For example, is *codo* 'elbow' now encoded as [koðo] rather than [kodo] or [koɾo] and is *coro* 'choir' encoded as [koɾo]? If these changes have occurred, trainees will activate *coro,* not *codo*, from their lexicon after hearing the auditory prime [koɾo], resulting in faster lexical decision reaction times for the visual target *coro* and slower for *codo.* Likewise, if trainees have categorized the tap and trill as separate phonemes and encoded these separately in words like *coro* and *corro* 'I run', reaction times will be slower in response to a visual target *corro* after an auditory prime [koɾo].

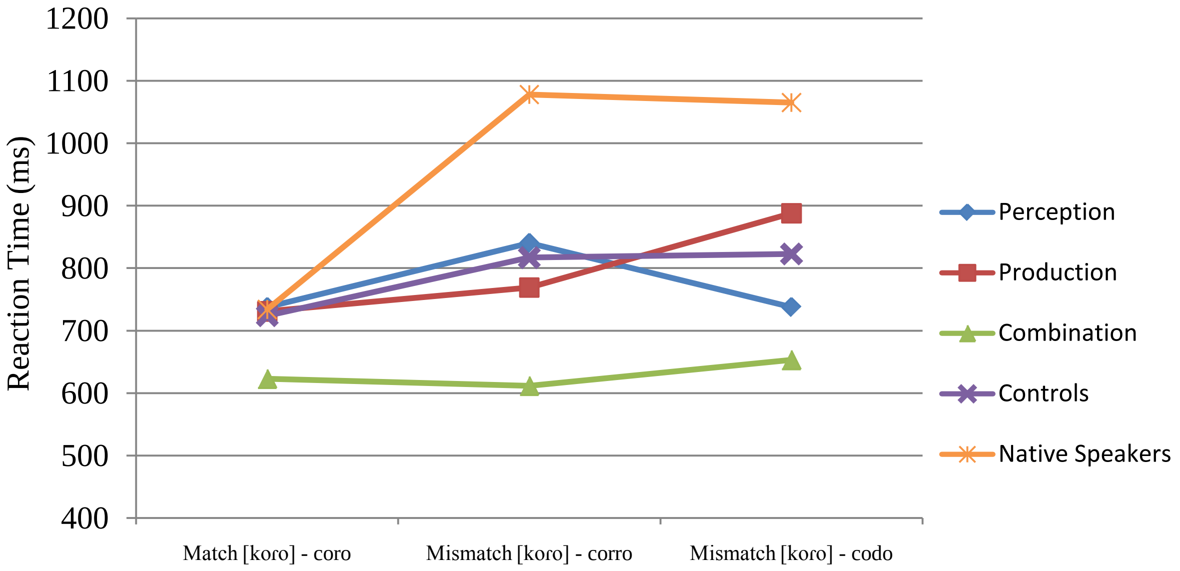
The analyses of reaction time that follow are based on correct responses, and the percentages of correct responses by Target Type are listed in Table 4.3. Correct responses with reaction times beyond two standard deviations below or above each participant's mean were also excluded. This resulted in a loss of 10.6% of the data and using 1030 of the 1152 available data points [(2 pre/post x 12 tokens x 42 Spanish learners = 1008) + (12 tokens x 12 native Spanish speakers = 144) = 1152].

Table 4.3. Mean accuracy rate of the matching and mismatching visual targets with standard deviations.

Target Type	Mean Accuracy (s.d.)
coro (Match)	96.7% (7)
corro (Mismatch)	85.1% (15)
codo (Mismatch)	89.6% (12)

To determine if participants responded with faster reaction times to the matching pairs (e.g., auditory prime [koɾo] followed by visual targer *coro*) than to mismatching pairs (e.g., auditory prime [koɾo] followed by visual target *corro* or *codo*) before training, one-way ANOVAs including the three Target Types were performed on reaction times of each group of participants. For the native Spanish speakers shown below in Figure 4.15, the main effect of Target Type proved significant [F (2, 33) = 3.886, p = 0.031]. Pairwise comparisons indicated that Spanish speakers responded more quickly to the matching targets (coro) than to mismatching 'rr' (*corro*, p = 0.019) and mismatching'd' (*codo*) targets (p = 0.024).

Figure 4.15. Cross-modal priming reaction times (ms) from the pretest.



However, the Spanish learners did not exhibit significant differences in reaction times between matching and mismatching targets. The Perception Trainees [$F(2, 27) = 0.737$, $p = 0.448$], Production Trainees [$F(2, 27) = 1.279$, $p = 0.295$], Combination Trainees [$F(2, 30) = 0.146$, $p = 0.865$], and Controls [$F(2, 30) = 0.340$, $p = 0.714$] responded with similar reaction times regardless of Target Type.

Another series of one-way ANOVAs including Target Type was conducted on the posttest cross-modal priming reaction times. Because the native Spanish speakers only completed the pre-cross-modal priming task, the results from Figure 4.15 are again reported here to illustrate the difference between native Spanish speakers and Spanish learners. As illustrated in Figure 4.16, training resulted in no changes in reaction time for the English learners of Spanish; trainees still exhibited similar reaction times to visual targets regardless of Target Type. Table 4.4 details the statistical results.

Figure 4.16. Cross-modal priming reaction times (ms) from the posttest.

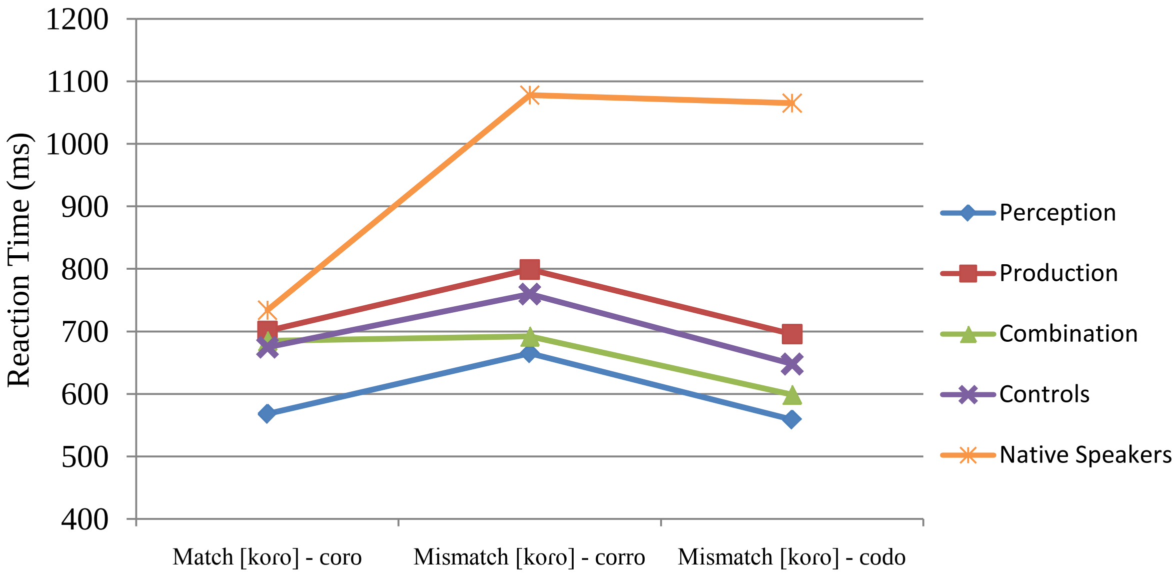


Table 4.4. Results of the one-way ANOVA conducted on the post-cross-modal priming task reaction times with Target Type as a factor.

Group	Statistical Results
Perception Trainees	F (2, 27) = 0.661, p = 0.550
Production Trainees	F (2, 27) = 1.249, p = 0.303
Combination Trainees	F (2, 30) = 0.669, p = 0.521
Controls	F (2, 30) = 1.105, p = 0.344

With respect to overall reaction time, the Spanish learners actually responded significantly more quickly overall (708 ms) than the native Spanish speakers (959 ms) according to an independent samples t-test [t (51) = 3.579, p = 0.001]. This could mean that the auditory prime [koɾo] actually primes all three visual targets (i.e., *coro, corro, codo)* for learners of Spanish. Alternately, it is likely that the native Spanish speakers' responses to mismatch targets were inhibited by the preceding auditory prime while the Spanish learners did not exhibit this inhibitory effect.

