The Effect of Diacritic-Enhancement on L2 Pronunciation of Shared L1/L2 Graphemes

By

Paul Dion Grosse

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Spanish)

at the

UNIVERSITY OF WISCONSIN-MADISON

2023

Date of final oral examination: 08/18/2023

The dissertation is approved by the following members of the Final Oral Committee:
Rajiv Rao, Professor, Spanish & Portuguese
Catherine Stafford, Associate Professor, Spanish & Portuguese
Fernando Tejedo-Herrero, Professor, Spanish & Portuguese
L. J. Randolph Jr., Assistant Professor, Curriculum & Instruction

Table of Contents

List of Figures

List of Tables

DEDICATION

To my Lord and Savior, Shepherd and Friend, Jesus Christ;

to my wife, Emily;

to my parents, Tess and Carolyn;

and to my children, Noah and Evangeline.

"Oh, the depth of the riches and wisdom and knowledge of God! How unsearchable are His judgments and untraceable His ways! For who has known the mind of the Lord, or who has become His counselor? Or who has first given to Him, and it will be repaid to him? Because out from Him and through Him and to Him are all things. To Him be the glory forever. Amen." (Romans 11:33-36, Holy Bible Recovery Version).

ACKNOWLEDGEMENTS

First and foremost, thank you to my amazing advisor, Dr. Rajiv Rao, for your ever-cheerful encouragement, patience, and guidance throughout this process. I'd also like to extend my special thanks to the other distinguished professors who kindly agreed to serve on my dissertation committee: Dr. Cathy Stafford, Dr. Fernando Tejedo-Herrero and Dr. L. J. Randolph Jr..

I'm very grateful to all of the linguistics professors in the department who taught me so much about the field and inspired my thinking: Dr. Rajiv Rao, Dr. Cathy Stafford, Dr. Fernando Tejedo-Herrero, Dr. Grant Armstrong, Dr. Diana Frantzen, and Dr. Sandro Sessarego. Thank you, also, to the many other professors and colleagues I had the pleasure to take a class with here at UW-Madison, as well as the helpful staff members in the Department of Spanish and Portuguese.

I would like to especially acknowledge all of the help I received from Dr. Kyle Parrish in analyzing the data, producing so many customized data visualizations, and even providing some coaching in Bayesian statistics.

In order to both start and finish this part of my journey in life, I needed a lot of prayer, encouragement and support from so many of my brothers and sisters in Christ. While I cannot name them all here, I want to especially thank Robert Hengst, David and Jean Olson, Frank and Gloria Liao, and the campus team: Jacob, Jackie, Bryan, Mercy, Claire, and Tyler.

Finally, thank you so much to my beloved wife, Emily, for encouraging me, listening to me, and doing everything you could to make this possible.

ABSTRACT

Within the field of linguistics, whether considering language contact situations (Weinreich, 1979) or foreign language education (Lado, 1957), the topic of language transfer, especially as it relates to pronunciation, has always been an item of particular interest. While research on such transfer has mostly focused on various phenomena of the L1 and L2 phonological systems (Flege, 1987; Flege & Bohn, 2021), in the past number of decades a number of studies have highlighted the effects caused by a shared orthographic system, whenever language pairs share such a system (Vokic, 2011; Young-Scholten and Langer, 2015). Such effects have been found to be persistent, not disappearing even after years of residence in the L2-speaking environment, even though the pronunciation issue itself is not problematic for the learner at all.

Because this effect is embedded in the orthography itself, the present study sought to explore the possible benefits of manipulating the visual information provided to L2 learners in their orthographic input in order to assist them in their acquisition of a more native-like L2 pronunciation. Looking at a small, homogenous group (n = 25) of university-level, English-speaking learners of L2 Spanish, an experiment was designed focusing on the pronunciation of Spanish intervocalic <d> as [ð], which should not be a problematic sound for these learners phonologically speaking. Participants were divided between three groups in order to test for changes in pronunciation behavior over a 6-week period, either: by normal classroom experience alone, by the help of a brief pronunciation lesson on the target sound, or by combining a pronunciation lesson with the use of a diacritic-enhanced <d̰> both in the lesson as well as in subsequent paragraph-reading tasks. It was discovered that a brief pronunciation lesson alone was unable to make lasting differences in students' pronunciation as measured by a paragraph-

reading posttest, but the group who interacted with diacritic-enhanced text showed substantial improvement in pronunciation of the item. These improvements were found to still be intact in a paragraph-reading task after the diacritics were removed, suggesting that learners had associated the target-like [ð] pronunciation with the authentic <d> itself.

Because the results of the present study were obtained using a relatively small sample size, some caution should be exercised in the interpretation of these results, and further study of this particular effect must be carried out in order to corroborate what was found here. Nonetheless, the results of this study indicate the potential for real benefits to incorporating diacritic-enhancement in L1/L2 learning situations where the two languages share an orthographic system but have diverging phonetic realizations of graphemes.

CHAPTER 1: INTRODUCTION

Over the past thirty years, the field of second language (L2) phonological acquisition has explored and documented in increasing detail the relationship between L2 orthographic input (OI) and L2 phonology (e.g., Bassetti, 2006; Rafat, 2011; Vokic, 2011; Young-Scholten, 1995; Zampini, 1994). More precisely, a number of studies have examined how first language (L1) grapheme-phoneme correspondence (GPC) rules, that is, the association of specific graphemes to specific phonemes (e.g., English <t> = /t/), can affect pronunciation of L2 items when the L1 and L2 share the same orthographic system, such as the Roman alphabet used across a wide range of language pairs. This interaction has proven to be a source of both positive and negative phonological transfer, as learners have a tendency to carry L1 sounds into the L2 for the shared graphemes. Because written language is one of the primary ways a classroom learner interacts with the L2, understanding these transfer effects, and potentially mitigating any hindrances they create, could be a real asset to effective L2 instruction. In the case of negative transfer, that is, L1 sound insertions that do not correspond to native L2 phonology, the strength of L1 GPC rules has been shown to be considerable, such that even with extensive L2 experience the transfer is often not overcome (Vokic, 2011; Young-Scholten and Langer, 2015). Usage-based theories such as connectionism (Koda, 2007) seem to provide a rational explanation for this transfer phenomenon in that a strong neurological mapping has already developed between an L1 grapheme and a related phoneme (e.g., English <z> and /z/), making the automatized L1 pronunciation for the shared L1/L2 grapheme difficult to overcome.

One intervention that has recently demonstrated potential in helping learners avoid such transfer is the use of diacritically-enhanced L2 graphemes (Showalter and Hayes-Harb, 2013), optionally combined with explicit phonetic instruction to imbue the diacritic with phonetic

information, which may help learners associate a more target-like sound with the problematic graphemes (Jackson 2016). The idea of enhancing or making modifications to L2 orthography in such things as foreign language textbooks is certainly not a new one for language educators. Language instruction in L2 Chinese often uses a pinyin alphabet that is enhanced with diacritics over the vowels to help learners acquire the lexical tone (e.g., high <mā> "mother", rising <má> "hemp", low-falling-rising <mǎ> "horse", and high-falling <mà> "scold"). Pronunciation exercises in an L2 have also been enhanced with other formatting distinctions, such as changes in color to draw attention to target graphemes (see for example Morgan, 2010, pp. 169-170). Yet surprisingly it is only very recently that the effectiveness of L2 grapheme enhancement has begun to be empirically investigated.

So far, what research has been done has focused on the effects of grapheme enhancement on L2 perception, and there has been some early evidence for benefits to L2 perception while interacting with modified orthography. Showalter and Hayes-Harb (2013) investigated whether or not English-background learners of Mandarin would benefit from the use of diacritics to distinguish lexical tone. In two picture-auditory matching experiments, the researchers found that a learner group who had interacted with diacritic-enhanced orthography outperformed a comparison group who had interacted with the same materials, but without diacritic enhancement. In another picture-auditory matching investigation, Jackson (2016) found that when learning a new L2 Arabic phonemic contrast (namely, velar plosive /k/ vs. uvular plosive /q/ - which does not exist in L1 English), learners who both received explicit phonetic instruction and also interacted with a diacritically enhanced grapheme <ḳ> for the new phoneme /q/ outperformed those who received the same phonetic instruction but interacted with an unenhanced orthography (<k> = /q/).

While these L2 perception studies do give some initial indications of potential benefits to orthographic enhancement, there have not been any production studies for this intervention, and the degree to which such improvement applies to the wide-ranging variety of nuanced L1/L2 GPC interactions remains to be thoroughly studied and understood. Furthermore, while some benefit has been documented for L2 learners during the initial stages of interaction with enhanced learner texts, there has not been any investigation of whether those benefits persist once learners go back to interacting with authentic L2 orthography, or whether the perception improvements seen in experimental tasks actually translates to the learner's L2 phonological system.

In terms of present day language-learning experiences and the potential effects of OI and GPC interactions, the Spanish-English language pair seems especially relevant, as Spanish and English share the same orthographic system, and Spanish has by far the highest foreign language enrollment figures in colleges in the United States (Looney & Lusin, 2019). Within this L1-L2 context, one example of negative transfer that has been suggested to be in part due to OI and incongruent GPCs is <d>, especially in its fricative/approximant contexts. For example, such early L2 Spanish learners will often pronounce Spanish <cada> ("each") as either [ˈka.da] or [ˈka.ɾa], when in fact the typical Spanish pronunciation would be [ˈka.ða]. When mispronounced as [ˈka.ɾa], it even encroaches on the Spanish lexical items *cara* and *caro* ("face" and "expensive," respectively). Such learners have to disassociate or inhibit L1 English phonetic realizations of this grapheme (namely [d] or [ɾ]) and reassociate it with a different sound ([ð]). What is somewhat intriguing about the target L2 sound in the case of intervocalic <d> is that it is not a novel or articulatorily difficult L2 sound, but rather is very similar to the fricative [ð] of the L1 phonetic repertoire of English, in which it also appears intervocalically (e.g., "rather" and

"bathing"). This particular grapheme thus provides an opportunity to test for the possible effects of using artificially modified L2 orthography in order to attain a more nativelike L2 phonology.

The present study therefore seeks to shine light on the effects of OI especially in the L1 English - L2 Spanish context, and on any possible benefits that might be achieved with explicit pronunciation instruction and the help of diacritic-enhanced orthography. The grapheme <d> will specifically be targeted, which presents a particular challenge to said learners, who must merge an L1 English allophone [ð], corresponding to <th> in English, with the L2 Spanish GPC <d> = /d/. At the same time, these learners will have to disassociate L1 [ɾ] from Spanish <d>. See Figure 1.1 for a visualization of this learning process. In addition to providing additional empirical research on the subject, the present study also points to possible pedagogical improvements that may assist these learners.

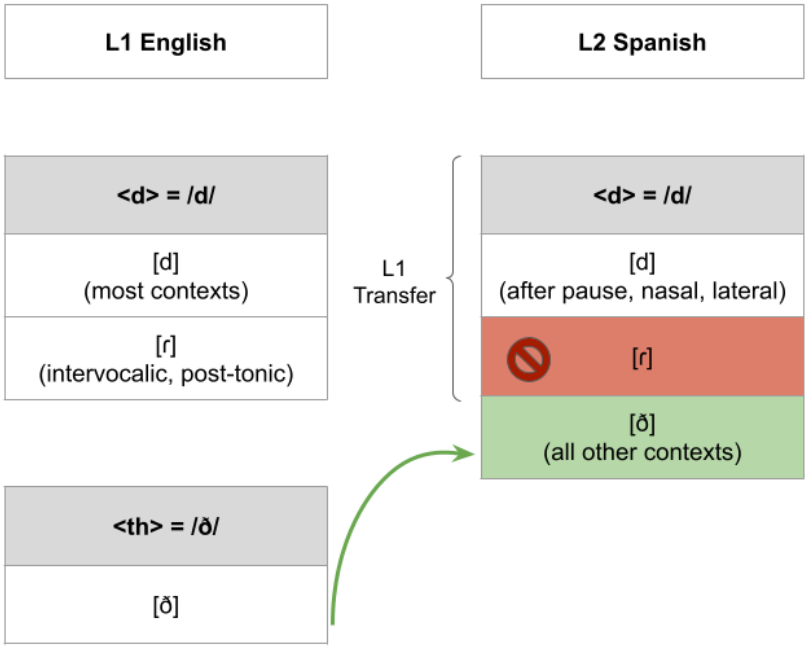


Figure 1.1 Diagram of transfer of L1 <d> = /d/ and needed relearning of L2 /d/

The remainder of this dissertation is organized as follows: Chapter 2 provides a review of relevant literature related to the effects of OI in the L2 learning process, presenting a theoretical framework for L1 and L2 phonological issues as well as a review of relevant pedagogical interventions, especially through explicit phonetic instruction and orthographic enhancement; Chapter 3 describes the methodology of the present study, in which an experiment was designed to test the effects of phonetic instruction and orthographic enhancement on participants enrolled in intermediate Spanish classes at a university in the Midwest United States; Chapter 4 presents the results of the experiment; and Chapter 5 discusses the implications of the findings for the field of SLA, presents some limitations and directions for future work, and then provides a conclusion to this study.

CHAPTER 2: LITERATURE REVIEW

2.0 Overview

This chapter provides a review of past research concerning the effects of OI on L2 phonology and also lays out a theoretical framework in which these effects can be understood. Section 2.1 gives an extensive review of the documented effects of OI on L2 phonology. Section 2.2 presents a theoretical framework that describes underlying psycholinguistic processes and structures that are most relevant to OI and L2 phonological acquisition. Section 2.3 presents various pedagogical approaches that have sought to either mitigate the negative effects or else capitalize on the positive effects of OI on L2 phonology. Section 2.4 introduces the object of the present study, namely the orthography-related dynamics of Spanish L2 phonological acquisition for English-background learners, focusing on the particular case of <d>.