Prior to the cross-modal priming practice task, participants were instructed to hold both hands above the button box and to use both hands (i.e., left hand for leftmost button and right hand for rightmost button) during the lexical decision task in order to speed responses. Many of the Spanish learners replied that this was "good gaming strategy." By contrast, none of the native Spanish speakers made this comment, and most seemed less sure in their use of the button box during the practice period. It is unclear whether this was due to unfamiliarity with the task or unfamiliarity with a button-box-type controller. Perhaps the Spanish learners exhibited faster reaction times due to their online gaming experience.

Regardless of the reason, the key point is that learners of Spanish did not exhibit a reaction time pattern similar to that of native Spanish speakers either before or after training. However, these results do not necessarily rule out the possibility that trainees have recategorized their L1 allophone [ɾ] as a separate L2 phoneme and created a new phonemic category for the L2 /r/. Instead, this simply means that if restructuring of the allophonic and phonemic categories has taken place, this change has not yet been encoded in the learners' lexicons. Considering that fluent bilinguals who have spoken an L2 since childhood still exhibit priming differences when compared to L1 speakers (Pallier, Colomé, and Sebastián-Gallés 2001), the present results are not surprising. The MMN experiment should be able to provide evidence whether the /ɾ/ contrast has been recategorized.

4.4 MMN Task

The MMN task measured to what extent 9 native speakers of Spanish and 9 Perception Trainees exhibited pre-attentive responses to standard and deviant stimuli. Of the twenty [eɾe] and [ede] tokens repeated by speaker F1, the seven best repetitions of each were selected based on visual and auditory inspection using Praat (Boersma & Weenink 2005). The amplitude of all tokens was adjusted to 70 dB. Paired samples t-tests confirmed that the different token types ([eɾe] vs. [ede]) did not differ significantly in total duration, first vowel duration, amplitude, fundamental frequency, or vowel formants (F1, F2, F3). However, the difference in consonant duration [t (12) = -7.524, p = 0.084] and second vowel duration [t (12) = 0.026, p = 0.060] neared significance. These differences were unavoidable as /ɾ/ is a shorter consonant than /d/ and the total duration of the tokens was kept constant.

The experiment was presented in two blocks, with a 7:1 [ede] standard to [eɾe] deviant in block one and a 7:1 of [eɾe] standard to [ede] deviant in block two. The ISI averaged 500 ms and randomly ranged from 400 ms to 600 ms in 10 ms intervals in order to reduce the effect of overlap. Eight hundred stimuli were presented in each block, and the blocks were counter-balanced across participants. EEG data were collected at the scalp using 32 Ag/AgCl electrodes attached to an Electro-Cap and recorded using a Neuroscan Synamps 2 system. Six additional electrodes were attached above, below, and to the side of each eye to capture eye blinks (VEOG) and horizontal eye movement (HEOG). Two reference electrodes were also placed on the left and right mastoid. The impedance of the VEOG and HEOG electrodes was kept below 10 kilo-ohms and that of the scalp and mastoid electrodes was kept below 5 kilo-ohms. The EEG data was later re-referenced to the average of the left and right mastoids then filtered. Recordings were segmented into 600 ms epochs (100 ms preceding the uniqueness point of the auditory stimulus to 500 ms following the uniqueness point), and epochs were rejected due to artifacts if electrical activity exceeded a 120 μV threshold. More than fifty epochs from each condition were accepted for each participant.

To determine if native Spanish speakers, Perception Trainees at pretest, and Perception Trainees at posttest exhibit different MMN responses, the mean amplitude of electrical response to standard [ede] was compared to that of deviant [ede] for each group. Likewise the mean amplitude of electrical response to standard [eɾe] was compared to that of deviant [eɾe]. Since /ɾ/ and /d/ are separate phonemes in Spanish, the native Spanish participants were expected to exhibit a robust MMN response, with deviant stimuli eliciting a more negative response than standard stimuli. For the Perception Trainees, who categorize /ɾ/ is an allophone of /d/ in English, no MMN responses were expected at pretest. As the Perception Trainees significantly improved their perception and production according to the previous behavioral tasks, they were expected to exhibit MMN responses to deviant stimuli at posttest.

To test these predictions, paired samples t-tests were conducted on the amplitude of electrical activity in a time window from 100 – 200 ms as measured in microvolts (μV) at electrode sites over the left hemisphere (FP1, FC3, P3, C3, F3, FT7, T5, O1, F7, CP3, TP7, T3, as in Figure 4.17) for standard [ede] vs. deviant [ede] and standard [eɾe] vs. deviant [eɾe]. For the native Spanish speakers, the amplitude of response to deviant [ede] was significantly

Figure 4.17. Electrode site map for 32-channel Electro-Cap.

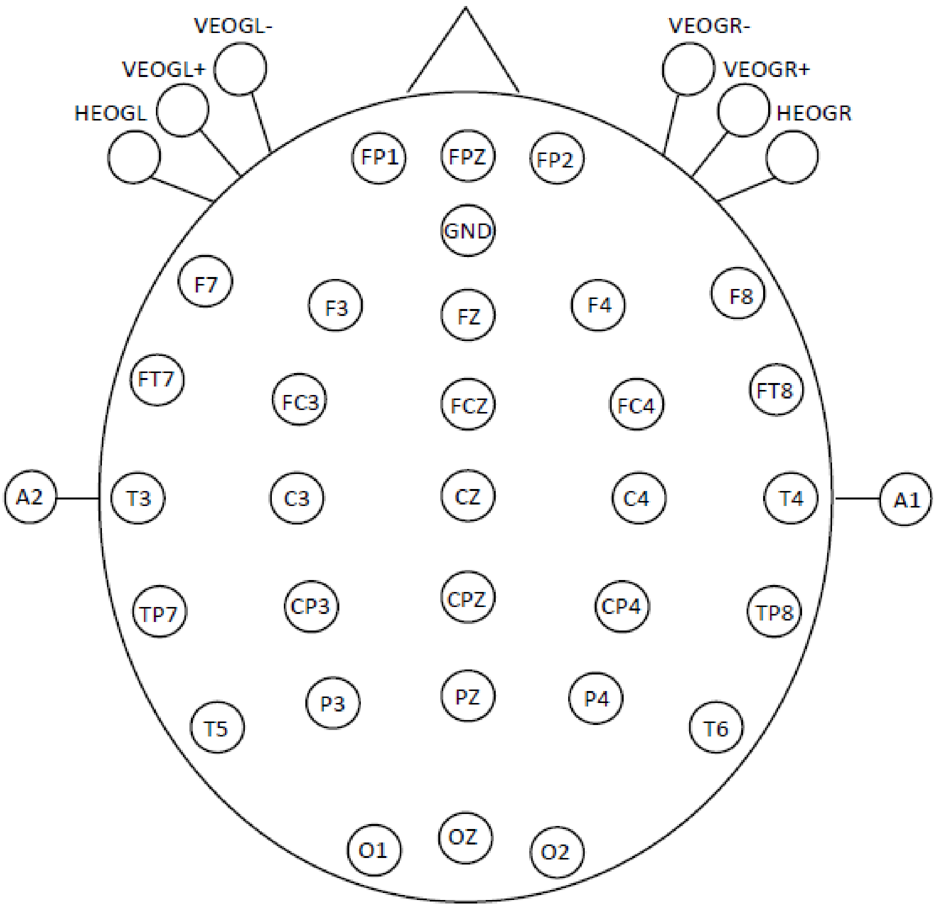
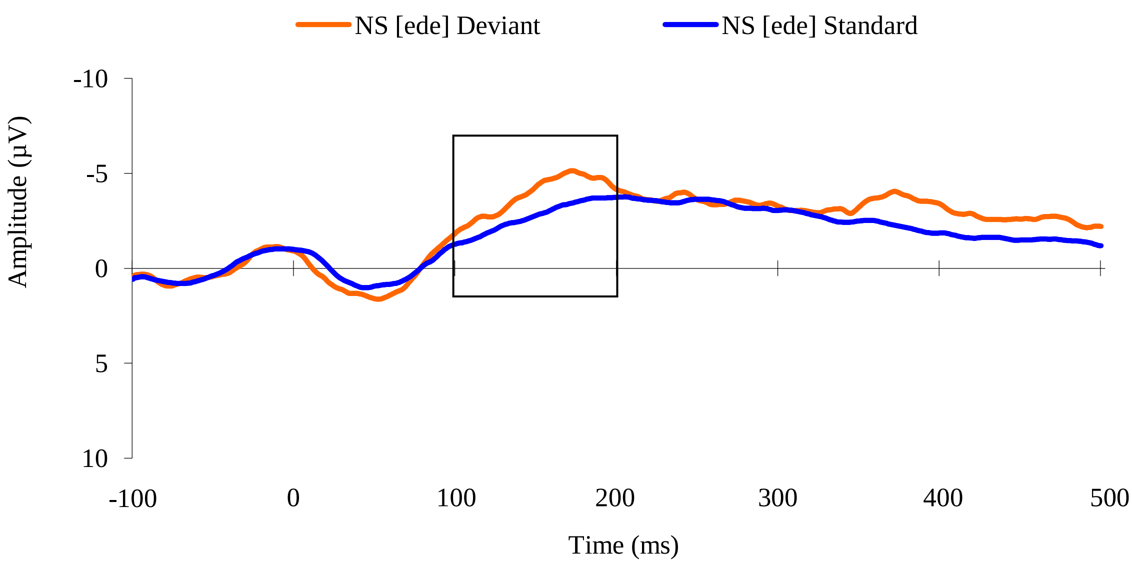
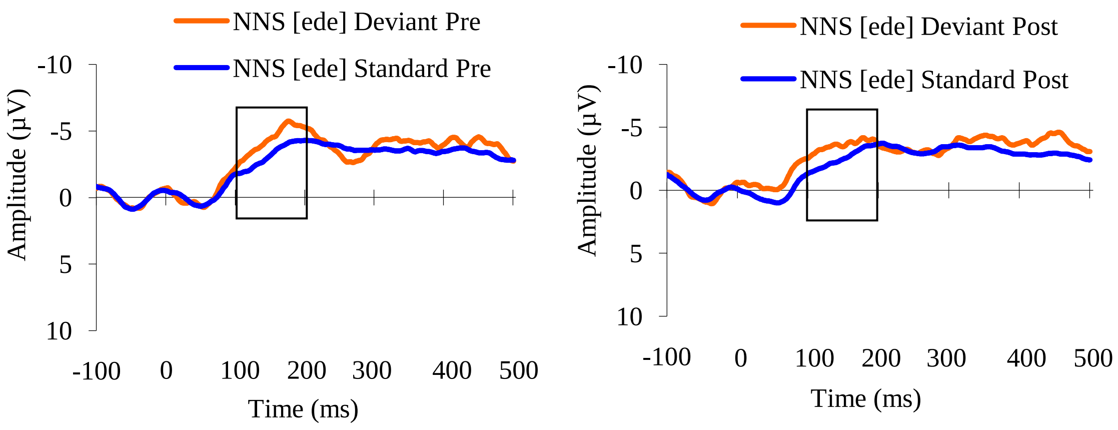


Figure 4.18. Grand-average waveform at F3 for native Spanish speakers presented with standard and deviant [ede].



more negative than that to standard [ede] as seen in Figure 4.18 above [t (8) = 2.784, p = 0.024]. While this effect was predicted, Perception Trainees also exhibited significantly more negative responses to deviant [ede] than standard [ede] at pretest [t (8) = 2.375, p = 0.045] and posttest [t (8) = 3.258, p = 0.012]. The left panel of Figure 4.19 illustrates that the Perception Trainees already exhibited an MMN response to deviant [ede] and that the response remained present at posttest. It is important to note here that [ɾ] is an allophone of

Figure 4.19. **Grand-average waveform at F3 for Perception Trainees presented with standard and deviant [ede] at pretest (on left) and posttest (on right).**



both /t/ and /d/ in English. It is possible that a deviant [ede] preceded by standard [eɾe] elicited an MMN response at pretest, because the English allophone [ɾ] is associated with more than one phonemic category.

Native Spanish speakers presented with deviant [eɾe] again resulted in a significantly more negative response than standard [eɾe] as predicted [t (8) = 3.594, p = 0.007]. This MMN response, which we expected since /ɾ/ and /d/ are separate phonemes in Spanish, can be seen in Figure 4.20. The Perception Trainees also performed as expected when presented with standard and deviant [eɾe] at pretest. As shown in the left panel Figure 4.21, Perception Trainees did not exhibit a significantly different response to deviant [eɾe] and standard [eɾe], indicating that /ɾ/ and /d/ were categorized as the same phoneme at pretest [t (8) = 0.125, p = 0.904]. Crucially, this was not the case at posttest. After training, Perception Trainees, like

Figure 4.20. Grand-average waveform at F3 for native Spanish speakers presented with standard and deviant [eɾe].

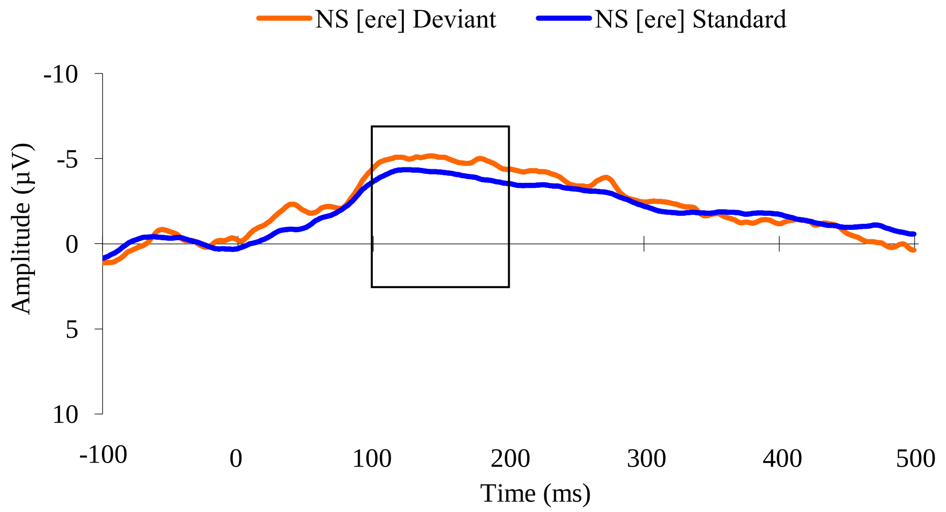
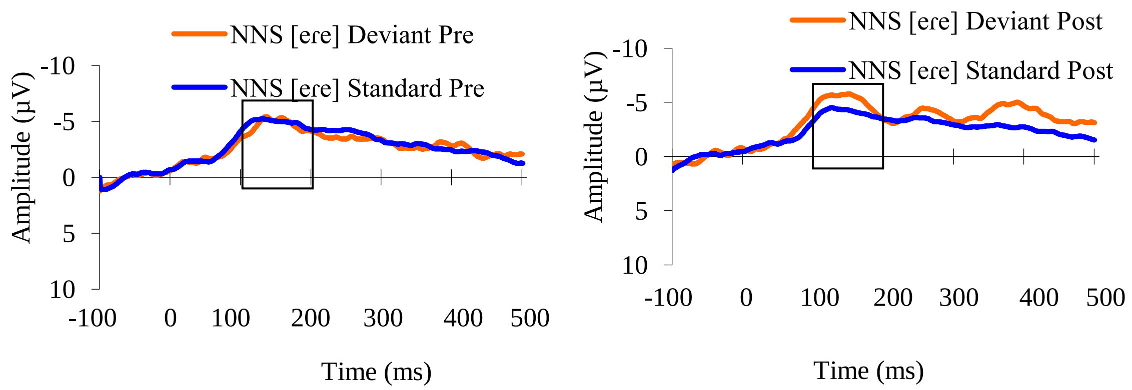


Figure 4.21. Grand-average waveform at F3 for Perception Trainees presented with standard and deviant [eɾe] at pretest (on left) and posttest (on right).



the native Spanish speakers, exhibited a significant MMN response when presented with deviant [eɾe] compared to standard [eɾe] [t (8) = 2.520, p = 0.036]. While [ede] and [eɾe] were categorized as the same phoneme at pretest, Perception Trainees appear to process these sounds differently at posttest.

Training not only improved their perception and production of Spanish contrasts, but it also resulted in changes in neural activation, suggesting that the Perception Trainees now categorize /ɾ/ and /d/ as separate phonemes. As to the asymmetry between pretest MMN responses to deviant [ede] and [eɾe], since /ɾ/ is an allophone of /t/ as well as /d/ in English, it is likely that [ete], like [ede], would also elicit an MMN response. Basically, since /ɾ/ belongs to two phonemic categories in English, participants are unable to build a mental representation of one phonemic category when presented with this allophone as the standard in the oddball paradigm. Returning to the consonant duration and second vowel duration differences in the MMN stimuli, these acoustic differences clearly did not have an effect on the MMN results. If this task were sensitive to these small acoustic differences, Perception Trainees would have shown an MMN response to deviant [eɾe] at pretest since the consonant duration and vowel duration were systematically different than those of the standard [ede]. Clearly, the changes in neural activation reported here are due to perception training.