2.1 Effects of OI on L2 phonology

2.1.1 Background - L2 phonological transfer

Transfer effects, that is, the influence that L1 features and learning has on the development of an L2, have been a major topic of study in the field of linguistics for decades (Weinreich, 1979; Lado, 1957; Odlin, 1989; MacWhinney, 1997). While Weinrich (1979) referred to these effects as *interference*, Odlin (1989) used the term *transfer*, which he defined as "the influence resulting from similarities and differences between the target language and any other language that has been previously (and perhaps imperfectly) acquired" (Odlin, 1989, p.27). In that same book, Odlin demonstrates such L1 influence on the acquisition of L2 grammar, vocabulary, and pronunciation.

Concerning transfer effects on the acquisition of L2 phonology, Lado (1957), in his Contrastive Analysis Hypothesis, proposed that the learner would have the most difficulty with L2 items that had the greatest contrast with related L1 items. In a later study, which has become a foundation for L2 phonological acquisition, Flege (1987) argued that it was not always the most contrastive L2 sounds that gave learners difficulties, but that it was those that were similar to L1 sounds that caused the most trouble, as the learner assumes an "equivalence" of the two sounds and uses L1 sounds that are similar to the L2 target but actually still inaccurate. Indeed, the function of "equivalence" classification as a stage in L2 phonetic development still features prominently in Flege and Bohn's (2021) revised Speech Learning Model (SLM-r). MacWhinney (1987) uses a connectionist framework (presented later in this chapter) to explain a possible underlying reason for such transfer phenomena, asserting that because the L1 cognitive structures must be the base for the development of the L2, the initial state of L2 phonology will be a complete transfer of the L1 phonological system, which must henceforth be reconfigured through L2 learning.

But in the past three decades, researchers have also begun to uncover a more specific issue tied up with L1 phonological transfer, namely, those effects that are not merely related to L1-L2 phonological differences, but those that are related to interactions between L1 and L2 orthography. After some initial considerations related to OI were commented on in various L2 phonological acquisitions research in the mid-1990s, more recently the importance of these orthographic effects has led to entire special issues dedicated to the subject in the journals *Language and Speech* (introduced by Simon & Van Herreweghe, 2010) and *Applied Psychology* (introduced by Basetti et al., 2015). In the next section, a summary of the research focused on orthographic effects in L2 phonology will be presented.

2.1.2 Orthographic effects in L2 phonological transfer

One of the first documented instances of what was suggested to be L1 transfer due to OI and L1 GPC is in Zampini (1994). Testing the pronunciation of L1 English learners of Spanish, she found that her two groups of participants, university students in their second and fourth semesters of Spanish instruction, were frequently pronouncing <v> as [v] in Spanish, which for most native speakers would be pronounced as [b/β]. Interestingly, these students pronounced [v] for <v> more often in the reading task than in the conversational one, leading Zampini to suggest that the visual input of orthography was influencing such [v] pronunciations. Perhaps an even more intriguing finding was that the students who were in their 4th-semester were pronouncing <v> as [v] even more frequently than the 2nd-semester students, leading her to further postulate that the emphasis on reading and writing in the 3rd and 4th semesters may have reinforced errant pronunciation due to increased OI over that period.

A more elaborated commentary on the effects of OI on pronunciation was presented in Young-Scholten (1995), a chapter detailing the many ways that a learner's L2 input diverges from the typical native speaker's L1 learning experience. Here, the author explains why OI would have a potentially significant effect on L2 learners who are literate in an L1 that uses the same alphabet as the L2, pointing out how classroom learners are often inundated with OI via textbooks, worksheets, and classroom displays from the very beginning of their studies. Given their literacy in the L1, which built up strong associations between L1 graphemes and phonemes, and "because the L2 phonology has not yet been established, the learner will only be able to access the L1 phonology" as they interact with L2 orthography (Young-Scholten 1995, p. 113). In this way, learners unwittingly bring L1 realizations of those graphemes into their L2 phonological systems. Based on this realization, the author pointed out the need for educators to

study the dynamics of the L1 and L2 orthographic systems, anticipating phonological transfer effects and seeking to mitigate them with an informed approach to language instruction, even stating that "the teaching of L2 speech is essentially instruction in the pronunciation of target language graphemes" (Young-Scholten 1995, p. 112). The author's identification of the great influence of the all-pervasive written word on literate L2 learners broke ground on a new topic of research in the SLA community, which over the past three decades has seen increasing interest and development as researchers seek to understand the reasons for these effects and apply this knowledge in a way that improves L2 teaching and learning.

In a recent state-of-the-art review on the influence of orthography on L2 phonological acquisition, Hayes-Harb & Barrios (2021) provided a categorization of the various strains of research that have been carried out. They identified various lines of study on the effects of orthographic input, including on L2 phonological awareness, on L2 phoneme perception, on the pronunciation of L2 words, and related to the acquisition of L2 phonological processes.

2.1.2.1 Effects of orthographic input on L2 phonological awareness

Phonological awareness is a kind of metalinguistic awareness related to the sound structure of words, including the ability of a learner to identify the number of sounds (phonemes) and syllables in a word. Bassetti (2006) examined the effect of OI on both L2 phonological awareness and phonological production in her investigation of university-level English-background learners of Mandarin (average of 24 months of prior learning). Many Mandarin learners are presented with Hanyu Pinyin orthography, a romanized method of alphabetically representing Mandarin words, and these participants all reported using pinyin 'very often' or 'often.' She conducted two experiments that centered on various rimes ([iou], [uei], [uən]) in Chinese, which are represented in pinyin by just two graphemes after a consonantal onset (e.g.,

<ui> = [uei] for <gui>) but as three graphemes in the absence of an onset (e.g., <wei> = [uei]). Crucially, native speakers pronounce the rime as [uei] in both cases, despite the difference in pinyin orthographic representation. In both a phoneme-counting task as well as a phoneme segmentation task, L2 learners identified more sounds within the target rime when it was represented by 3 graphemes in pinyin, but fewer sounds when the same rime was represented by just 2 graphemes. In a second experiment in the study, Bassetti (2006) also tested the learners' pronunciation of these target items, and found that they also tended not to pronounce the triphthong for items where the pinyin only used two characters for the vowels. Bassetti (2007) replicated this study with Italian-speaking learners of Mandarin and found the same effects. Crucially, in both studies, the participants were not exposed to pinyin during the experiment itself, but only the logographic Hanzi characters, and thus the effects were shown to be most likely due to the considerable interactions they had with pinyin during their years of education. Thus, Bassetti (2006) and Bassetti (2007) demonstrated that these learners to some degree depended on orthographic representation in order to build up their underlying phonological representation of these items. Interestingly, these studies showed that the pinyin OI could have a positive effect on learner phonology, since for the rimes written with 3 characters, the orthography actually helped learners to count the correct number of sounds.

In another study on orthography's influence on phonemic awareness, Pytlyk (2017) investigated the perception of L2-Russian and L2-Mandarin Chinese phonemes by native speakers of Canadian English. In her introduction, the author lays out the relevant research to date showing a strong, perhaps even "unbreakable bond between letters and sounds" formed when alphabetic knowledge is acquired during L1 acquisition (Pytlyk 2017, p. 233). She points to a line of studies which show how learning to read in different writing systems effects the

subword language units in different ways, such that learning a syllabary in the L1 (e.g., Japanese *kana* and *hiragana*) creates a greater awareness of consonant-vowel units (i.e., syllables), and learning an alphabet (e.g., Roman and Cyrillic) creates a greater awareness of phonemes. Concerning those who learn an L1 that uses an alphabet, it has been found that "the ability to count phonemes only surfaces once children learn to read via an alphabet" (Pytlyk, 2017, p. 233), and that only individuals who go through the experience of learning and using an alphabet are able to breakdown words into their component sounds "because alphabets sensitize individuals to the phonemic level" (Pytlyk, 2017, p. 234).

Pytlyk (2017) took this understanding of the power of alphabetic learning in the L1 and tested it on L2 learning. In this study, she showed that two groups of L2 learners, one learning Russian (using the Cyrillic alphabet) and the other using Mandarin (using pinyin), were sensitive to the L2 orthographic representation in terms of their phonemic awareness. In a phoneme-counting task, these learners were more successful when there was a consistent number of L2 graphemes and phonemes in the word (e.g., *всё* /fsʲɔ/, three letters/three phonemes) than in words with inconsistent correspondences (e.g., *звать* /zvatʲ/, five letters/four phonemes). In this way, the author found evidence that L2 phoneme awareness is influenced by L2 orthography, and that "orthographic effects are not limited to the L1" (Pytlyk, 2017, p. 233).

2.1.2.2 Effects of orthographic input on L2 phonological perception

In addition to the studies on phoneme awareness, there has also been a line of work documenting the effects of orthography on L2 phoneme perception, the ability of a learner to accurately perceive (hear and categorize) L2 phonemes in auditory input. Weber and Cutler (2004) carried out an eye-tracking study that showed that Dutch-background learners of English treat English /ɛ/ and /æ/ asymmetrically. Dutch contains a vowel /ɛ/ (which is similar though not

identical to English /ɛ/) but does not contain /æ/. English, however, possesses both /ɛ/ and /æ/. This particular English contrast has been noted to be a difficult one for Dutch-background learners to perceive, as the Dutch /ɛ/ is somewhat in between English /ɛ/ and /æ/. Learners were presented auditory forms along with pictures representing English words such as 'pencil' and 'panda,' and eye-tracking data was used to see what learners looked at during the brief period in which they heard just the first syllable, either [pɛn] or [pæn], which are not phonemically contrastive in Dutch. Interestingly, Weber and Cutler found that despite the fact that these front vowels are hard for such learners to discern, they did not respond to the syllables the same way. When hearing /pæn/ (of "panda"), participants looked at both the pencil and panda pictures, but when hearing /pɛn/ (of "pencil"), they tended to just look at the pencil picture. In reflecting upon these findings, Cutler et al. (2006) proposed that one possible explanation for this asymmetric behavior was that the L2 orthographic representations of these two sounds (<e> vs. <a>) might be interacting with the Dutch L1 GPC patterns of these graphemes, providing a facilitating effect to resolve /ɛ/ as <e> and not <a> (since <e> is mapped to /ɛ/ in Dutch), but no such helpful effect for resolving /æ/ as <a> (since <a> corresponds to a low back vowel in Dutch).

While Cutler et al. (2006) settled instead on another explanation for the asymmetry (namely, that Dutch /ɛ/ was closer to English /ɛ/ than /æ/, making English /ɛ/ more discernible), Escudero et al. (2008) carried out a near replication of Weber and Cutler (2004). They recruited a very similar sampling of English-learners from a Dutch background, and instead of testing them with known words, the researchers used some English non-words paired with pictures of non-objects and had participants pass through a learning phase with the new items. Crucially, one group of learners saw orthographic representations to go along with the picture an auditory form of the words, while a comparison group did not see the orthographic form. Interestingly,

the comparison group did not show any asymmetry, looking to pictures of words starting with /ɛ/ or /æ/ indiscriminately, whereas the experimental group, which had seen orthographic forms for the words, manifested the same asymmetry demonstrated in Weber and Cutler (2004), providing further evidence for the L2 phonological effects of L1 GPC rules interacting with L2 OI.

In another study looking at L2 perception, Hayes-Harb et al. (2010) began looking at how GPC incongruencies between languages could affect recognition of the audio form of L2 words. They tested "naive" learners of new words in a "foreign," pseudo-English non-language. In the study, participants were presented new words via a picture of an object along with an audio form of that word's pronunciation. Additionally, learners were exposed to the orthographic form of the word, which was either congruent or incongruent with typical English GPC rules. A control group was presented with a null "XXXX" orthographic representation for the words. After the word-learning phase was completed, participants completed a picture-auditory form matching test. Results showed that words whose orthographic form was incongruent with English GPC rules were less accurately matched with their auditory form compared with congruent words, demonstrating that the incongruent orthography presented at learning had affected how learners internalized the phonological form of the word.

2.1.2.3 Effects of orthographic input on L2 phonological processes

There has also been a growing body of research looking more specifically at the influence of orthography on phonological production, especially as it pertains to phonological processes in the L2 where an L2 phoneme can have different allophonic realizations. Some examples of this, which are specifically investigated in the studies below, include German devoicing of word-final /b, d, g/ (realized as [p, t, k]), Korean weakening/deletion of word-internal, intervocalic /h/ (often realized as ∅), Spanish spirantization of intervocalic /b, d, g/ (realized as [β, ð, ɣ], and English

flapping of word-internal /t, d/ (realized as [ɾ]). In all of these cases, the orthographic representation of the letters has been shown to affect L2 pronunciation of various learners.

Young-Scholten (2002) carried out a longitudinal case study in which she demonstrated possible effects of OI in three high-school L1 English learners of German in a one-year study abroad program in Germany. Seemingly due to their L1 GPC rules, these learners tended to pronounce word-final <b, d, g> as [b, d, g], even though native speakers of German systematically devoice these segments to [p, t, k]. Interestingly, the learner that had the lowest rate of devoicing was also the one who reported doing the most reading in German. The author pointed out this negative correlation between a greater rate of L2 orthographic input (that is, spending more time reading German text) and a lower rate of final devoicing, suggesting that the increased frequency of reading word-final <b, d, g> had also reinforced an erroneous pronunciation of word-final [b, d, g], despite the learner's simultaneous experience with native speakers who systematically devoiced these segments.

Expanding a bit on Young-Scholten (2004), wherein English-background learners of German showed difficulty in pronouncing word-final <b, d, g> as [p, t, k], Young-Scholten and Langer (2015) went on to find that such learners also have difficulty in pronouncing word-initial <s> as [z], favoring instead the English GPC of [s] for <s>, even after living for an extended period of time living in Germany and interacting with native speakers. What is intriguing in these cases is that the native pronunciation is well within L1 phonological inventory; word-final [p, t, k] and word-initial [z] are very frequent in English. Thus, it seems the OI combined with the conflicting GPCs in English and German are leading to non-native pronunciation of these items.