Chapter 5: Conclusions and Discussion

This dissertation sought to tease apart the effects of perceptual and production training and to investigate whether native speakers of American English can be trained to perceive and produce the three-way /ɾ, r, d/ contrast in Spanish. Moreover, if trainees succeed in improving their perception and production of the contrast on behavioral tasks, this study further set out to examine the results of psycholinguistic and neurolinguistic tasks to evaluate whether trainees had developed automaticity in their perception of the L2 contrasts (MMN) and / or their perception of lexical items containing the contrasts (cross-modal priming).

The perception, production, and generalization results strongly indicate that all three training types (Perception, Production, and Combination) improve trainees' ability to perceive and / or produce the /ɾ, r, d/ contrast in Spanish. Whereas the Controls exhibited no improvement in the perception of the three-way contrast, all three training groups improved in the overall perception of the sounds. More specifically, the training groups significantly improved in their ability to distinguish the /ɾ, r/ contrast on both the perception posttest and the perception generalization tests, indicating that they can apply benefits of training to new contexts with new speakers, new dialects, and new words.

While training did not result in an improvement in the perception of the /ɾ, d/ contrast, the Perception Trainees maintained their pretest perception levels while the Production and Combination Trainees declined significantly in their perception of this contrast. One might assume that the /ɾ, r/ contrast evinced more improvement than the /ɾ, d/ because the participants perceived these contrasts with significantly different levels of accuracy on the pretest. However, the /ɾ, d/ contrast actually exhibited the lowest perception accuracy, allowing more room for improvement. Since the /ɾ, d/ contrast appeared to be more resistant to training than the /ɾ, r/, it leads one to ask how the two pairs of sounds differ.

As mentioned above, in the case of acquiring the /ɾ, r/ contrast, participants simply have to distinguish the /ɾ/, a familiar sound for English speakers, from the /r/, an unfamiliar but distinctive L2 sound. In the case of the /ɾ, d/ contrast, however, the participants have to distinguish the /ɾ/, an allophone of /d/ in English, from a Spanish L2 /d/ variant. This appears to be a much more difficult process as predicted in Best's Perceptual Assimilation Model (PAM), which predicts that separating two allophonic variants of the same phoneme into different categories will prove difficult (Best, McRoberts and Goodell 2001). However, PAM also predicts that participants will experience less difficulty when they are distinguishing two L2 sounds that map onto different sounds in their L1, which should also be the case with the pronunciations of *cada* [kaɾa] 'each' and *cara* [kaða] 'face'. Trainees should have been able to map [ɾ] and [ð] onto different categories (/d/ and /ð/) in their L1, facilitating the acquisition of the intervocalic /ɾ, d/ contrast. Clearly, the present study shows the /ɾ, d/ contrast's resistance to training.

In addition to directly comparing perception and production training, this study also verified that native speakers of American English experience difficulties acquiring the ability to perceive the /ɾ, r, d/ contrast in Spanish. While previous research has documented the pronunciation difficulties of English learners of Spanish (Face 2006, Waltmunson 2005, and Zampini 1993, 1994, 1997) and the perceptual confusion for monolingual speakers of English (Boomershine et al. 2008), no prior research has documented English learners of Spanish experiencing difficulties with the perception of the /ɾ, r, d/ contrast. The perception pretest of this dissertation establishes that English learners of Spanish experience difficulties distinguishing both the /ɾ, r/ (81% at pretest) and /ɾ, d/ (70% at pretest) contrasts.

Turning to production, the Controls again evinced no overall improvement while all trainees' productions improved significantly overall. The Production Trainees significantly improved their production of the trill /r/ based on native Spanish speaker evaluations, native

Spanish speaker ratings, and acoustic analysis. The Perception Trainees significantly improved their pronunciation of the trill /r/ according to native Spanish speaker identification and made near-significant trill /r/ improvement according to acoustic analysis. Similarly, the Combination Trainees made near-significant improvement on their pronunciation of the trill /r/ based on native Spanish speaker identifications and significant improvement based on acoustic analysis. While all groups, including the Controls, improved their pronunciation overall on the Generalization Test, only the three training groups exhibited improvement on a specific contrast (/r/), and the Controls did not exhibit the same magnitude of improvement seen in the trainees.

The highly significant improvement of the trill by the training groups combined with the lack of /ɾ, d/ improvement appears to support Flege's (1995) hypothesis that novel L2 sounds are easier to acquire than sounds that are similar to phonemes in the learners' L1. Flege's Speech Learning Model (1995) can also capture the reason the /ɾ, r/ contrast exhibited significant improvement in perception while the /ɾ, d/ did not. When distinguishing the former, learners are distinguishing a familiar sound (/ɾ/) from a novel phoneme (/r/), a relatively easy task. In the latter case, though, learners are distinguishing sounds that are similar to sounds that already exist in English. It seems unlikely, however, that learners would have difficulty distinguishing a Spanish-like [ð̞] in *cada* [kaða] 'each' from a Spanish-like [ɾ] in *cara* [kaɾa] 'cara' when they easily distinguish words like *though* [ðoʊ] and *dough* [doʊ] in English. It is not the case that learners must distinguish the Spanish [ð̞] from the English [ð], which would be more difficult and more like the French /u, y/ contrast Flege's uses as an example. Again, the participants' difficulties acquiring the /ɾ, d/ contrast seems to call for another explanation.

Although the acoustic analysis revealed that the Combination Trainees made significant improvement in their pronunciation of the /d/ and /ɾ/, these results did not translate into higher identification rates or rating scores by the native Spanish-speaking judges. As mentioned in Chapter 4, this lack of effect on the native Spanish speaker identification task may be due to the high pretest scores exhibited by all participants. Since the learners often mispronounced the /d/ as [d] and the /ɾ/ as [ɹ], it would have been easy for native Spanish speakers to identify which sounds the learners of Spanish intended. As for the native Spanish speaker ratings, the eight native Spanish judges first completed an identification task which consisted of 3 minimal pairs x 2 productions (pre and post) x 42 learners of Spanish, or 504 words. After completing this identification task, they were presented with a rating task for the same 504 words. This very well could have caused fatigue, resulting in inconclusive native Spanish speaker ratings.

Having reviewed the perception and production findings, one can now evaluate which of the three training types proved most effective and which of the two, production or perception, transferred better to the other modality. As mentioned above, all three training types proved highly effective; however, since the Production and Combination Trainees actually declined significantly in their perception of the /ɾ, d/ contrast while the Perception Trainees maintained their pretest levels of perception, Perception Training proved the most effective. Likewise, Perception Training transferred better to production, resulting in near-significant production improvement overall according to acoustic measures and significant (native speaker identification) or near-significant (acoustic analyses) improvement in the production of the trill /r/. Crucially, Perception Trainees exhibited no decline in the production of any contrast.

With the addition of the production data, the question as to why the /ɾ, d/ contrast proved more resistant to training than the /ɾ, r/ contrast arises again. Similarly, the production of the /ɾ/ and the /d/ also seemed resistant to training. Even though this lack of improvement could be due to a ceiling effect as mentioned above, another explanation is possible. In her dissertation, Swanson (2007) investigated whether it was more difficult for French learners of English to acquire aspiration and palatalization in L2 English or for English learners of French to suppress these rules in L2 French. In her study, which focused on production, she found that acquiring an L2 phonological process was more difficult than suppressing one from the L1. Although both suppressing and acquiring a phonological process could be trained, acquiring a phonological process required time to develop. This predicts that English learners of Spanish would have an easier time suppressing English flapping when producing /d/ in L2 Spanish than acquiring spirantization, which is exactly what occurred. Native Spanish speakers could easily identify the /d/ productions of the Spanish learners, because they were produced as [d]. Only 23% of the Spanish learners' pretest and posttest /d/ productions were produced as flaps, compared to 94% - 99% of the /d/ productions reported in English (Patterson and Connine 2001; Connine 2004; Zue and Laferriere 1979; Byrd 1994; and Herd, Jongman, and Sereno 2010). This supports Swanson's claim that it is easier to suppress an L1 phonological process (English flapping) than to acquire an L2 phonological process (Spanish spirantization).

However, the opposite appears true of perception. Although participants did not improve in their perception of the /ɾ, d/ contrast, they demonstrated no difficulty in perceiving the Spanish productions containing [ð̞] as /d/. Instead, their errors involved perceiving Spanish productions containing [ɾ] as /d/. When participants heard a [ð], they perceived it as a /d/ 94% on pretest and posttest, but when they heard a [ɾ], they also perceived that as a /d/ 59% on pretest and posttest. In other words, they were unable to suppress English flapping when they were performing a perception task. They categorized /ɾ/ as /d/ above chance levels, but they had no difficulty categorizing [ð̞] as /d/, or acquiring spirantizaion for the purposes of the identification task. Although these results are based on a forced-choice paradigm, the possibility that the ease of acquiring or suppressing a phonological process could vary with modality warrants further research.