2.1.2.4 Effects of orthographic input on the pronunciation of L2 lexical items

Finally, in addition to orthography's problematic interactions with allophonic variation in the L2, researchers have also found that it can affect the pronunciation of L2 lexical items as a whole. Han and Kim (2017) examined the interaction of orthography with the acquisition of L2 phonological processes and word pronunciation. They based their approach on Bürki et al. (2012), in which the authors investigated the possible effects of orthography on French speakers who were auditorily taught novel (pseudo) French words. The words contained what auditorily was pronounced as a consonant cluster (e.g., [pl]) that could potentially represent either a true phonemic cluster (/pl/) or a reduced (allophonic) variant pronunciation of two consonants separated by an unpronounced, underlying schwa (e.g., [pəl]/[pl] in *peluche* "cuddly, plush toy"). Bürki et al. (2012) discovered that even after the participants had learned the pronunciation of a word through four days of hearing the auditory form without a schwa (e.g., /pluʁ/), those who were afterwards exposed to a spelling of the target word one time with a potential schwa <e> (e.g., <pelour>) were subsequently more likely to pronounce a schwa in between the consonants, even though they had never been exposed to any auditory forms in which the schwa was pronounced. In this way, the authors demonstrated the strong effect of orthography upon an already developed phonological form of a lexical item.

Han and Kim (2017) carried out a similar study, expanding on the L1 findings of Bürki et al. (2012) to test whether L2 learners might also be affected by a single exposure to an orthographic form after having learned a word auditorily. They specifically looked at the phonological process of Korean /h/ weakening/deletion in particular phonetic contexts. Because this phoneme is usually not pronounced in certain intervocalic and other contexts, if a new word was only presented in auditory form, the learner would normally presume that there was not a

phoneme there. On the other hand, as a learner gains more experience with the language, they may become aware of the possible presence of a silent /h/ in that context. The authors tested two groups of Mandarin-speaking learners of Korean, a beginner group (6 months or less of learning) and an advanced group (more than 1 year of learning), following the pattern of Bürki et al. (2012). Participants completed multiple days of auditory learning of new words, and were in a subsequent session exposed one time to an orthographic form. The authors had three orthographic conditions for each proficiency group: either seeing the word with no character in the potential /h/ spot, seeing the /h/ character there, or else not ever being exposed to orthography. It was found that among the beginning learners, those who saw the word without an /h/ character pronounced an /h/ 4% of the time, those who saw the word with an /h/ character pronounced it 26% of the time, and those who never saw the spelling of the word only pronounced an /h/ 6% of the time. Again, these participants never heard an auditory form containing [h], and had been trained fully with a pronunciation without such a sound, but one exposure to an orthographic representation had seemingly restructured their phonological representation of the words. In the case of these particular L2 learners, the L1 of Mandarin Chinese used a completely different, logographic script, and the effects were entirely due to the new, L2 syllabary script of Korean.

Coming to cases where the L1 and L2 share the same script, Vokic (2011) looked at advanced L2 learners of English whose L1 was Spanish. English and Spanish, of course, both use the same Latin alphabet, with only a few diacritic elements distinguishing them (e.g., Spanish <ñ, ü, á, é, etc.>). The participants in her study had begun English instruction around middle-school age and were now mostly in their 30's and had been residing in the US for about 4 years, on average. She looked at these learners' production of a target, word-internal English

flap ([ɾ]) represented by <t>, <d>, <tt>, and <dd>. If these learners were to rely on Spanish GPC rules, these graphemes would be pronounced as either [t], [d], or [ð]. In a sentence-reading task with embedded target words that balanced each of the four possible orthographic representations of the target flap, Even though English [ɾ] is very similar to Spanish /ɾ/, in Spanish it is only represented by <r>. Vokic showed that only a few learners produced the tap consistently, with most participants producing [t], [d], or [ð] for many test items. An interesting finding was that participants produced the tap more frequently with the digraphs, <tt> and <dd>, which do not have a GPC in Spanish, compared to the individual graphemes <t> and <d>, which do have Spanish GPCs of [t] and [ð], respectively. An additional item-based analysis was carried out to see if the relative frequency of the English words, as measured by corpus studies, had any interaction with pronunciation. Indeed, she found that these learners produced the flap 57% of the time in frequent lexical items, and 34% of the time in infrequent ones, demonstrating perhaps that language experience with native speakers may have improved the phonological form of some high-frequency items. This would also indicate that the pronunciation of the tap is sensitive to the lexical item, which suggests that even if the English phonological process of tapping is not yet acquired, it may be possible for the phonological form of particular words to be corrected through experience. But generally speaking, in the absence of such word-specific learning, the L1 GPC rules were more strongly affecting how low-frequency words were being sounded out.

The results of Vokic (2011) thus point to a persisting transfer effect for L2 graphemes with conflicting L1 GPCs, despite a number of years of residence in a country where their L2 is the majority language. Like Young-Scholten (2002) and Young-Scholten and Langer (2015), increased L2 interaction with native speakers had not brought these learners to a native-like

pronunciation of the target phoneme, even though such pronunciations fall well within their L1 phonological system, demonstrating the considerable power that OI and L1 GPC rules have on their L2 phonology. But the fact that word frequency did have a positive effect on target flap production does indicate that increased experience in the L2 did help learners with specific lexical items - if not their L2 phonological system as a whole. It is especially interesting that Vokic (2011) found that when the problematic graphemes <t> and <d> appeared as digraphs <tt> and <dd>, the learners seemed to more readily produce the target sound - a sound that in no way corresponds to the L1 GPC. This suggests that learners attend to the visual context around a single grapheme and are not locked into pronouncing a grapheme according to the L1 GPC if other contextual information points them away from doing so.

In a more extensive look at L1 English-L2 Spanish orthographic transfer effects, Rafat (2011) identified a number of problematic graphemes where English and Spanish GPCs were incongruent (e.g., <v> = Spanish [b] vs. English [v]; <z> = Spanish [s] vs. English [z]). Working with English-background participants with no knowledge of Spanish, she experimented with the learners' exposure to orthography in various ways. In one of the experiments, she designed a word-learning task in which participants saw a picture, heard the pronunciation of the corresponding word, and then later tried to produce it during a picture-naming test. Four groups of participants were formed in which the author modulated the exposure to orthography both at the time of learning [+/- ortho-learning] and at the time of production [+/- ortho-production]. This gave a total of four possible scenarios of the presence of orthography: ortho-learning and production (OLP); ortho-learning only (OL); ortho-production only (OP); and no orthographic input (O-).

Rafat found that both the OLP and OL groups evidenced strong transfer effects in the presence of orthography (with mean proportions of transfer of .53 and .54, respectively) for items that contained one of the problematic graphemes. The OP group also showed a substantial transfer effect (.43), while the O- group, as might be expected, showed a much weaker effect (.08), attributable to other sources of L1 phonological transfer. For example, learners who never saw orthographic forms still showed significant L1 transfer for intervocalic <d>, only producing the target [ð] 45% of the time. This might be due to the rarity of English words with word-medial /ð/ before vowels other than a vocalized rhotic. Nonetheless, when orthography was presented, these learners only produced [ð] between 2-10% of the time. Overall, her findings demonstrated a considerable increase in L1 phonological transfer for incongruent GPCs when orthography was presented.

One interesting finding was that some graphemes led to relatively less transfer than others. For instance <ll> = /j/ only saw L1 transfer 1%, 21%, and 18% of the time in OLP, OL, and OP conditions, respectively (compare to <z> = /s/: 69%, 67%, 64%; and <v> = /b/: 99%, 92%, 77%). Rafat suggests that the differential phonological transfer for these graphemes may be due to the degree of salient phonetic difference between L1 and L2 GPCs. In other words, the English GPC for <ll> (/l/) is more saliently different from the Spanish GPC (/j/), and thus learners more readily notice the difference (Spanish <ll> ≠ /l/) and seem to open up a new L2 GPC rule (<ll> = /j/). With <z>, on the other hand, because the L2 sound is not as saliently different (/z/ ≈ /s/), the learner may not break free from incorporating the L1 GPC (<z> = /z/) in their L2 representation.

Taken together, the findings in this section have demonstrated that orthographic input can be the direct cause for both positive and negative transfer effects in L2 phonology, affecting

phonological awareness, perception, and production. Concerning the negative effects on phonological production, OI has been shown to sometimes hinder or prevent the acquisition of L2 phonological processes and also to alter the pronunciation of L2 words, even lexical forms already acquired through auditory input. For language pairs that share an orthographic system, L1 transfer effects related to incongruent L1-L2 GPC is prevalent. These negative effects, however, appear to sometimes be mitigated by increased experience with specific words (as seen in high-frequency words being more native-like than low-frequency). It is also possible for the learner to notice a salient difference between an L1-transferred sound and the L2 sound (e.g., Spanish <ll> ≠ English /l/), and potentially overcome incongruent L1-L2 GPC even at the outset of L2 learning. However, even with significant native L2 experience, for pronunciation differences that are less salient, proficiency in L2 pronunciation may still be hindered by transfer effects rooted in orthographic phenomena.

2.2 Theoretical approaches to mapping

The many observed effects of orthography on L2 pronunciation compel us to seek an understanding of the underlying processes by which such developments are manifested. Turning to prevalent theories in SLA, those that have emerged from a psycholinguistic perspective seem to offer the most help in understanding possible relationships between the components of the L1 and L2 systems as they relate to their visual components of orthography. In this section, some usage-based theoretical constructs will be presented, along with possible explanations for the findings related to orthographic-related language transfer.

2.2.1 Form-function mapping

Functionalism defines language learning as a process of form-function mapping through experience and usage of a language. An example of this in spoken English would be learning that pronouncing the phonemic form /d/ at the end of a present tense verb such as *raise* has a specific function - it communicates that the verb happened in the past. Even pre-literate children comprehend and actively use this linguistic feature. During the subsequent process of learning to read in English, the child would come to learn that the orthographic form <d> is pronounced /d/ (mapping alphabetic form to phonetic form), and eventually they will learn that <d> added to the end of <raise> would carry out the same linguistic function. Thus, according to a psycholinguistic framework for language systems, a learner acquires a wide array of visual and auditory mappings and builds up the capacity to interact in this system in an efficient way. But how is it that these mappings develop?

Koda (2007) suggests that connectionism is one such theory that provides a rational explanation for how relationships emerge. The essence of this usage-based theory is that by cumulative experience with a language, the learner internalizes the mappings between form and function, with the elements that are contacted the most frequently having the strongest association in the mind of the learner. For example, in English monosyllabic C(C)VC(C) words, the letter <i> is strongly associated with the sound /ɪ/, as in "hid," "pig," "king," and "wind," yet it also has some association with /aɪ/, such as in "kind" and "pint." Connectionist theory would postulate that when reading aloud such words, the form-function mapping of <i> = /ɪ/ would be stronger than <i> = /aɪ/, simply because the learner would have contacted the association of <i> = /ɪ/ much more frequently. This stronger connection would be manifested in faster processing

times when reading aloud <i> = /ɪ/ words (compared to <i> = /aɪ/ words), and also a propensity for pronouncing unknown <i> words with /ɪ/ rather than /aɪ/.

Concerning language systems and the form-function mapping, some specific forms of the language include orthography, phonology, and morphology, all of which must become systematically mapped both to one another and to their semantic meaning. Taking the example of "king" above, through the learner's interaction with the English language over time, the phoneme string /kɪŋ/ would be mapped to its morpheme, "king". The letter string <king> would also become mapped to that morpheme semantically (<king> = "king"), as well as phonemically (<king> = /kɪŋ/). Since English has an alphabetic orthography, the individual letters and digraphs would also get mapped to specific phonemes (i.e., <k> with /k/, <i> with /ɪ/, and <ng> with /ŋ/). The theory thus describes how, over time, the learner will build up a system which allows him/her to produce and perceive its linguistic forms and their related meaning both efficiently and accurately.

Connectionist theory further describes the learning process as "a gradual transition from deliberate efforts to automatic execution, and its outcome as a dynamic, ever-changing state, rather than a static entity" (Koda, 2007, p. 10). The process of arriving at automatic execution is fittingly referred to as *automatization* (Logan, 1988), whereby frequent input causes nondeliberate, nonvolitional activation of the constellation of form-function mappings. Ellis (2006) characterizes this process as tuning the learner's developing system to the cues of the L1, automatizing its representation and production. He provides the example of the English learner interacting with <one> = *one*. The first time that this new form-meaning pair is noticed, the learner may form an explicit memory of the relationship between its form and meaning. In this way a detector unit is added to the learner's perception system such that if the learner interacts

with the grapheme string <one> again in the future, this detector could now signal the word's presence in the input. At the initial stage, it may take the learner more effort to detect and interpret this new form in the environment, but the more frequently the learner encounters this grapheme string <one> with the interpretation of "one," the higher the "resting level of activation" of this detector becomes, and the more easily the learner will detect this form and jump to its interpretation in the future.

2.2.2 Grapheme-phoneme mapping

The building up of such mappings for languages with a writing system, of course, also becomes developed between orthographic forms and phonological forms. There is consensus in the literature that through progressive experience in a language, strong associations are built up between graphemic representations and their phonological representations. According to Seidenberg (1985), "As salient orthographic units are recognized, they activate their phonological representations" (Seidenberg, 1985, p. 5). Because alphabetic languages represent sounds with letters, if there are consistent mappings of those graphemes to particular sounds, then grapheme-phoneme correspondences will be built up. For Seidenberg, these grapheme-phoneme correspondences become the means by which newly confronted or less familiar words are phonologically processed. He states that while reading low-frequency (or newly learned) words, the grapheme-phoneme correspondence of the language will be drawn upon in order to access the proper phonological form associated with the meaning of the word. In other words, the visual representation of the letters/words combined with the sounding out of those letters help establish lexical meaning. For example, at the early stages of learning English writing, an L1 learner may have to sound out <digger> based on known grapheme-phoneme correspondences, pronouncing /dɪgɹ/, and eventually arriving at its interpretation: "digger." As experience with the

new word increases, its orthographic form eventually arrives at a condition of "direct access" to the lexical entry, such that while the grapheme-phoneme relationships are activated, they are not relied upon for lexical access.

But how exactly does a letter arrive at its related sound, and is the individual letter the unit for sound representation? Glushko (1981) firstly contested that although theories at that time had maintained an assumption that languages generated a set of grapheme-phoneme rules, that these correspondences are better described as activations of known pronunciations. Furthermore, he suggests and then demonstrates that it is not so much individual letters that are processed for pronunciation (e.g., <g> and <h> in "laughable"), but rather words (e.g., <laughable>), morphemic word stems (<laugh->), or strings of letters (<gh>). He describes a phonological activation mechanism whereby the reading of letter strings (words or morphemes) causes the activation of the known pronunciation of those strings. Returning to the question of rules, while prior theories had said that pronouncing new words would need the assistance of a separate mechanism which applies letter-to-sound rules (that is, explicit knowledge about how letters "should" sound), Glushko's (1981) experiments demonstrated that such a mechanism does not manifest itself, but instead, he claims that it is the activation of the (implicitly) known pronunciation of other strings of letters that feeds the process of new word pronunciation.

Glushko's (1981) proposal for letter strings over letters as the unit of GPC may be somewhat anglocentric, as English is known to be an orthographically deep language; that is, its graphemes are infamous for their ambiguity of pronunciation (as seen in the range of pronunciations of <gh> in "laugh," "bought," "ghost"). Thus, it is possible that for learners of English, with individual graphemes that do not have a reliable grapheme-phoneme correspondence, it is more efficient to consider whole letter strings (words or morphemes) to

arrive at a proper pronunciation. This is part of the proposal of Katz and Frost (1992) in their orthographic depth hypothesis. Since learning to read and pronounce words is different in terms of the grapheme-phoneme faithfulness between an orthographically deep language such as English and a language with a more shallow orthography such as Spanish (where grapheme-phoneme relations are much more transparent), Katz and Frost propose that this will lead to language-specific differences in how learners process GPCs.

But whether the learner uses individual letters or letter strings, eventually a system seems to emerge whereby, upon reading a new, unfamiliar word, a plausible (though not necessarily accurate) phonological form can be generated. In fact, Share (1995) proposes that learners use this process of phonologically recoding newly confronted orthographic words in what he presents as The Self-Teaching Model. When further describing this model in Share (1999), he asserts that "phonological recoding acts as a self-teaching device or built-in teacher enabling a child to independently develop the word-specific orthographic representations essential to skilled reading and spelling" (Share, 1999, p. 96). In other words, as a learner confronts a new word while reading, he/she will immediately decode the orthography and associate a self-generated phonological form, which greatly assists in internalizing the new word into his/her memory and lexicon. For example, if a learner bumps into the new word "respite" while reading <respite>, they may derive an interpretation from context, and also form a self-taught phonological form by sounding it out (probably as /ɹi'spaɪt/, whether that is a correct pronunciation of the word or not). Share demonstrates how learners seem to employ this self-teaching function right away with new words in order to quickly and efficiently retrieve the new word when it is encountered again.

Taken together, usage-based research related to the internalization of the orthographic and phonological forms of words seems to paint the following picture. During the early,

literacy-acquisition stage, as the orthographic form of words is mapped to the phonological form (thus building an orthographic-phonemic lexicon), the frequency of specific associations between orthographic and phonological forms builds up the automatic activation of letter-to-sound and letter-string-to-sound-string mappings. These mappings are heavily relied upon as learners take the coordinated efforts of orthographic and phonological familiarity to arrive at the intended meaning of words and build up their L1 vocabulary. As certain words are read more and more frequently, eventually the orthographic form of the word gains *direct access* to the lexicon, such that while the mappings of letters and sounds is activated, the intended meaning of the orthographic word is reliably activated without reliance on the "sound it out" process. New or low-frequency words, however, would still rely on such a process, and upon contacting new words in orthographic forms, learners will immediately decode them into phonological forms in order to lexicalize the newly encountered word to quickly and efficiently recognize the word in the future.

2.2.3 Modeling the language system of reading and speaking

Based on the cognitive views of language and reading described above, various computational models have been developed to encapsulate the cognitive systems involved in reading, with a view to replicating human patterns of learning. One of those models, showing how the learner reads words differently depending on whether a word is known or not, is called the Dual Route Cascaded (DRC) model (Coltheart et al., 1993; Coltheart et al., 2001). A simplified version of that model from Castles et al. (2018) is presented in Figure 2.1. Starting at the bottom of the diagram, the model illustrates that when reading (the letter units at the bottom of the model), an individual will either process the letters/words as already known lexical items by the leftward branch (route 1), or if the word is unknown to the reader, they would sound it out

by using the Grapheme-Phoneme Rule system of the rightward branch (route 2). Either way, a phonological form will be activated and the reader will seek to arrive at the intended meaning (Semantic system) of what they are reading in the most efficient way.

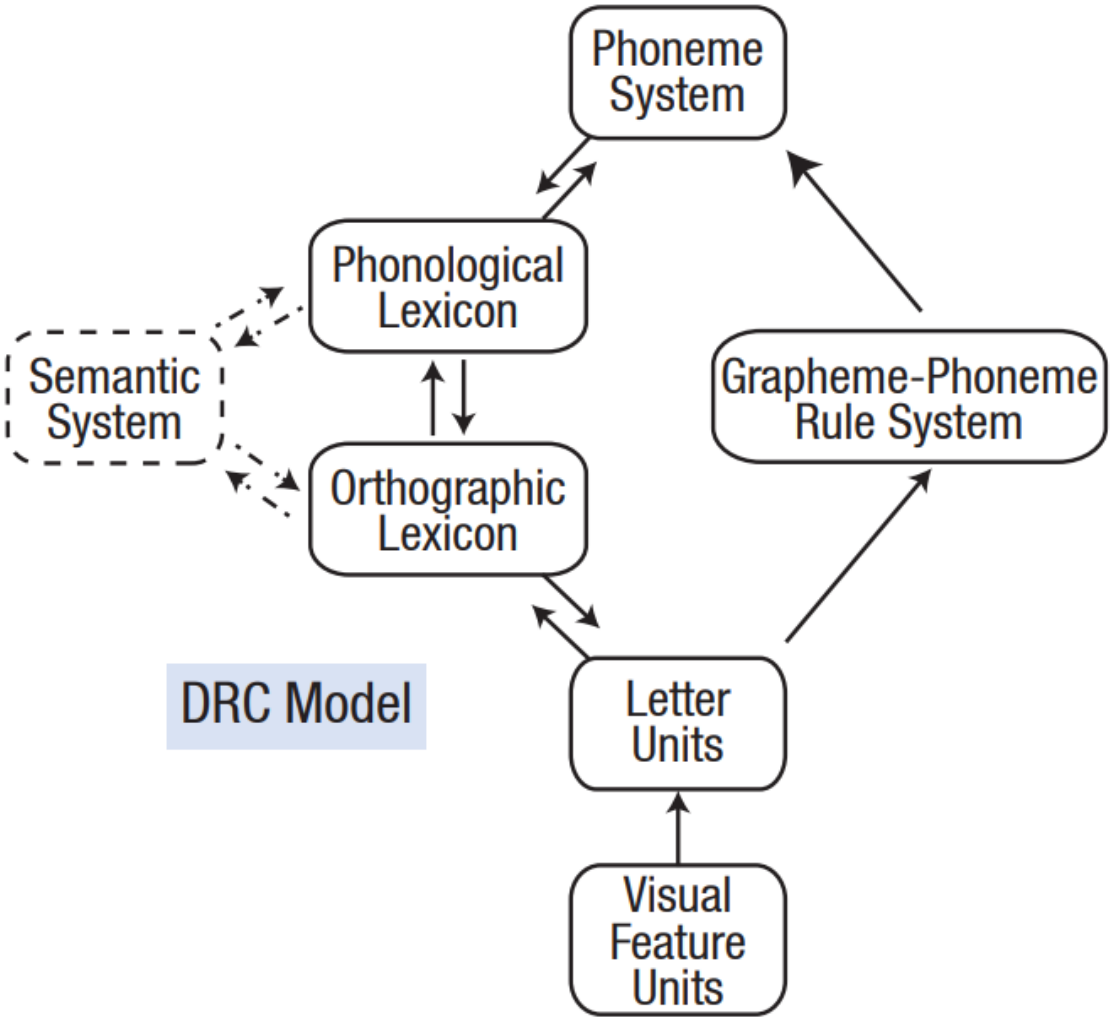


Figure 2.1 Simplified illustration of the DRC Model, from Castles et al. (2018)

Coltheart et al. (1993) present a number of psychological studies that give evidence for there being two such routes, namely in the studies of surface dyslexia and phonological dyslexia, and the way that individuals with such impairments process regular words (words or non-words that follow the normal grapheme-phoneme rules, such as “king,” “wind”) and exception words (a term used to refer to those words that break the typical pronunciation rules, such as “kind,” and

"pint"). In surface dyslexia, which is an acquired dyslexia (a reading impairment caused by brain damage in a previously literate person), the individual has trouble recognizing known words if they do not follow the typical GPC rules. For example, an individual with surface dyslexia would have no trouble reading "bite," but would struggle to read a word like "flood," whose GPC rules would point towards a pronunciation rhyming with "mood." Looking at the DRC model, individuals with surface dyslexia would have impairment to the left-branching neurological link in route 1, which links the letters directly to the orthographic and phonological lexicons. Instead, known lexical items can only be retrieved by first passing through the right-branching route 2. Additionally, because these individuals do have access to the GPC rules, they do not have difficulty reading nonwords out loud according to those rules (e.g., *nust*, *cobe*, or *ploon*). But if a previously known exception word does not follow the normal GPC rules, its retrieval from the lexicon would be hindered or even blocked.

Another kind of reading impairment is phonological dyslexia. In this acquired dyslexia, individuals maintain the ability to read out loud words that are known to them, whether they are regular words or exception words, even long, more complex ones. But they have difficulty reading aloud non-words (like those listed above), or words that are unknown to them. They may also have trouble telling what sounds individual letters make, even though they can name the letter itself. In this case, looking at the DRC model, such individuals seem to have damage to their neurological link to the right-branching route 2.

While the above two cases of dyslexia give evidence of two distinct routes of pronunciation and semantic retrieval, this does not mean that typical, unimpaired function of this system chooses only one route or the other. Rather, in normal cases, both the grapheme-to-phoneme assembling route (via GPC rules) as well as the lexical route (direct recognition of

orthographic lexemes) are activated during reading. Once an individual starts to read and process the letters, both routes become activated and the process of phonological decoding and semantic retrieval "cascades" along in both directions, potentially aiding one another in the retrieval process. "When the DRC model is reading aloud or performing the lexical decision or Stroop tasks, phonology is always partly assembled and always partly lexical" (Coltheart et al., 2001, p. 212).

If we presume that the linguistic systems used in reading resemble what is described in the DRC model, we could revisit our previous description of L1 language and alphabetic literacy development in English and trace how the pathways and systems would possibly emerge. Of course, a child would begin at the earliest ages with interactions between a semantic system and a phonological lexicon (e.g., /kat/ = "cat"). The phoneme system would begin to take shape, especially as the child begins to speak, though they would likely have little to no phonemic awareness. They will eventually begin to learn the shapes and names of letters, forming the system of letter units (e.g., <c>, <a>, <t>), and probably shortly thereafter they will begin to be taught the grapheme-phoneme rules (e.g., <c> = /k/, <a> = /æ/, <t> = /t/). In this way a Grapheme-Phoneme Rule system will begin to be internalized, thus developing the right-branching route 2. They will increasingly contact the language in orthographic form (e.g., <cat>), and any printed words contacted frequently will gain direct access to the emerging orthographic lexicon, beginning the development of the left-branching route 1, both in learning to recognize the meaning of orthographic lexemes (e.g., <cat> = "cat"), and also fusing them with their phonological lexemes (<cat> = /kat/ = "cat"). As the practice of reading increases, the child will also be able to independently sound out words according to the grapheme-phoneme rules (route 2) and potentially connect the sounded out word with its meaning , thus "closing the

circuit" and connecting the rightward branch with the lexicon and the semantic system (e.g., <c, a, n> = /k/, /æ/, /n/ = /kan/ = "can"). As L1 experience progresses day by day and year by year, strong neural networks will develop within and between each of these systems, and the activation and function of the systems will become automatized and tuned to the patterns of L1.

2.2.4 The DRC model and L2 orthography/phonology

While this summary has described theoretical acquisition processes related to orthography and phonology in the L1, how do they affect SLA? Usage-based models assert that an L2 is necessarily shaped by L1 knowledge and experience. According to Ellis, "the initial state for SLA is no longer a plastic system, it is one that is already tuned and committed to the L1. Our later experience is shaded by prior associations; it is perceived through the memories of what has gone before" (Ellis, 2006, p. 109). The effects of such L1 tuning on L2 acquisition contribute to what is commonly called L1 interference and L2 transfer. Koda (2007) describes this as nonvolitional L1 activation when processing L2 information. Applying this principle to orthographic exposure to an L2, it implies that previous learning of L1 grapheme-phoneme correspondence (whether at the individual grapheme or the letter string level) will affect learning outcomes for the L2, especially if the two languages share the same set of graphemes.

The DRC model provides a theoretical framework of word recognition and reading aloud that is particularly helpful when considering the OI-related transfer in L2 phonology. If indeed the starting point of the L2 is a system that is already tuned and committed to the L1, we can look at the systems defined in the DRC model and easily see how OI would very likely have a significant influence on the learners developing L2 phonology. The Grapheme-Phoneme Rule System is especially problematic for language pairs that share the same writing system, such as Spanish and English. The learner already has powerful neurological relationships between

letters and their sounds, both in their underlying phonemic representations and the motor cortex for pronouncing them. Thus, when L2 text is read, the already *entrenched* activation of letter pronunciations will cause the right-branching route to be activated as they attempt to pronounce the word and seek any L2 lexical entries based on L1 GPC rules. For example, a novice English-speaking learner of Spanish, when reading aloud the word <hora>, would probably pronounce [ˈho.ɹə], and look for a lexical entry that matched. If the learner knows, say from a Spanish textbook, that the text <hora> means "hour/time," then there would already be an entry in their Orthographic Lexicon, linking <hora> to its relevant Semantic System entry. Thus, if an accurate entry of /ˈo.ɾa/ has not been built up in the Phonological Lexicon through statistical frequency, it's possible that an alternative, L1 GPC-generated entry of /ˈho.ɹə/ would develop as an entry there. Furthermore, it's likely that at the early stages of learning, such a learner would speak and use this pronunciation of the word with others, building up further frequency and strength of the erroneous pronunciation. If L2 orthographic input were left unchecked by sufficient native auditory input, the learner's Phonological Lexicon would develop many L1-influenced pronunciations of words.