Although the cross-modal priming task did not elicit any reaction time changes in the trainees from pretest to posttest, it did confirm the way native Spanish speakers process the /ɾ, r, d/ contrast. When an auditory prime containing a tap /ɾ/ was followed by the orthographic representation of that word with a 'r' (cross-modal repetition priming), native Spanish speakers responded significantly more quickly in a lexical decision task than when the prime was followed by the orthographic representation of a trill 'r' or a 'd'. While this result was expected, it had not been documented in Spanish with these sounds.

The MMN task greatly contributed to this work in several ways. First, the native Spanish speakers' results established that /d/ and /ɾ/, which are separate phonemes in Spanish, yield an MMN in our experimental paradigm, and that canonical MMN responses can be elicited using naturally produced stimuli. Second, the emergence of an MMN response in the Perception Trainees, the most successful of the training groups, clearly indicates that neural plasticity still exists in adult learners and that it is possible to recategorize L1 contrasts when learning an L2. The inability of adult L2 learners to overcome their foreign accents led to the strongest version of the Critical Period Hypothesis, the idea that acquiring new contrasts in an L2 is impossible after a critical age, like puberty. This research shows that intensive input, like the type presented during perception and production training, can lead to improved perception and production of new L2 contrasts as well as a reorganization of the L1 phonemic categories. Most importantly, this acquisition and emergence of neural activation can occur regardless of a learner's age.

References

Best, C. 1991. The emergence of native-language phonological influences in infants: A Perceptual Assimilation Model. *Haskins laboratories status report on speech research* 108, 1 – 30.

Best, C., McRoberts,G., & Goodell, E. 2001. Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America* 109:2, 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.

Bluhme, H., & Burr, R. 1971. An audio-visual display of pitch for teaching Chinese tones. *Studies in Linguistics* 22, 51 – 57.

Boersma, P., & Weenink, D. 2005. Praat: Doing phonetics by computer (Version 4.3.12) [Computer Program]. Retrieved May 11, 2005, from http://www.praat.org/.

Bongaerts, T., van Summeren, C., Planken, B. & Schils, E. 1997. Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition* 19, 447 – 465.

Boomershine, A., Hall, K.C., Hume, E., & Johnson, K. 2008. The impact of allophony versus contrast on speech perception. In *Contrasts in phonology: Theory, perception, acquisition,* eds. P. Avery, B. Dresher, and K. Rice, 145 – 171.

Bradlow, A., Pisoni, D., Akahane-Yamada, R., & Tohkura, Y. 1997. Training Japanese listeners to identify English /r/ and /l/: Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America* 101, 2299 – 2310.

Bradlow, A., Akahane-Yamada, R., Pisoni, D., & Tohkura, Y. 1999. Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception and Psychophysics* 61. 977 – 985.

Byrd, D. 1994. Relations of sex and dialect to reduction. *Speech Communication* 15, 39 – 54.

Connine, C. 2004. It's not what you hear but how often you hear it: On the neglected role of phonological variant frequency in auditory word recognition. *Psychonomic Bulletin & Review* 11:6, 1084 – 1089.

Davies, M. 2002. *Corpus del Español*. [Online Corpus]. Retrieved March 15, 2011, from http://www.corpusdelespanol.org/.

Díaz-Campos, M. 2004. Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition* 26, 249 – 273.

Face, T. L. 2006. Intervocalic Rhotic Pronunciation by Adult Learners of Spanish as a Second Language. In *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*, eds. Carol A. Klee and Timothy L. Face, 47-58. Somerville, MA: Cascadilla Proceedings Project.

Flege, J. 1995. Second language speech learning: Theory, findings, and problems. In *Speech perception and linguistics experience: Issues in cross-language research,* ed. Winifred Strange, 233 – 277. Timonium, MD: York Press, 1995.

Flege, J., Munro, M., & MacKay, I. 1995. Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America* 97:5, 3125 – 3134.

Grainger, J., Kang, M., & Segui, J. 2001. Cross-modal repetition priming of heterographic homophones. *Memory and Cognition* 29, 53 – 61.

Herd, W., Jongman, A., & Sereno, J. 2010. An acoustic and perceptual analysis of /t/ and /d/ flaps In American English. *Journal of Phonetics* 38:4, 504 – 16.

Hirata, Y. 2004. Computer assisted pronunciation training for native English speakers learning Japanese pitch and durational contrasts. *Computer Assisted Language Learning* 17, 357 – 376.

Ioup, G., Boustagui, E., El Tigi, M., & Moselle, M. 1994. Reexamining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition* 16, 73 – 98.

Jamieson, D., & Morosan, D. 1986. Training non-native speech contrasts in adults: Acquisition of the English /T/-/D/ contrast by Francophones. *Perception and Psychophysics* 40, 205 – 215.

Johnson, J. & Newport, E. 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, 60 – 99.

Kazanina, N., Phillips, C., & Idsardi, W. 2006. The influence of meaning on the perception of speech sounds. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 11381 – 11386.

Lenneberg, E. 1967. *Biological Foundations of Language*. New York: Wiley.

Logan, J., Lively, S., & Pisoni, D. 1991. Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America* 89, 874 – 886.

Lord, G. 2005. (How) Can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania* 88, 557 – 567.

MacKay, I., Flege, J., Piske, T., & Schirru, C. 2001. Category restructuring during second-language speech acquisition. *Journal of the Acoustical Society of America* 110 (1), 516 – 528.

Marslen-Wilson, W. 1990. Activation, competition, and frequency in lexical access. In *Cognitive Models of Speech Processing: Psycholinguistics and Computational Perspectives*, ed. G. Altmann, 148 – 172. Cambridge, MA: MIT Press.

Marslen-Wilson, W., & Zwitserlood, P. 1989. Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance* 15, 576 – 585.

Moyer, A. 1999. Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition* 21, 81 – 108.

Munro, M. & Mann, V. 2005. Age of immersion as a predictor of foreign accent. *Applied Psycholinguistics* 26, 311 – 341.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Iimoniemi, R., Luuk, A, Allik, J., Sinkkonen, J., & Alho, K. 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385, 432 – 434.

Oldfield, R. C. 1971. The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia* 9, 97 – 113.

Oyama, S. 1976. A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research* 5 (3), 261 – 283.

Pallier, C., Colomé, A. & Sebastián-Gallés. 2001. The influence of native-language phonology on lexical acces: Exemplar-based versus abstract lexical entries. *Psychological Science* 12 (6), 445 – 449.

Patterson, D. & Connine, C. M. 2001. Variant frequency in flap production: a corpus analysis of variant frequency in American English flap production. *Phonetica* 58, 254 – 275.

Peltola, M. S., Kujala, T., Tuomainen, J., Ek, M., Aaltonen, O., & Näätänen, R. 2003. Native and foreign vowel discrimination as indexed by the mismatch negativity (MMN) response. *Neuroscience Letters*, 352, 25 – 28.

Peltola, M. S., Kuntola, M., Tamminen, H.,Hämäläinen, H., & Aaltonen,O. 2005. Early exposure to non-native language alters preattentive vowels discrimination. *Neuroscience Letters* 388, 121 – 125.

Peltola, M.S., Tuomainen, O., Kroskinen, M., & Aaltonen, O. 2007. The effect of language immersion education on the preattentive perception of native and non-native vowel contrasts. *Journal of Psycholinguistic Research* 36, 15 – 23.

Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., McGinnis, M., & Roberts, T. 2000. Auditory cortex accesses phonological categories: An MEG mismatch study. *Journal of Cognitive Neuroscience* 12 (6), 1038 – 1055.

Quilis, A. 1993. *Tratado de fonología y fonética españolas.* Madrid: Gredos.

Seliger, H., Krashen, S., & Ladefoged, P. 1975. Maturational constraints in the acquisition of second language accent. *Language Sciences* 36, 20 – 22.

Snow, C. & Hoefnagel-Höhle, M. 1978. The critical period for language acquisition: Evidence from second language learning. *Child Development* 49 (4), 1114 – 1128.

Strange, W., & Dittmann, S. 1984. Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception and Psychophysics* 36, 131 – 145.

Swanson, K. 2007. *Acquisition versus suppression of phonological processes in the second language acquisition of French and English*. Ann Arbor, MI: UMI.

Tabossi, P. 1996. Cross-modal semantic priming. Language and Cognitive Processes 11, 569 – 76.

Tagliaferri, B. 2008. Paradigm: Perception Research Systems [Computer Program]. Retrieved March 23, 2008, from http://www.perceptionresearchsystems.com/.

Tremblay, K., Kraus, N., Carrell, T.D., & McGee, T. (1997) Central auditory system plasticity: Generalization to novel stimuli following listening training. *Journal of the Acoustical Society of America* 102, 3762 – 3773.

Tremblay, K., Kraus, N., & McGee, T. 1998. The time course of auditory perceptual learning: Neurophysiological changes during speech-sound training. *NeuroReport* 9, 3557 – 3560.

Waltmunson, J. 2005. *The relative difficulty of L2 Spanish /d,t/, trill, and tap by L1 English speakers: Auditory and acoustic methods of defining pronunciation accuracy*. Ann Arbor, MI: UMI.

Wang, Y., Jongman, A., & Sereno, J. 2003. Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America* 113, 1033 – 1043.

Wang, Y., Sereno, J., Jongman, A., & Hirsch, J. 2003. fMRI evidence for cortical modification during learning of Mandarin lexical tone. *Journal of Cognitive Neuroscience* 15 (7), 1019 – 1027.

Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., Czigler, I., Csépe, V., Ilmoniemi, R., & Näätänen, R. 1999. Brain responses reveal the learning of foreign language phonemes. *Psychophysiology* 36, 638 – 642.

Winkler, I., Kujala, T., Alku, P., & Näätänen, R. 2003. Language context and phonetic change detection. *Cognitive Brain Research* 17, 833 – 844.

Zampini, M. 1993. Spanish voiced stop phonemes and spirantization: A study in second language acquisition. Ann Arbor, MI: UMI.

Zampini, M. 1994. The role of native language transfer and task formality in the acquisition of Spanish spirantization. *Hispania* 77, 470 – 481.

Zampini, M. 1997. L2 Spanish Spirantization, prosodic domains and interlanguage rules. In *Focus on Phonological Acquisition*, eds. S. J. Hannahs and Martha Young-Scholten, 209 – 234.

Zhang, Y., Kuhl, P., Imada, T., Iverson, P., Pruitt, J., Stevens, E., Kawakatsu, M., Tohkura, Y., & Nemoto, I. 2009. Neural signatures of phonetic learning in adulthood: A magnetoencephalography study. *NeuroImage* 46, 226 – 240.

Zue, V. W. & Laferriere, M. 1979. Acoustic study of medial /t,d/ in American English. *Journal of the Acoustical Society of America* 66, 1039 – 1050.

Appendix A: Dialect Questionnaire (Native Spanish Speakers)

Questionnaire

Gender: __________

Age: __________

Native country/state ________________

Native language ______________

Knowledge of OTHER languages / dialects: Write the name of the language in the blank, and indicate your approximate abilities in each of the four areas for each language.

1. Language: ________________________

Speaking	*Listening*	*Reading*	*Writing*
□ Poor	□ Poor	□ Poor	□ Poor
□ Fair	□ Fair	□ Fair	□ Fair
□ Good	□ Good	□ Good	□ Good
□ Near-Native	□ Near-Native	□ Near-Native	□ Near-Native

2. Language: ________________________

Speaking	*Listening*	*Reading*	*Writing*
□ Poor	□ Poor	□ Poor	□ Poor
□ Fair	□ Fair	□ Fair	□ Fair
□ Good	□ Good	□ Good	□ Good
□ Near-Native	□ Near-Native	□ Near-Native	□ Near-Native

3. Language: ________________________

Speaking	*Listening*	*Reading*	*Writing*
□ Poor	□ Poor	□ Poor	□ Poor
□ Fair	□ Fair	□ Fair	□ Fair
□ Good	□ Good	□ Good	□ Good
□ Near-Native	□ Near-Native	□ Near-Native	□ Near-Native

This section is for native Spanish speakers.

1. What was your age when you started learning English? _______________

2. Did you take English in Elementary School? Yes No

   If yes, where? ___________________

   For how many years? ___________________

3. Did you take English in High School? Yes No

   If yes, where? ____________________

   For how many years? ____________________

4. Did you study English at the college level? Yes No

   If yes, where? ___________________

   For how many years? _____________________

5. Have you lived in an English-speaking country (including your current stay in the United States)? Yes No

   If yes, where and when? ____

   For how many years? ____

6. Have you had any informal, out of classroom, exposure to English? Yes No
   → If yes, please mark all exposure you have had.

   --------- Music in English.
   --------- English-speaking relatives (Relationship _______________).
   --------- English-speaking friends.
   --------- Vacation travel to English-speaking country.
   --------- English language magazines/ newspapers.
   --------- English language TV.

7. Do you have a foreign accent in English? Yes No
   → If yes, please rate the strength of your accent.

   □ No Accent □ Slight Accent □ Moderate Accent □ Strong Accent

Thank you for your participation.

Appendix B: Dialect Questionnaire (Native English Speakers)

Questionnaire

Gender: __________

Age: __________

Native country/state ________________

Native language ________________

Knowledge of OTHER languages / dialects: Write the name of the language in the blank, and indicate your approximate abilities in each of the four areas for each language.

1. Language: _________________________

Speaking	*Listening*	*Reading*	*Writing*
□ Poor	□ Poor	□ Poor	□ Poor
□ Fair	□ Fair	□ Fair	□ Fair
□ Good	□ Good	□ Good	□ Good
□ Near-Native	□ Near-Native	□ Near-Native	□ Near-Native

2. Language: _________________________

Speaking	*Listening*	*Reading*	*Writing*
□ Poor	□ Poor	□ Poor	□ Poor
□ Fair	□ Fair	□ Fair	□ Fair
□ Good	□ Good	□ Good	□ Good
□ Near-Native	□ Near-Native	□ Near-Native	□ Near-Native

3. Language: _________________________

Speaking	*Listening*	*Reading*	*Writing*
□ Poor	□ Poor	□ Poor	□ Poor
□ Fair	□ Fair	□ Fair	□ Fair
□ Good	□ Good	□ Good	□ Good
□ Near-Native	□ Near-Native	□ Near-Native	□ Near-Native

This section is for native English speakers.

1. What was your age when you started learning Spanish? _______________

2. Did you take Spanish in Elementary School? Yes No

   If yes, where? ___________________

   For how many years? ___________________

3. Did you take Spanish in High School? Yes No

   If yes, where? ____________________

   For how many years? ____________________

4. Did you study Spanish at the college level? Yes No

   If yes, where? ___________________

   For how many years? _____________________

5. Have you lived in a Spanish-speaking country? Yes No

   If yes, where? ________________

   For how many years? _____________________

6. Have you had any informal, out of classroom, exposure to Spanish? Yes No
   → If yes, please mark all exposure you have had.

   --------- Music in Spanish.
   --------- Spanish-speaking relatives (Relationship: ______________)
   --------- Spanish-speaking friends.
   --------- Vacation travel to Spanish-speaking country.
   --------- Spanish language magazines/ newspapers.
   --------- Spanish language TV.

7. Do you have a foreign accent (i.e., American accent) in Spanish? → Yes No
   If yes, please rate the strength of your accent.

□ No Accent □ Slight Accent □ Moderate Accent □ Strong Accent

Thank you for your participation.

Appendix C: Edinburgh Handedness Inventory

Name: ___________________________________ Date: ______________________

Date of Birth: _______________

Edinburgh Handedness Inventory

Please indicate your preferences in use of hands in the following activities by putting + in the appropriate column. Where the preference is so strong that you would never try to use the other hand unless absolutely forced to, put ++. If in any case you are really indifferent, put + in both columns.

Some of the activities require both hands. In these cases, the part of the task or object for which hand preference is wanted is indicated in parentheses.

Please try to answer all the questions, and only leave a blank if you have no experience at all with the object or task.