But Rafat (2011) shows us that auditory input can at least to some extent pull learners away from pronunciations that are fully according to L1 GPC rules. She found that for words with the digraph <ll>, which would generally activate /l/ for the English-speaking participants, novice Spanish learners pronounced /j/ between 79% and 99% of the time even though they were exposed to the orthography alongside the audio input. Compare that to the grapheme <z>, in which these participants only pronounced the target /s/ between 31% and 36% of the time when exposed to orthography, which pulled pronunciations to /z/, following English GPC rules. Rafat

attributed this difference to the more salient difference between the target L2 sound /j/ and English /l/, as compared to the difference between L2 /s/ and L2 /z/.

Applying this finding to the DRC model, it seems possible that early learners can overcome L1 GPC rules if the L2 auditory input noticeably differs from pronunciations that the L1 GPC would dictate. If that is indeed what is going on in these learners, it would also mean that the learner could begin constructing a new, native-like L2 GPC Rule System as they increase frequency of fusing the L2 phonemes with the shared L1-L2 graphemes (e.g., the previously mentioned L2 learners develop a new Spanish GPC of <ll> = /j/). It should be noted that in Rafat (2011), learners were carrying out a simple picture-naming, word-learning task. They were presented with the auditory input alongside a picture, either with or without the orthography. Therefore, the lack of strong transfer effect for the case of <ll> in this study does not necessarily mean that the L2 GPC of <ll> = /j/ has been fully formed. It could be that if further L2 input were heavily orthographic, such as in some sort of reading tasks, that L1 GPC rules would still have an effect for this grapheme. Furthermore, another possible reason for these learners' success with <ll> may be that English has already adopted Spanish words that incorporate the Spanish GPC rule for <ll> (e.g., *tortilla* = /toɹˈti.ja/), and there would thus be some familiarity with such a pronunciation.

But while these learners did succeed in avoiding the L1 GPC rules for <ll>, they demonstrated transfer effects for <v> (pronouncing /v/ instead of /b/), /z/ (/z/ instead of /s/), <d> ([d] or [ɾ] instead of [ð]), and <h> (/h/ instead of ∅). Indeed, in prior studies, transfer effects for <v> pronounced as /v/ were found in Spanish learners in their 4th semester and beyond (Zampini, 1994). Face and Menke (2009) found that Spanish learners pronounced intervocalic <d> as [d] rather than target [ð] long into their L2 experience. In their analysis of university

Spanish learners at three levels, namely, fourth-semester (average of 5.15 years of Spanish study), graduating Spanish majors (7.95 years of Spanish study), and Ph.D. students (average of 13.1 years of Spanish study), the authors found that when reading a story out loud, these participants still pronounced intervocalic <d> as a stop for 67%, 34%, and 16% of the tokens, respectively. Their results demonstrate the difficulty for these learners in overcoming L1 GPC rules even after thirteen years of instruction and significant L2 experience.

These studies, along with the many presented earlier in this chapter concerning the effects of OI on pronunciation, can now be analyzed through the lens of the DRC model in order to highlight specific systems that may be at the root of the problem. If we ignore the case of cognates (e.g., Spanish and English <motor>), where the printed L2 word could be recognized as an L1 word (which would activate the left-branching, direct access to the orthographic/ phonological lexicons), all other L2 words would at their initial stages of learning have to be sounded out according to the right-branching, Grapheme-Phoneme Rule System. While auditory input seems to mitigate the discrepancies that are especially salient (e.g., Spanish <ll> = /j/ and not /l/), for less-noticeable phonological differences, reading L2 text may reinforce errant pronunciations. If the Phoneme System and the Phonological Lexicon are built up based on statistical frequency, then every time the learner reads and associates the letters and words with an errant pronunciation, these systems will develop with significant effects caused by L1 GPC rules, especially if these pronunciations are accepted and even pronounced by interlocutors.
In this light, we can see the considerable power that orthographic input can have on the learner's emerging L2 system, and we can understand why Young-Scholten (1995) would say that "the teaching of L2 speech is essentially instruction in the pronunciation of target language graphemes" (Young-Scholten, 1995, p. 112). Because the typical auditory and orthographic

input that learners encounter, whether in instructed settings or in real-world, study abroad programs, has proven to be insufficient in helping them overcome many L1 GPC transfer effects, it begs the question of whether alternative pedagogical interventions could be helpful.

2.3 Pedagogical approaches to OI related transfer

2.3.1 Audio only, or audio first

Based on their findings related to the effects of OI on L2 phonology, researchers have sometimes discussed strategic pedagogical approaches to possibly lessen or avoid such effects. Rafat (2011), for example, proposed that teachers expose learners to the auditory input prior to exposing them to orthographic input. Of course, in terms of what has become the typical foreign language classroom experience, this would involve a radical change, as textbooks and other text-based instructional media are relied upon even from the onset of learning. It also remains to be seen how much "audio-only" instruction would be necessary to build up a strong enough L2 phonological base in order to survive the effects of OI once it is introduced. Furthermore, as was found in the case of Bürki et al. (2012) and Han and Kim (2017), it is possible that even after a word has been phonologically acquired through audio-only instruction, even one exposure to orthography could rewrite the phonological representation of the word according to L1 GPC rules. For all of these reasons, it seems that audio-only instruction may be difficult to carry out and any potential benefits may not survive the eventual exposure to orthography.

2.3.2 Explicit GPC instruction and practice

Another intervention that has been carried out is to give the students explicit instruction concerning the difference between L1 and L2 grapheme-phoneme relations and supplementing the lesson with pronunciation practice. Though not en vogue during the peak of communicative

pedagogical approaches, which sought to tap into implicit learning through naturalistic language experiences, explicit pronunciation instruction has more recently regained momentum in SLA research. According to Ellis (2011, p. 35), "Adult acquisition of second language (L2) is a different matter in that what can be acquired implicitly from communicative contexts is typically quite limited in comparison to native speaker norms, and adult attainment of L2 accuracy usually requires additional resources of consciousness and explicit learning." While studies on explicit pronunciation instruction have differed in their methodologies, there is evidence for a generally positive effect of instruction on learner pronunciation (for a review see Thomson and Derwing, 2015, or Lee et al., 2015).

2.3.2.1 Skill Acquisition Theory

Concerning explicit instruction, a highly relevant psycholinguistic framework for learning is presented in DeKeyser (2020) as Skill Acquisition Theory. In this approach the learner is first given explicit instruction concerning the knowledge being taught, and is subsequently led to practice the skill in light of that knowledge, seeking to build up statistical frequency of accurate execution of the skill being learned. Building on the findings of behaviorism, cognitivism, and connectionism, the theory describes three stages of the learning process: declarative learning, procedural learning, and automatization. At the declarative stage, knowledge about the new skill is explained and learned, such that the learner conceptualizes the knowledge related to the skill and can verbalize it, though they may not yet be able to carry out the new skill itself. This knowledge is stored in long-term memory and can be drawn upon via executive control as the learner then applies the knowledge while practicing the skill in the procedural stage. Reaction time and error rates will be relatively high at the beginning of this stage, but these will decrease as frequency of practice increases according to a power law of

learning. The power law describes a mathematical tendency in learning whereby in the first few practice instances a sharp decline in reaction time and error rate is manifested, but as practice continues a very slow but steady decline will ensue, bringing the learner to a stage where executive control diminishes, declarative knowledge is relied on less and less, and eventually the learned behavior is automatized.

Central to Skill Acquisition Theory is the learning of explicit knowledge about the skill in the beginning stage, which can help guide the learner as they practice the new skill in order to *proceduralize* it. The interface between explicit (declarative) knowledge and procedural knowledge is what Ellis (2011) refers to as "The Interface Question," which he worded as: "Just how much do explicit learning and explicit instruction influence implicit learning, and how can their symbiosis be optimized?" (Ellis, 2011, p. 36). After summarizing the various research carried out that scrutinized the role of explicit instruction in learning in the three decades prior, Ellis (2011) concludes that what was left standing is something of a "Weak Interface" position, "whereby explicit knowledge plays [a role] in output, with explicit knowledge coaching practice, particularly in initial stages, with this controlled use of declarative knowledge guiding the proceduralization and eventual automatization of language processing, as it does in the acquisition of other cognitive skills" (Ellis, 2011, p. 38).

Applying this to language learning and specifically to L2 phonological acquisition, explicit knowledge would be embodied in the metalinguistic explanations of how L2 phonology or pronunciation patterns work, and possibly even including explanations of how L2 sounds map to L2 orthography, and then practicing those mappings. If the proper mappings were to be proceduralized and even automatized at an earlier stage of L2 acquisition, it could help learners

internalize a more native-like phonological form for words when they are first encountered via reading.

2.3.2.2 Explicit Pronunciation Teaching in L2 Spanish

Offerman (2020) carried out a dissertation study which included a summary of pronunciation instruction in L2 Spanish. In it, she provided a helpful definition of explicit pronunciation as "the 'conscious awareness' of one's own pronunciation… drawing L2 learners' attention to specific phonetic features in their L2, creating an environment in which learners must consciously focus on these features, as well as receiving direct explanations of how phonetic features in the L2 differ from the L1" (Offerman, 2020, p. 19). She identified four common elements to such instruction, namely: 1) transcription learning/exercises (such as using the IPA alphabet), 2) explanations of articulatory gestures, 3) repetition exercises, and 4) corrective feedback. In this section, some studies that have been carried out in pronunciation instruction in L2 Spanish will be presented, while highlighting findings that are relevant to the effects of orthographic input.

One such study carried out with Spanish learners at the university level is Elliott (1997). In it, the author demonstrated that intermediate students in the experimental group, who received a series of 10-15 minute pronunciation lessons during the semester, made significant improvements in pronunciation, whereas a control group who did not receive such instruction over that time period did not show improvement. Of the four elements of explicit instruction presented in Offerman (2020), Elliott (1997) included explanations of articulatory gestures, repetition exercises, and corrective feedback. Elliott does not mention any training in the IPA or transcription exercises, though part of the instruction included relating proper pronunciation to the Spanish graphemes and comparing them with the pronunciation of these graphemes in

English (e.g., “graphemes 'b' and 'v' are pronounced as 'b'; intervocalic 'r' sounds like the 't' in city” (Elliott, 1997, p. 101)). While the experimental group did show improvement in some of the posttest measures, this improvement only showed up in word-reading and sentence-repetition tasks, not in a more informal, free-elicitation task. Elliott explains that this could be evidence in support of theories related to formal vs. informal speech and the different levels of attention that are applied in such language tasks, with high-attention tasks (e.g., word-reading or sentence-repetition) receiving a more nativelike L2 pronunciation and relatively lower-attention tasks (e.g., speaking freely), where learners are focused more on communicating meaning and less on form, experiencing more L1 transfer phenomena. In addition, Elliott (1997) found that while some graphemes did see improved pronunciation, others, such as the intervocalic <b, d, g>, were not helped by the pronunciation lesson.

Returning to the connectionist theories presented earlier, it may be that for word-reading tasks, the learner’s attention to pronunciation is heightened and he/she applies explicit knowledge about the pronunciation as they sound it out - as if they were learning a new word for the first time. Crucially, Elliott (1997) related pronunciation lessons to the L2 graphemes, even drawing upon L1 GPC rules (e.g., his reference to borrowing the pronunciation of “t” in “city”). In terms of the DRC model, it could be that the explicit pronunciation instruction presented in the 15-minute lessons gave the learner usable information in their long-term memory, especially concerning the mapping in the L2 Grapheme-Phoneme Rule system. During word-reading tasks, the learner could then apply this information and force themselves to pronounce words according to the new L2 Grapheme-Phoneme Rule system as they attend to the orthographic input. But when speaking freely and focusing more on the communication of meaning, the learner would by necessity be drawing from their Semantic System in tandem with their Phonological Lexicon,

which had been internalized and rehearsed (likely with L1 transfer effects) for many semesters before taking this intermediate-level course. Thus, Elliott's (1997) findings may indicate that for L2 learners that have already built up inaccurate pronunciation of words, pedagogical interventions need to not only reconstruct the L2 Grapheme-Phoneme Rule system, but also somehow repair the L2 Phonological Lexicon.

Another study on L2 Spanish pronunciation is Lord (2005), who, like Elliott (1997), looked at undergraduate students, though in this case the learners were enrolled in an upper-division Spanish phonetics class. Her study did not have a control group of L2 learners, but rather investigated any within-group changes over the course of a semester in the course, with a control group of native speakers. The course included explanations of articulation, pronunciation practice exercises, and IPA-transcription. Additionally, the course had some sessions which incorporated voice analysis software, where students could view spectrographs that gave visual feedback concerning voice-onset time (VOT) and the difference between continuant and occlusive realizations. At the end of the semester, the group showed significant improvement in all of the target areas (VOT of /p, t, k/; trill /r/; diphthongs; and [β, ð, ɣ]), as measured by their reading of a short paragraph. Amidst many other lessons and activities, the students read the paragraph three times during the course of the semester for part of a self-analysis project that contributed to their grade in the course, though she states that "the reading and the aspects of their own pronunciation of the reading were never discussed or analyzed as a class" (Lord, 2005, p. 561). By the end of the course, the VOT levels for /p, t, k/, as measured by the experimental task, were not significantly different from those of the control group of native speakers. For trill /r/, accuracy improved from 26% pre-instruction to 39% post-instruction. Within-word diphthong production went from 60% accuracy pre-instruction to 86% accuracy post-instruction.

Between-word diphthongs, improved from 3% to 30% accurate pronunciation. And the approximate allophones [β, ð, ɣ] went from pre-instruction 44%, 34%, 28% accuracy, respectively, to 66%, 48%, 53%, all of which were found to be significantly different from their pre-instruction measures.

The results from Lord (2005) indicate that when given enough focused instruction and practice, L2 Spanish learners were able to make significant improvement in their pronunciation abilities, though for many of these sounds, these learners were still far short of native levels of accuracy. In terms of the amount of instruction and practice that was needed to attain this improvement, Lord (2005) does not specify the amount of in-class and outside-of-class time was required, but typical upper-division courses could have 14 weeks of instruction, with perhaps 2.5 hours of in-class time per week, totaling perhaps 35 hours of in-class instruction, all devoted to improved pronunciation of 7 consonant phonemes plus the diphthongs. The fact that the measurements were all done via paragraph-reading means that we are only given a window into changes in the ability to read text out loud, which, in terms of the DRC model, would mostly demonstrate changes related to the Grapheme-Phoneme Rule system. We would need data from more open-ended speaking tasks in order to judge whether any improvements were seen in the Phonological Lexicon or the Phoneme System.

Looking more closely at the target consonant phonemes in Lord (2005) in relation to the grapheme-phoneme dynamics, we can categorize them in the following way. For the changes in VOT of /p, t, k/, these English speaking learners of Spanish had to become aware of the phenomenon of VOT and practice pronouncing the letters <p, t, k> somewhat differently. The Spanish pronunciation of /p, t, k/ is actually very close to the English pronunciation of /b, d, g/ when looking at VOT (both being short-lag pronunciations of occlusives). Thus, these learners

were learning to map an articulation similar to an L1 pronunciation of /b, d, g/ to the letters <p, t, k>, respectively, and to inhibit the aspirated (long-lag) L1 pronunciation of these letters. For trill /r/, these learners had to acquire a foreign, articulatorily difficult sound, and map it to the letter <r> (word-initially) and <rr> (word internally). Lord (2005) did not mention the Spanish tap /ɾ/, which also is mapped to Spanish word-internal <r>, thus making the learning and mapping task for Spanish <r> even more complex. For the approximant [β, ð, ɣ] allophones, learners have the difficult task of mapping two different allophones, an occlusive and an approximant ([b/β, d/ð, g/ɣ]), to each of the graphemes <b, d, g>, respectively. Furthermore, for [β] and [ɣ], the sounds are not familiar to native English speakers, even if they may be pronounced as varieties of /b/ and /g/ in informal speech of some English speakers, and /ð/ is a separate English phoneme, mapped to English <th>.

2.3.2.3 Teaching L2 Spanish <d> for English-background learners

L2 Spanish <d> presents an interesting case of negative transfer for English-background learners. What makes this grapheme especially intriguing is its relation to a mix of both similar and distinct sounds in Spanish and English. In English, the grapheme <d> corresponds to the /d/ phoneme, which is generally realized as a short-lag, unaspirated stop [d] in stressed syllables (e.g., "dinner"), but lenited to a tap (flap) [ɾ] in unstressed syllables (e.g., "medic"), or sometimes even before stressed syllables when crossing a word boundary (e.g., "I'd often go," "Did I do that?") or in rapid speech (e.g., "medicinal," "cadaver"). Herd et al. (2013) reported that based on various studies, flapping of English /d/ occurs between 94% and 99% of the time in post-tonic, intervocalic position.

In Spanish, <d> also corresponds to the underlying /d/ phoneme, though it is realized as a dental fricative/approximant [ð] after vowels and continuous consonants (e.g., /s, z, ɾ/), which

makes [ð] the most frequent realization of the phoneme, and only as a [d] after a pause, lateral, or nasal. Waltmunson (2005) reported that among native speakers of Spanish, spirantization of intervocalic /d/ occurs 99% of the time. Thus, intervocalic /d/ demonstrates a major difference in phonological processes between English and Spanish, with English expressing it as [d/ɾ] pretonically/posttonically, respectively, and Spanish expressing it as [ð].

Considering the process of L2 phonological acquisition for the English-background Spanish learner, we need to analyze the somewhat complex intersection of phonemes, allophones, and graphemes involved with <d>. As presented above, English <d> represents the English phoneme /d/ which is realized most frequently as [d] and [ɾ]. Spanish, however, actually possesses a separate phoneme for /ɾ/, which is mapped to the grapheme <r>. This Spanish phoneme contrasts with /d/, such as in the minimal pair /todo/ 'all' (<todo>) and /toɾo/ "bull" (<toro>). Thus, avoiding the English GPC for <d> being realized as [ɾ] is crucial to avoiding confusion in L2 Spanish.

Another complexity to this L2 phonological acquisition emerges when considering the sound [ð]. In Spanish it is an allophone of /d/ (represented by <d>), but in L1 English, /ð/ is a phoneme (represented by <th>), which contrasts with English <d> = /d/ in various positions in the word (e.g., "<u>th</u>ere" vs. "<u>d</u>are," "la<u>th</u>er" vs. "la<u>dd</u>er," "loa<u>th</u>e" vs. "loa<u>d</u>"). This means that the Spanish learner, in building up the L2 phonology, must reallocate the English sound [ð], dropping its L1 phonemic status and merging it into Spanish /d/ = <d>. Furthermore, they must associate the sound with a new grapheme, Spanish <d>.

One more possible phonological hindrance for learning this new association may lie in the low frequency of L1 English /ð/ in the intervocalic, word-medial context. English /ð/ most frequently occurs word-initially. In fact, in a seminal study on the frequency of occurrence of

English phonemes, Mines et. al (1978) found that this phoneme was by and large associated with just two words: “the” and “that,” which accounted for 50% of all occurrences of /ð/ in their study. Such a high association between a phoneme and just a few functional words was rare, showing that English /ð/ is rather weak in its breadth of use. When /ð/ does occur word-medially, it is usually before vocalic [ɹ] (e.g., “lather,” “leather,” “heather”), and not the canonical vowels. This may help explain Rafat’s (2011) finding that in a word-repetition task, without the presentation of any orthography, novice L2 Spanish learners only reproduced the intervocalic [ð] about half of the time, often substituting other sounds.

Together, the above observations reveal a particular complexity surrounding the acquisition of L2 Spanish <d> = [ð] for learners with an L1 English background, which can be visualized in Figure 2.2. Nonetheless, one thing that makes the learning task easier for the learner is that Spanish GPCs are highly reliable. English has a relatively *deep orthography* (Katz and Frost, 1992), such that the orthographic form alone cannot reliably lead an L2 learner to the phonological form. Spanish, on the other hand, has a *shallow orthography*, with graphemes pointing to a specific phoneme in a near one-to-one relationship. Spanish <d> is thus reliably realized as /d/, meaning that L2 learners can faithfully rely upon the orthography to guide the learner to the correct phonological representation.

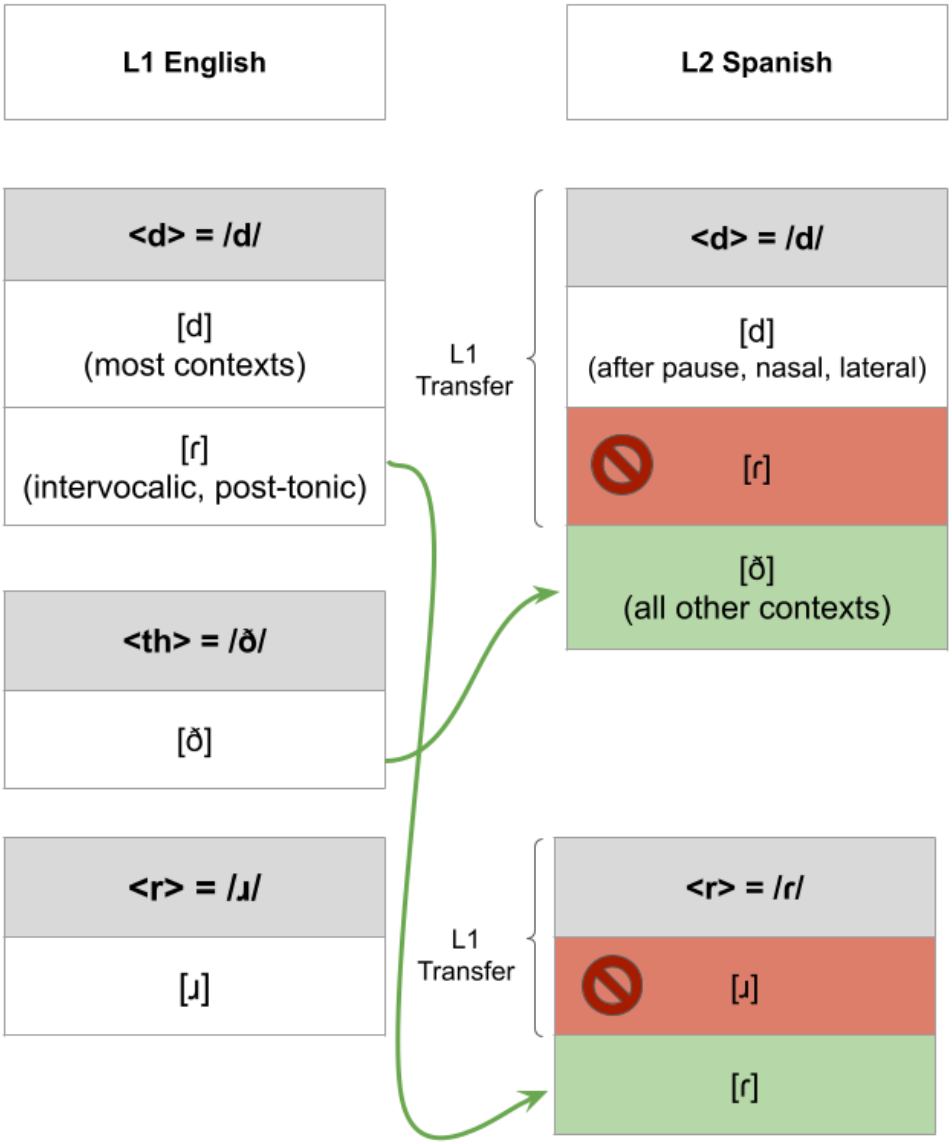


Figure 2.2 Diagram showing transfer and remapping processes related to L2 Spanish <d> for English-background learners.

Concerning explicit instruction for Spanish /d/, Bajuniemi (2013) conducted a study entirely focused on this problematic segment. She looked at Spanish students enrolled in a first-year, accelerated Spanish course at the university level. To enroll in this course, students had to have taken at least two years of Spanish study in high school. She set up her study to compare

two separate sections of the course. One section served as the experimental group, which received some explicit instruction in pronunciation. The other section did not receive any pronunciation instruction, but only received the implicit learning of hearing input from a native Spanish speaker.

The course covered all of the normal areas of Spanish instruction including reading, grammar, and writing practice. For the experimental group, the teacher incorporated a 10-minute lesson on pronunciation into the class, twice a week for 5 weeks (thus, totaling about 100 minutes of pronunciation instruction). The instructor presented a handful of common issues (tap /ɾ/, occlusive /d/, approximant /d/, and orthographic <h>), but the pronunciation lessons focused on approximant /d/. The instruction included: explanations of when and how to make the [ð] sound, repetition exercises as a class and in pairs, and practice in producing and listening to exaggerated target articulations. There is no mention of any transcription exercises. Progress was measured via recordings of a paragraph reading before and after the 5-week lesson period, with productions being coded by raters as either "target" or "non-target" based on the audio playback of the isolated word, as well as viewing the spectrogram.

Results for Bajuniemi (2013) showed that students in the experimental group improved their percent of "target" pronunciations of intervocalic /d/ by 24.1%, while those in the control group saw no improvement after the 5-week period (average of 1.6% fewer "target" productions). In her conclusion, she concedes that measuring pronunciation in a paragraph-reading exercise does not necessarily indicate changes in spontaneous speech, but she argues that these learners will now have the ability to recognize the allophone when they hear it in the future and possibly incorporate it into their spontaneous speech with more practice.

Returning to an interpretation of these findings in light of the DRC model, the improvement in pronunciation seen by these first-year students would primarily be in their Grapheme-Phoneme Rule System, such that they have learned to more frequently pronounce approximant [ð] for intervocalic <d>. This improvement could lead to additional improvement in their Spanish Phoneme System (strengthening [ð] as the correct allophone for intervocalic /d/, instead of [d] or [ɾ]) as well as strengthen associations of native-like pronunciations of words containing intervocalic /d/ with their entries in the Phonological Lexicon and the Semantic System (e.g., [ˈla.ðo] = /lado/ = *lado,* "side"). However, without further testing in other contexts such as spontaneous speech, any claims to improvements in these other systems is only suppositional.

Herd et al. (2013) also studied explicit training methods related to perception and production of L2 Spanish intervocalic /d/ (which in this study is always to be realized as [ð]), along with related phonemes intervocalic /ɾ, r/. Their study was with students enrolled in a university-level intermediate Spanish course and were native speakers of American English. The authors describe a very detailed methodology in which laboratory perception and/or production training were modulated, and compared against a control group which received no training at all. A pretest and posttest were carried out to ascertain their ability both to perceive the difference between three intervocalic contrasts (/ɾ/ - /r/, /d/ - /ɾ/, and /d/ - /r/), and also to accurately produce intervocalic /d, ɾ, r/. Participants were assigned either to one of three experimental groups (receiving training sessions either via perception training, production training, or a combination of the two), or else to the control group, which received no training. Participants in the experimental groups then attended six training sessions. These varied in length depending on group assignment, with perception training sessions lasting 20-30 minutes each, and production

training sessions lasting 60-75 minutes for the first one, and 35-45 minutes in subsequent sessions.