PREFERENCE = + In the Appropriate Column
STRONG PREFERENCE = ++ In the Appropriate Column
NO PREFERENCE = + In Both Columns

	Task	Left	Right
1.	Writing		
2.	Drawing		
3.	Throwing		
4.	Scissors		
5.	Toothbrush		
6.	Knife (without a fork)		
7.	Spoon		
8.	Broom (upper hand)		
9.	Striking Match (match)		
10.	Opening box (lid)		
i.	Which foot do you prefer to kick with?		
ii.	Which eye do you use when using only one? (microscope)		

Appendix D: Pretest and Posttest Stimuli

Tap /ɾ/ - Trill /r/

Word	Gloss	Word	Gloss
ahora	now	**ahorra**	he saves
boro	boron	**borro**	I erase
carera	rip-off artist	**carrera**	race
churo	handsome	**churro**	pastry
coral	coral	**corral**	farmyard
curar	to cure	**currar**	to work
encerar	to polish	**encerrar**	to lock up
enterado	well-informed	**enterrado**	buried
eres	you are	**erres**	r's
fiero	fierce	**fierro**	iron
foro	forum	**forro**	lining
Gora	name	**gorra**	(ball) cap
morena	dark	**morrena**	moraine (glacial)
morón	moron?	**morrón**	knock/hit
soro	year-old hawk	**zorro**	fox-m
Taro	Place/Name	**tarro**	jar
turón	polecat/fieldmouse	**turrón**	candy (nougat)
vara	stick	**barra**	rail

Trill /r/ - /d/

Word	Gloss	Word	Gloss
arroba	at symbol	**adoba**	He pickles
arrobo	bliss	**adobo**	marinade
berro	watercress	**vedo**	I prohibit (hunting)
borrega	sheep	**bodega**	winery
burra	female donkey	**buda**	Buddha
carrete	cadet	**cadete**	spool
cerramos	we saw	**cedamos**	Let's transfer!
cirro	cirrus	**sido**	was (pp of ser)
correra	He will run	**codera**	elbow patch
corrillo	small group of people	**codillo**	elbow/knuckle
error	error	**hedor**	stench
gorro	cap	**godo**	goth/gothic
irreal	unreal	**ideal**	ideal
porra	nightstick	**poda**	pruning (season)
puerro	leek	**puedo**	I can
sarro	plaque	**sado**	sado (masochism)
serrar	to saw	**sedar**	to sedate
torrero	lighthouse keeper	**todero**	jack of all trades

Tap /ɾ/ - /d/

Word	**Gloss**	**Word**	**Gloss**
aire	air	**hay de**	There are..
armara	He will arm	**armada**	armada
chiro	cool	**chido**	fantastic
Daro	name	**dado**	dice
duroso	painful	**dudoso**	doubtful
hiere	injures (herir)	**hiede**	stinks (heder)
ira	rage	**ida**	outward journey
lirio	iris	**lidio**	I fight
loro	parrot	**lodo**	mud
mera	mere	**meda**	from ancient Media (Mede)
mora	blackberry	**moda**	style
Nara	name	**nada**	nothing
oro	gold	**odo**	Name
parecer	to seem	**padecer**	to suffer
peral	pear tree	**pedal**	pedal (bike, piano)
quería	wanted	**qué día**	what day
tajara	He will sharpen	**tajada**	slice
violara	He will violate	**violada**	violated

Appendix E: Generalization Stimuli

New Words – Familiar Speaker

Word	Gloss	Word	Gloss
boro	boron	**borro**	I erase
curar	to cure	**currar**	to work
fiero	fierce	**fierro**	iron
morón	moron?	**morrón**	knock/hit
torear	to fight	**torrear**	to fortify with towers/turrets
birete	NW	**birrete**	mortar board
euro	euro	**eurro**	NW
jera	amount of field that can be plowed by a pair of oxen in one day	**jerra**	NW
merar	mix liquors	**merrar**	NW
uraca	NW	**urraca**	magpie
arrobo	bliss	**adobo**	marinade
carrete	cadet	**cadete**	spool
corrillo	small group of people	**codillo**	elbow/knuckle
parrilla	grill	**padilla**	small frying pan
serrar	to saw	**sedar**	to sedate
arriós	NW	**adiós**	goodbye
forraje	forage	**fodaje**	NW
larro	NW	**lado**	side
narrir	NW	**nadir**	opposite of zenith
zurriar	to hum/buzz	**zudiar**	NW
armara	He will arm	**armada**	armada
duroso	painful	**dudoso**	doubtful
lirio	iris	**lidio**	I fight
Nara	name	**nada**	nothing
quería	wanted	**qué día**	what day
dero	NW	**dedo**	finger
fura	shy	**fuda**	NW
Mario	name	**madio**	NW
oriar	NW	**odiar**	to hate
roreo	NW	**rodeo**	detour

New Words – New Speaker

Word	**Gloss**	**Word**	**Gloss**
carera	rip-off artist	**carrera**	race
encerar	to polish	**encerrar**	to lock up
foro	forum	**forro**	lining
paral	wooden trough	**parral**	vine arbor
turón	polecat/fieldmouse	**turrón**	candy (nougat)
claro	clear	**clarro**	NW
ferar	NW	**ferrar**	garnish with iron
larín	NW	**larrín**	NW
naria	NW	**narria**	sledge/fat woman
riera	laugh (form of reir)	**rierra**	NW
berro	watercress	**vedo**	I prohibit (hunting)
cerramos	we saw	**cedamos**	Let's transfer!
error	error	**hedor**	stench
porra	nightstick	**poda**	pruning (season)
terrero	terrace	**tedero**	iron candlestick
borrador	eraser	**bodador**	NW
gorrena	NW	**godena**	rich
morrada	butting of heads	**modada**	NW
nirrito	NW	**nidito**	little nest
sábarro	NW	**sábado**	Saturday
cerilla	match	**zedilla**	cedilla
entrara	he will enter	**entrada**	entrance
loro	parrot	**lodo**	mud
oro	gold	**odo**	Name
tajara	He will sharpen	**tajada**	slice
dorado	gilted	**dodado**	NW
gerente	manager	**gedente**	NW
meria	NW	**media**	half
ruera	NW	**rueda**	wheel
niro	NW	**nido**	nest

New Words – New Dialect

Word	**Gloss**	**Word**	**Gloss**
churo	handsome	**churro**	pastry
enterado	well-informed	**enterrado**	buried
Gora	name	**gorra**	(ball) cap
soro	year-old hawk	**zorro**	fox-m
vara	stick	**barra**	rail
dares	given	**darres**	NW
gara	NW	**garra**	claw
siere	NW	**cierre**	Close!
norabuena	congratulations	**norrabuena**	NW
rural	rural	**rurral**	NW
borrega	sheep	**bodega**	winery
cirro	cirrus	**sido**	was (pp of ser)
gorro	cap	**godo**	goth/gothic
puerro	leek	**puedo**	I can
torrero	lighthouse keeper	**todero**	jack of all trades
charra	gaudy	**chada**	NW
hurra	hurray	**huda**	NW
merrio	NW	**medio**	half
orraca	alcohol distilled from coconut	**odaca**	NW
salarra	NW	**salada**	salty
chiro	cool	**chido**	fantastic
hiere	injures (herir)	**hiede**	stinks (heder)
mera	mere	**meda**	from ancient Media (Mede)
parecer	to seem	**padecer**	to suffer
peral	pear tree	**pedal**	pedal (bike, piano)
ensalara	NW	**ensalada**	salad
narie	NW	**nadie**	noone
pare	stop	**pade**	NW
saluro	NW	**saludo**	greeting
zurana	stock-dove	**zudana**	NW

Appendix F: Cross-Modal Priming Stimuli

Lexical Item Targets:

Auditory Prime	**Gloss**	**Word Freq.**	**Visual Target**	**Gloss**	**Word Freq.**	**Target Type**
para	for	237280	**para**	for	237280	/ɾ/
duro	hard-m	1795	**duro**	hard-m	1795	/ɾ/
miro	I look	565	**miro**	I look	565	/ɾ/
hora	hour/time	10433	**hora**	hour/time	10433	/ɾ/
pero	but	131191	**perro**	dog	1992	/r/
coro	choir	1391	**corro**	I run	414	/r/
caro	expensive	849	**carro**	car	1102	/r/
varo	peso (in Mexico)	34	**barro**	I sweep	774	/r/
toro	bull	1068	**todo**	all	74189	/d/
cara	face	7134	**cada**	each	27660	/d/
moro	moor	863	**modo**	style	14442	/d/
dura	hard-f	1847	**duda**	doubt	8028	/d/

Nonword Targets:

Auditory Prime	**Gloss**	**Word Freq.**	**Visual Target**	**Gloss**	**Target Type**
cere	NW		**cere**	NW	/ɾ/
dera	NW		**dera**	NW	/ɾ/
fara	NW		**fara**	NW	/ɾ/
liro	NW		**liro**	NW	/ɾ/
luro	name	2	**lurro**	NW	/r/
rara	bizarre/rare	1300	**rarra**	NW	/r/
sara	name	236	**zarra**	NW	/r/
ore	ore	17	**horre**	NW	/r/
flora	flora	419	**floda**	NW	/d/
paro	strike	465	**pado**	NW	/d/
fero	NW		**fedo**	NW	/d/
curo	I cure	21	**cudo**	NW	/d/