Concerning changes in overall production (of /d, ɾ, r/ combined), the only group that showed significant gains was the combination group, however when /d/ was isolated, none of the experimental groups significantly differed from the control group. Production was measured by simply reading out loud a word on a screen containing the target phoneme. Concerning the type of instruction these participants received, it should be noted that the perception training group did not receive any sort of metalinguistic instruction. Instead, they were simply presented with audio input (e.g., [ˈko.ɾo]) and asked to choose the correct spelling (either <coro> or <codo>), with immediate feedback given after each item. Each session was dedicated to just one contrast (e.g., /ɾ/ vs. /d/). Thus, this group did not get explicit metalinguistic instruction about the articulatory gestures, or even about the L2 sounds they were learning. The production training group, by design, did not receive any audio input, but was only shown waveform and spectrogram visualizations to view the correct sound pattern, along with the orthographic form of the word. They would then record their own speech, and see if their own waveform and spectrogram could replicate those of the native speakers. The combination group received three sessions of perception training, and then three sessions of production training. Concerning the categories of explicit instruction listed in Offerman (2020), the combination group in Herd et al. (2013) received practice repetitions and a form of corrective feedback (either the correct/incorrect for the perception training, or the self-generated waveform/spectrogram data in the production training), as well as some indirect instruction in articulatory gestures in their first production training session, in which a teacher showed them the difference in waveforms/spectrograms of [ð, ɾ, r], with some explanation of why they looked different, though

that was without any auditory input for those sounds. Despite this highly directed acoustical training, which included 6 training sessions focused on these three phonemes, lasting a total of 190-255 minutes over the course of 2-3 weeks, this group saw no significant advantage in /d/ production accuracy compared to the control group.

The difference in results between Bajuniemi (2013) and Herd et al. (2013) may point to the crucial effect of explicit, metalinguistic knowledge in pronunciation instruction, especially with a problematic phoneme such as /d/ for English-speaking learners of Spanish. As Ellis (2011) pointed out, explicit knowledge seems to serve the important function of coaching the individual's pronunciation practice. The participants in Herd et al. (2013) did practice the /d/ sound extensively, always in the presence of L2 orthography, and also got the binary feedback of whether productions were correct or incorrect as well as the visual feedback of their waveforms and spectrograms, but they received very little declarative knowledge to inform them about the phonetic properties of /d/ and orthographic <d> in the L1 or L2. It is notable that even in the simplified task of learning to correctly pronounce items on the word list, they did not show any significant improvement in pronouncing intervocalic Spanish <d> as [ð]. The participants in Lord (2005) and Bajuniemi (2013), on the other hand, did receive explicit instruction concerning the L1 and L2 sounds and orthography, complementing the pronunciation practice exercises and other feedback that they received, and in both cases participants made significant improvements in pronouncing intervocalic Spanish <d>as [ð].

This section has presented the theorized mechanics of explicit instruction as it applies to L2 pronunciation, as well as documented evidence for the effectiveness of explicit instruction in L2 pronunciation. The case of L2 Spanish <d> and /d/ were also presented as one of the more challenging items for English-background learners, with some examples of studies that tested the

phonological acquisition of this L2 item. While explicit instruction has shown some promise as a helpful teaching intervention in helping learners to pronounce this grapheme/phoneme, in the following section another related intervention will be presented, in which instructors have attempted to modify the orthography itself.

2.3.3 Textual enhancement and modified orthography

A somewhat innovative intervention for L2 phonological acquisition that has only recently been subjected to empirical investigation is the use of modified learner orthographies, which either enhance the L2 graphemes with diacritics, or else replace certain graphemes with novel, artificial ones. Since the visual information of the L2 orthography seems to be leaned upon so heavily, these researchers are studying how the visual stimuli might be altered in order to lead learners toward a more nativelike L2 phonology.

One of the first known studies to empirically test such effects is Showalter and Hayes-Harb (2013), who investigated whether or not novel orthographic features could help learners associate novel phonological features to lexical items. Looking at naive English-background learners of Mandarin, they tested whether a group that learned words presented with diacritics to distinguish lexical tone would perform differently than a group that learned the same words with the same pinyin characters but without the diacritics. The crucial test items were a set of nearly homophonous monosyllabic words that were pronounced with the same segments and stress but with different lexical tones. During the word-learning phase, in which all participants would see a picture and hear the word spoken, the experimental group would see the pinyin text along with diacritic marks above the central vowel, which distinguished (and somewhat exemplified) the tonal shape of the vowel. The comparison group, on the other hand, would see pinyin text without any diacritics, and thus these near homophones, though distinct in their lexical tones,

would appear exactly the same orthographically. After the word-learning phase was completed, participants were given a picture-auditory form matching test in Experiment 1, and an orthography-auditory form matching test in Experiment 2 (which was a separate study with a somewhat similar experimental setup as Experiment 1 but composed of a different group of participants).

In Experiment 1, the experimental (diacritic) group outperformed the comparison group in matching the pictures with their auditory forms, indicating that the diacritics had helped these learners more accurately discern the tonal features of the L2. In Experiment 2, after having passed through the word-learning phase, participants were tasked with matching the auditory form with the diacritic-enhanced pinyin. The comparison group had not seen the diacritics before, but rather only the plain pinyin, and thus they were guessing which auditory forms corresponded to the various diacritic-enhanced orthographic forms. They were predicted to perform at chance level, which was indeed borne out in their results. The experimental group, on the other hand, had been presented with the diacritic-enhanced pinyin along with the auditory forms and pictures during the word-learning stage. This group did perform above chance in matching the auditory forms with the diacritic-enhanced pinyin forms, demonstrating that there had been some association of the tonal pronunciations with the related diacritics. The results of the two experiments demonstrated that modifying the orthography with novel features assisted the learners in associating a novel phonological feature with the words.

A psycholinguistic analysis of Showalter and Hayes-Harb (2013) might suggest that the learners who saw diacritically differentiated orthographic forms noticed the novel orthographic features and also discerned new L2 differences (i.e., different tones) in the auditory forms. To some extent, they were able to correctly relate some of those diacritics to the distinct tones that

they were associated with. Thus, they were mapping entirely new phonological features to modified (via diacritics) orthographic forms, seemingly building new GPCs that encoded L2 phonological features with modified L2 versions of their already familiar L1 orthography.

The question arises as to why diacritics could help learners in their perception of L2 sounds. Hayes-Harb and Barrios (2021), in their review of the effects of orthography on L2 phonological acquisition, gave a simple but helpful definition of orthographic input, stating: "orthographic input…is the input that is available to learners in written form" (Hayes-Harb & Barrios, 2021, p. 297). Diacritics give visual information that, if it communicates linguistic information, can be considered additional linguistic input for learners. Solier et al. (2019), point out that while learners may encounter various auditory input of a word, the visual orthographic trace of the word acts as "a permanent external memory" which learners can refer to (Solier et al., 2019, p. 1460). Furthermore, the alteration of already familiar graphemes may heighten the learner's consciousness and attention. In other words, the fact that the learner both recognizes a known grapheme, and also notices that it has been modified, may cause the learner to increase attention resources, and thereby change their processing of the item. Such a phenomenon was borne out in the results of Alsadoon and Heift (2015), whose eye-tracking study showed that by textual enhancement in the form of bold font and a contrasting red color for certain problematic vowels, Arabic-background learners of English fixated on enhanced vowels about 13 times longer than the unenhanced version.

In terms of the DRC model, the results of Experiment 2, which isolated the association between the diacritics and the auditory form, points to L2 phonological learning related to the Grapheme-Phoneme Rule System. The results of Experiment 1, however, which isolated the association between the auditory form and the picture, point to more accurate entries in the

Phonological Lexicon, since the experimental group was more accurate than the comparison group at identifying pictures (the semantic forms) based on auditory input. By implication, it is also thus possible that the learners in the experimental groups were forming a more accurate Phonological System than those in the comparison groups.

The same researchers followed up with another study (Showalter & Hayes-Harb, 2015) that investigated similar questions of how learners relate new phonological contrasts to new orthographic forms. In this study, learners from an L1-English background were presented with Arabic minimal pairs that contrasted in one phoneme: /k/ vs. /q/. Like the tone-differentiated vowels in Showalter and Hayes-Harb (2013), the phonemic contrast here is outside of the L1 phonological categorization system (as English does not have an uvular plosive that contrasts with a velar one). An additional difference in Showalter and Hayes-Harb (2015) is that learners were presented with Arabic script, with entirely different graphemes and a different, right-to-left directionality. In this way the researchers tested whether learners could relate new phonemic contrasts with a completely novel set of graphemes. In Experiment 1, two groups saw pictures representing the words and heard the auditory forms, with the experimental group seeing the genuine Arabic script for the word and the control group seeing a meaningless <ططط> on the screen, roughly analogous to <XXXX>. Crucially, the Arabic script does differentiate the target phonemes with two distinct graphemes (<كـ> for word-initial /k/, <قـ> for word-initial /q/). Participants in the experimental group were not explicitly told anything about the new script they were reading from (e.g., that it is read in the opposite direction of English). Results showed no difference between the control and experimental groups, demonstrating that the learners had not benefited from the Arabic OI. In Experiment 2, the researchers gave participants a brief explanation of the new directionality of the Arabic script, to see if learners would pick up on the

Arabic GPC after being explicitly told where the first letter of the word appears in Arabic (which was the exact location of the contrasting phoneme in the two words). However, the results still showed no benefit for OI in helping discern the target graphemes. In Experiment 3, the researchers substituted a familiar Roman script, with the target phonemes /k/ and /q/ represented by <k> and <q>, respectively. These learners were not given any sort of explicit instruction concerning the new L2 associations of graphemes with phonemes. Again, no benefit was found for the OI, and, in fact, participants in Experiment 3 who were given the modified orthographic forms (<k> = /k/ and <q> = /q/) performed worse than participants who carried out the picture-matching without seeing any orthographic forms at all.

In their discussion, Showalter and Hayes-Harb (2015) pointed out that, unlike the participants in Showalter and Hayes-Harb (2013) who were able to associate a new phonemic contrast (lexical tone of vowels) with the diacritic-enhanced pinyin, it seemed that associating new sounds with the novel Arabic script was a much greater challenge for these learners. This may show that for an L2 with a novel script, even if it is alphabetic, there is a greater cognitive demand to associate sounds with an entirely new orthographic system (maybe especially if that script has a different directionality). They also point out that this phonemic contrast, a velar versus uvular plosive, was a particularly difficult one for L1 English participants to distinguish. Finally, it seemed to them that the graphemes <k> and <q>, which are both used to represent the velar plosive /k/ in English, was not helpful in leading learners to notice the distinct L2 phonemes, perhaps even erroneously signaling to these participants that they should associate the L1 sounds with these letters.

Jackson (2016) followed up on Showalter and Hayes-Harb (2015) concerning orthographic representations of Arabic /k/ and /q/, asking whether results might have been

different if the phoneme /q/ had been represented either by a completely new character, or else a known, Roman character that had been modified with a diacritic. Thus, learners would not see a grapheme exactly like one in their L1, and perhaps would not experience the interference that L1 GPCs seem to cause on L2 phoneme perception. Jackson therefore ran a similar experiment as Schowalter and Hayes-Harb (2015), testing participants' ability to discern /k/ and /q/, with both groups seeing /k/ as <k>, but with one group (N) seeing /q/ as a new character, roughly resembling <ᒧ>, and the other group (M) seeing /q/ as a modified character - a Roman character with a diacritic, <ḳ>. Jackson also included another treatment variable: explicit instruction concerning the new GPC. Subgroups of participants from the new character group and the modified character group received explicit instruction (+EI) concerning the new sounds they were being asked to discern, along with their correspondence to their respective new or modified graphemes. There were also participants in both orthography groups who did not receive any explicit instruction (-EI). Thus, there were four treatment groups: N + EI, N - EI, M + EI, and M - EI.

The experiment started with a word-learning phase, which is the only phase in which the groups received the variable treatments. They were exposed to a total of 12 words, and over the course of 4 blocks encountered each word 8 times. Participants heard each word's auditory form, saw its corresponding picture, and saw the written form of the word, with either the new or modified orthography for /q/. The explicit instruction had taken place prior to the word-learning phase, for the subgroups who received it. Following the word-learning phase, participants had to score at least 90% on a criterion test in which participants were asked to correctly discern when an auditory form either did or did not match a picture that was displayed. All words were tested, but, crucially, participants were never asked to discern a picture/auditory form pair that differed

by the target contrast (/k/ vs. /q/). Participants who did not pass the criterion test had to repeat the word-learning phase again. Finally, after passing the criterion test, they completed a final test that exclusively tested their ability to discern the target contrast.

Results from Jackson (2016) demonstrated some interesting phenomena. First, irrespective of explicit instruction, the new grapheme group as a whole, those who had learned the /q/ words with <ЛI>, significantly outperformed the /q/ = <ķ> group. This seemed to support the notion that a new L2 sound can be more easily discerned if it is represented by a grapheme with which there is no L1 GPC. In the DRC model, this would be the creation of a new entry in the Grapheme-Phoneme Rule System, allowing for a brand new grapheme, such as <ЛI>, to receive its own phonemic association. In Jackson (2016), even a foreign L2 sound was able to be discerned with the help of such a new GPC.

Since the modified grapheme, the diacritic-enhanced <ķ>, has the L1 <k> as its main constituent, it still has the potential to pull learners towards the L1 GPC of <k> = /k/, which could cause confusion for the learners as they try to discern an L2 pattern of <k> = /k/ and <ķ> = /q/. But another finding was that while explicit instruction did not have a significant effect on discrimination accuracy for the new grapheme group, it did lead to a significant improvement in accuracy for the modified grapheme group. Interestingly, when comparing the two groups that received explicit instruction (novel grapheme plus explicit instruction and modified grapheme plus explicit instruction), there was not a significant difference between them.