Appendix G: Training Stimuli

Tap /ɾ/ - Trill /r/

Word	Gloss	Word	Gloss
chora	gutsy	**chorra**	silly
chorear	to swipe/steal	**chorrear**	to drip
choro	gutsy	**chorro**	stream
mirada	glance	**mirrada**	composed of myrrh
paral	wooden trough	**parral**	vine arbor
perico	parakeet	**perrico**	little dog
serado	parcel of baskets	**serrado**	serrated
torear	to fight	**torrear**	to fortify with towers/turrets
zorita	stock dove	**zorrita**	little female fox
zuro	spike, ear	**zurro**	I wallop
pera	pear	**perra**	dog-f
aras	altars	**arras**	security
Bora	Place/Name	**borra**	He sweeps
cerilla	match	**cerrilla**	die for milling coins
cero	zero	**cerro**	hill
cero	zero	**serro**	I saw
Cora	name	**corra**	Run!
jara	rock-rose	**jarra**	jug
Lara	name	**Larra**	name
lloro	I cry	**jorro**	bad tobacco
mira	He looks	**mirra**	myrrh
moral	moral	**morral**	rucksack
pira	pyre	**pirra**	crazy about
poro	leek	**porro**	lazy
vira	welt	**birra**	beer
voraz	voracious	**borras**	You erase
amarillo	yellow	**amarrillo**	NW
birete	NW	**birrete**	mortar board
claro	clear	**clarro**	NW
dares	given	**darres**	NW
enero	January	**enerro**	NW
era	was	**erra**	NW
erada	NW	**errada**	miss
euro	euro	**eurro**	NW
ferar	NW	**ferrar**	garnish with iron
gara	NW	**garra**	claw
horario	schedule	**horrario**	NW
iridio	iridium	**irridio**	NW
jera	amount of field; pair oxen	**jerra**	NW

larín	NW	**larrín**	NW
lironada	clean, neat	**lirronada**	NW
mara	NW	**marra**	deficiency
marón	NW	**marrón**	brown
merar	mix liquors	**merrar**	NW
naria	NW	**narria**	sledge/fat woman
norabuena	congratulations	**norrabuena**	NW
orador	orrator	**orrador**	NW
purela	NW	**purrela**	inferior wine
rareza	rarity	**rarreza**	NW
riera	laugh (form of reir)	**rierra**	NW
rural	rural	**rurral**	NW
salario	salary	**salarrio**	NW
siere	NW	**cierre**	Close!
soregar	NW	**sorregar**	to water in another course
teraje	NW	**terraje**	rent
tira	strip	**tirra**	NW
turar	NW	**turrar**	to roast
uraca	NW	**urraca**	magpie
veraco	NW	**verraco**	boar
zorollo	reaped while unripe (wheat)	**zorrollo**	NW
zura	NW	**zurra**	beating
zureo	billing and cooing	**zurreo**	NW

Trill /r/ - /d/

Word	Gloss	Word	Gloss
arrobe	Enter!	**adobe**	adobe
parrilla	grill	**padilla**	small frying pan
terrero	terrace	**tedero**	iron candlestick
borras	you erase	**bodas**	weddings
corra	He runs	**coda**	coda
narrado	narrating	**nadado**	swimming
narrar	to narrate	**nadar**	to swim
perrazo	large dog	**pedazo**	piece, bit
torra	he roasts	**toda**	all - f
zorra	fox-f	**soda**	soda (water)
zorras	foxes - f	**sodas**	sodas (water)
arras	security	**hadas**	fairies
borra	He sweeps	**boda**	wedding
cerrilla	die for milling coins	**zedilla**	cedilla
cerro	hll	**cedo**	I transfer
jarra	jug	**jada**	place
jarra	jug	**jada**	place (azada)
Larra	name	**lada**	(jara?) Place/Name
mirra	myrrh	**mida**	he measures
morral	rucksack	**modal**	modal
pirra	crazy about	**pida**	he orders
porro	lazy	**podo**	I prune
serra	he saws	**seda**	he sedates
serro	I saw	**sedo**	I sedate
armarrillo	NW	**armadillo**	armadillo
arriós	NW	**adiós**	goodbye
borrador	eraser	**bodador**	NW
charra	gaudy	**chada**	NW
derral	NW	**dedal**	thimble
dirruir	to destroy	**diduir**	NW
forraje	forage	**fodaje**	NW
gorrena	NW	**godena**	rich
hurra	hurray	**huda**	NW
irrogar	to cause	**idogar**	NW
jorro	bad tobacco	**jodo**	NW
larro	NW	**lado**	side
limonarra	NW	**limonada**	limonade
merrio	NW	**medio**	half
morrada	butting of heads	**modada**	NW

murrez	NW	**mudez**	muteness
narrir	NW	**nadir**	opposite of zenith
nirrito	NW	**nidito**	little nest
orraca	alcohol distilled from coconut	**odaca**	NW
purror	NW	**pudor**	shyness
rorro	suckling child	**rodo**	NW
ruerro	NW	**ruedo**	bull ring
sábarro	NW	**sábado**	Saturday
salarra	NW	**salada**	salty
serreta	dim	**sedeta**	NW
sorrio	NW	**sodio**	sodium
tirrena	relating to ancient Tuscany	**tidena**	NW
torrecilla	turret	**todecilla**	NW
turrell	NW	**tudel**	pipe in bassoon
verra	NW	**vera**	bank (river)
verraca	sailor's tent	**vedaca**	NW
verreja	NW	**vedeja**	lion's mane
virrente	NW	**vidente**	clairvoyant
zárriva	NW	**zádiva**	leaves that soften corns
zorrazo	big fox	**zodazo**	NW
zurrapa	dregs/sediment	**zudapa**	NW
zurriar	to hum/buzz	**zudiar**	NW
salirra	NW	**salida**	exit

Tap /ɾ/ - /d/

Word	Gloss	Word	Gloss
entrara	he will enter	**entrada**	entrance
doro	I gild	**dodo**	dodo
mura	Place/Name	**muda**	mute/he mutates
murar	to wall	**mudar**	to change
muro	wall	**mudo**	mute/I mutate
ara	altar	**hada**	fairy
aras	altars	**hadas**	fairies
Bora	Bora Bora	**boda**	wedding
cerilla	match	**zedilla**	cedilla
cero	zero	**sedo**	I sedate
Cora	name	**coda**	coda
era	era	**eda**	Name
jara	rock-rose	**jada**	place (azada)
Lara	name	**lada**	(jara?) Place/Name
lloro	I cry	**yodo**	iodine
mira	look	**mida**	measure
moral	moral	**modal**	modal
Noro	name	**nodo**	node
pira	pyre	**pida**	He orders
poro	leek	**podo**	I prune
sera	large basket	**seda**	silk
vera	bank (river)	**veda**	closed (hunting season)
vira	welt	**vida**	life
voraz	voracious	**bodas**	weddings
ciurad	NW	**ciudad**	city
dero	NW	**dedo**	finger
dorado	gilted	**dodado**	NW
ensalara	NW	**ensalada**	salad
ere	redundancy plan	**ede**	NW
erucación	NW	**educación**	education
fura	shy	**fuda**	NW
gerente	manager	**gedente**	NW
irioma	NW	**idioma**	language
jirafa	giraffe	**jidafa**	NW
loriga	coat of mail	**lodiga**	NW
Mario	name	**madio**	NW
meria	NW	**media**	half
narie	NW	**nadie**	noone
nariz	nose	**nadiz**	NW
niro	NW	**nido**	nest
oriar	NW	**odiar**	to hate

parada	stop (bus stop)	**padada**	NW
pare	stop	**pade**	NW
pureta	old fogey	**pudeta**	NW
rario	NW	**radio**	radio
roreo	NW	**rodeo**	detour
ruera	NW	**rueda**	wheel
saluro	NW	**saludo**	greeting
saraje	baskets	**sadaje**	NW
soroche	altitude sickness	**sodoche**	NW
surista	NW	**sudista**	Southerner
teriar	NW	**tediar**	to loathe
tiro	shot	**tido**	NW
torazo	large bull	**todazo**	NW
tostara	NW	**tostada**	toast
turesco	NW	**tudesco**	wide cloak
urana	coy, reserved	**udana**	NW
verado	NW	**vedado**	prohibited
viraje	curve/turn	**vidaje**	NW
zura	NW	**zuda**	Persian wheel
zurana	stock-dove	**zudana**	NW
zurito	little glass of wine	**zudito**	NW