Results of Jackson (2016) suggest that when explicit instruction is combined with a modified orthography, it may allow learners to distinguish the sounds associated with the modified and unmodified graphemes (e.g., <k> = /k/, <ķ> = /q/) to the extent that they perform just the same as learners who learned a sound association with a completely novel grapheme

(e.g., <k> = /k/, <ЈI> = /q/) . It is as if the explicit instruction gives the diacritic some phonetic information - at least enough to suggest to the learner that they should discern the diacritic-enhanced grapheme with its related L2 sound from the plain grapheme with its L1 (and L2) sound. This seems to somewhat correspond with the benefit experienced by learners who saw the diacritic-enhanced graphemes in Showalter and Hayes-Harb (2013). While Jackson (2016) did demonstrate the effectiveness of using a novel grapheme, if that grapheme is completely artificial, it is likely that any benefit would be lost once the learner interacts with the real orthography.

The combined findings of Showalter and Hayes-Harb (2013), Schowalter and Hayes-Harb (2015), and Jackson (2016) demonstrate in their perception studies that when an L2 has a new phoneme or phonemic contrast, beginning learners may struggle if those phonemes are represented by graphemes already present in the L1 that do not relate to contrasting phonemes. Of course, Vokic (2011) and Rafat (2011) showed that even if the learner is not trying to learn a novel sound, but is simply required to associate an already familiar sound that has an L2 GPC that differs from the L1, they will experience a considerable challenge in doing so. It was interesting that Vokic (2011) found that if the problematic grapheme (e.g., <t> = [ɾ] for Spanish-background learners of English) appeared as a digraph (e.g., <tt>), the learner was more successful in associating the correct L2 sound with the grapheme - suggesting that the learner processes not the grapheme alone but perhaps the visual context around it. Furthermore, Showalter and Hayes-Harb (2013) and Jackson (2016) demonstrated that when diacritics are added to an L2 grapheme that would otherwise be identical to an L1 grapheme, it can give the learner a visual cue that helps them associate the L2 grapheme with the L2 sound, especially if explicit instruction concerning the L2 phoneme is also provided.

Collectively, the body of research presented above does allow us to make some general categories related to the nature of the interaction between L1/L2 orthographic systems:

1) DIFFERENT: If an L2 uses an orthographic system that is different from the L1, forming new L2 GPC rules may take more time and L2 OI will not lead to either positive or negative transfer. (e.g., L2 Arabic, L1 English [Showalter and Hayes-Harb, 2015])

2) SAME: If an L1/L2 pair use the same orthographic system, for shared graphemes there will be a tendency both to pronounce the L1 sound when that grapheme appears in L2 contexts as well as insert L1 phonemes into underlying phonological representation of L2 lexical items containing that grapheme. If the L1 and L2 GPC rules are the same for that grapheme (that is, the phonological representations of that grapheme are similar), this will result in positive transfer effects, but if the GPC rules are different, negative transfer is likely. (e.g., L2 German final <b, d, g>, L1 English [Young-Scholten, 2004])

3) SAME + NOVEL: If an L2 uses an orthographic system that is similar to L1 but contains a novel grapheme that is entirely foreign to the L1, the learner may open up a slot for a new GPC between the novel grapheme and its corresponding L2 sound. (e.g., L2 synthetic "Arabic" <JI> = /q/ [Jackson, 2016])

4) SAME + MODIFIED: If an L2 uses an orthographic system that is similar to L1 but contains a grapheme that is a modified version of an L1 grapheme (e.g., an L1 grapheme with a novel diacritic), the learner may learn to associate a non-L1 phonetic realization with the L1 grapheme (e.g., L2 <ḳ> = /q/ [Jackson, 2016]).

If (3) and (4) above are artificially inserted as a pedagogical tool, questions of authenticity and ecological validity come to the forefront. While most studies on the effects of OI

looked at authentic L2 orthography, a handful of them considered artificial orthographic items, whether an entire romanized script such as pinyin in Showalter and Hayes-Harb (2013), the creation of a completely artificial grapheme for Arabic in Jackson (2016), or the modified, diacritic-enhanced graphemes in Showalter and Hayes-Harb (2013) and Jackson (2016). These studies demonstrate how educators could purposely alter the L2 orthographic system in order to take advantage of positive L2 transfer (as in pinyin for Mandarin - bringing instant GPC benefits), or else to avoid negative L2 transfer (by either substituting completely artificial graphemes in place of the L1/L2 grapheme or else artificially modifying the L2 grapheme). While such modifications could help students have more accurate L2 perception or production experiences, Koda (1992) points out that these benefits are temporary, as grapheme-phoneme relations would have to be reconstructed with the authentic L2 orthography later. This introduces the important questions of whether or not L2 phonetic learning with an artificial L2 orthography lead to lasting phonological developments, and how the hand-off between the artificial and authentic orthographies should be handled. Furthermore, Showalter and Hayes-Harb (2013) and Jackson (2016) looked only at learners' perception of L2 graphemes, leaving open the question of whether or not the benefits seen in using artificially modified L2 graphemes also extends to L2 production. Finally, these two studies looked at the acquisition of new L2 phonological material (lexical tone and a uvular plosive for L1 English), but many L1-L2 pairs mentioned in the literature deal with L2 GPCs that do not introduce new sounds, but rather involve the remapping of existing L1 sounds to different L2 graphemes.

2.4. Objectives and research questions of the present study

The effects of orthographic input on L2 phonological acquisition are still far from being fully understood, and research on pedagogical interventions related to those effects are nil. Thus,

based in a connectionist framework, the present study seeks to increase understanding of the possible benefit of explicit instruction and diacritic-enhanced texts on learners' L2 pronunciation for graphemes with incongruent L1/L2 GPC. Explicit instruction in this case will include both metalinguistic explanation of the target graphemes and phonemes, as well as practice and repetition to increase the statistical frequency of target pronunciations, as described in connectionist theory. Specifically, we will extend previous research on the acquisition of L2 Spanish /d/ by English-background learners and the phonological process of spirantization[1] in intervocalic context. This segment is particularly interesting because the target allophone of [ð] is already part of the L1 phonetic inventory, and thus the learning task involves remapping incongruent L1/L2 GPCs, or, more specifically, remapping the allophonic distributions of a mistakenly transferred L1 /d/. A number of researchers have pointed out that pronunciation errors that impact intelligibility should be the focus of pronunciation instruction (e.g., Thomson and Derwing, 2015; Colantoni et al., 2021). Intervocalic Spanish /d/ has been cited as a specific case in which mispronunciation as [ɾ] by English-background learners can have real communicative consequences (e.g., Hualde, 2005, pp. 150-151; Colantoni et al., 2021). Thus, if pedagogical aids to the acquisition of the proper pronunciation of intervocalic /d/ can be found, they could lead to real improvement in L2 communication. In order to find pedagogical aids with high ecological validity, pronunciation instruction will be kept brief, so as to be realistically included in normal instructional curricula.

The following specific research questions are put forth concerning incongruent L1/L2 grapheme-phoneme associations:

[1] For the sake of simplicity, in this pedagogical study only the [ð] realization was taught for intervocalic <d>. Such a generalization comes at the cost of appreciating variation, and the merits of possibly including variation are considered in 5.6.

1. Does the use of diacritics help learners to more frequently pronounce the target sound during the word-list-reading task in the pronunciation lesson compared with learners interacting with authentic orthography?

2. Does a short lesson with explicit instruction concerning intervocalic <d>, without any diacritic enhancement, help learners to pronounce more of the target [ð] in subsequent paragraph-reading tasks?

3. Does a short lesson with explicit instruction concerning intervocalic <d>, including diacritic enhancement, help learners to pronounce more of the target [ð] in subsequent paragraph-reading tasks which also include a diacritically enhanced <d>?

4. If there are any benefits gained by the inclusion of diacritics in learning/practice, do those benefits persist after diacritics are removed?

Hypothesis (1): Because reading from a word-list allows learners to focus on segment pronunciations, it is predicted that participants who receive a pronunciation lesson will generally produce target-like [ð], and that the inclusion of the diacritic will help learners to do so more frequently.

Hypothesis (2): Because explicit instruction has been shown to generally help learners' pronunciation, it is predicted that participants will produce more target-like [ð] on paragraph-reading tasks after the lesson.

Hypothesis (3): It is predicted that participants who received the pronunciation lesson with diacritics and then carry out paragraph-reading tasks with diacritics will improve [ð]

production rates due to the help of the visual cue reminding participants to apply the pronunciation lesson held in their long-term memory.

Hypothesis (4): It is predicted that for those participants who were reading with diacritics, since the visual cue will no longer be present, once they are removed there will be some diminishing of [ð] production, but not all the way to pretest levels, since due to some statistical frequency, some participants may have started to internalize a more accurate L2 allophonic distribution for /d/.

2.5 Summary

This chapter has presented a literature review concerning the effects of orthographic input on L2 phonological acquisition, including both theoretical explanations for these effects as well as pedagogical approaches that take such effects into account. Finally, the present study's objectives to investigate the hypothesized benefits of diacritic-enhanced pronunciation instruction and reading activities were presented. In the next chapter, I will present a methodology that was developed in order to answer the study's research questions and evaluate its hypotheses.

CHAPTER 3: METHODOLOGY

3.0 Overview

The objectives of the research questions are to demonstrate: 1) whether explicit instruction can help L2 Spanish students pronounce intervocalic <d> as [ð], 2) whether diacritics help these students with a target production on subsequent reading tasks, and 3) if the removal of diacritics would also diminish any benefits previously received. An experiment was designed in order to target these questions, as described below. Section 3.1 explains how participants were chosen for inclusion in the study. Section 3.2 describes the procedure and materials designed to document any possible effects related to the study's objectives. Section 3.3 describes how the data were analyzed, and Section 3.4 provides a summary of the chapter.

3.1 Participants

Students were recruited from various sections from the same 5th-semester Spanish course on grammar and composition at the University of Wisconsin, all taught by non-native Spanish speakers. All of these instructors were graduate students at the university whose L1 was English. Non-native Spanish speakers were chosen and controlled for under the assumption that doing so might maintain a certain degree of similarity in the type of input they would receive related to the phonological process in question. There are good reasons to carry out future research on Spanish learners of other levels, especially at the earlier stages of learning. In this study, 5th-semester students were chosen because they fall into the same general category of "intermediate" learners which have been a very common level for previous pronunciation research, giving us a way to compare the approach in this study more directly with other studies. (Admittedly, there is a good amount of variation in proficiency even when all students have landed in the same class

level. In future work, a quantitative proxy for proficiency such as the LEXTALE (www.lextale.com) vocabulary size test might be more suitable for controlling for proficiency.) Additionally, by looking at learners who have already established patterns of L2 pronunciation we can test the effectiveness of this study's treatment in repairing the negative transfer.

The course included a required, multi-part, online activity that was designed to improve pronunciation while reinforcing grammatical structures that were relevant to the course. Students were invited to participate by simply consenting to allow the research team to use their materials from that activity in the present study. A total of 51 students agreed to participate. Participants were asked to complete the Bilingual Language Profile (See Appendix A, Birdsong et al., 2012). This tool asks each participant their age, gender, and level of education, but also a number of questions including various L1, L2, or other language experiences, their level of identification with L1/L2 cultures, as well as personal goals and motivation related to the L2. This tool also computes a Dominance score between -218 (fully Spanish-dominant) and +218 (fully English-dominant), such that a perfectly balanced English-Spanish bilingual would score as 0. A subgroup of 39 participants who indicated they had little or no Spanish experience beyond the L2 classroom were identified as the target population of interest for the present study. The 12 participants who were set aside from the present analysis included 8 heritage Spanish speakers (for these 8, the average Dominance = 34.8) and 4 participants whose L1 was a language other than English (for these 4, the average Dominance = 99.6), for whom Spanish was their L3. This left a more homogenous group that would potentially demonstrate any possible effects for the experimental treatment without covariates related to other L1/L2/L3 language experiences. Among the 39 remaining participants, 7 did not complete or did not properly record a sufficient portion of the assignment and thus were not able to contribute to the analysis, leaving

32 participants. Finally, 7 of these would be removed from analysis due to their already performing at high proficiency levels on the target item (for these 7, the average Dominance = 118.6), leaving 25 participants for the study (for these 25, the average Dominance = 129.3). See Table 3.1 for a breakdown of participant information for those who were included in the study. Based on pretest scores (as described later in the chapter), participants were randomly distributed in a balanced way between three treatment groups. This was done in order to have a control group of learners who would show the behavior of learners who do not receive any explicit instruction on the target items, a comparison group to show the behavior of learners who receive explicit instruction, but without any diacritic-enhanced materials, and an experimental group to show the behavior of learners who received both explicit instruction and diacritic-enhanced materials. Unfortunately, because the number of participants who ultimately had incomplete data was not proportional between groups, group sizes became somewhat unbalanced. To summarize, the three study groups were as follows:

> **Comparison:** (n = 10) This group of participants would receive a true pronunciation lesson for the [ð] sound. Throughout the study they would interact with plain orthography. (Average Dominance = 135)
>
> **Control:** (n = 8) This group of participants would receive a "dummy" pronunciation lesson, for the [β] sound instead of [ð]. Throughout the study they would interact with plain orthography. (Average Dominance = 128.4)
>
> **Experimental:** (n = 7) This group of participants would also receive a true pronunciation lesson for the [ð] sound, but would also interact with a modified orthography, as described below. (Average Dominance = 122)

Table 3.1 Summary of Participant Subject Code, Gender, and Age, and Dominance by experimental group

Group	Participant #	Gender	Age	Dominance
Comparison	3	F	19	129
	7	F	18	137
	14	F	19	108
	16	F	18	123
	18	F	18	147
	35	F	20	119
	39	F	25	171
	41	F	18	147
	47	F	19	124
	49	F	20	148
Control	6	F	19	145
	17	M	18	130
	32	F	19	148
	42	F	19	154
	43	F	18	134
	44	F	18	83
	50	F	18	104
	51	M	18	130
Experimental	1	F	19	137
	8	F	19	131
	12	F	19	101
	13	F	18	117
	20	F	19	116
	28	F	18	131
	31	M	19	121

3.2 Procedure and Materials

3.2.1 Procedure

The asynchronous, online activity was composed of 6 sessions, spaced out at about 1 session per week for 6 weeks. Sessions 1 and 3-6 were all simply "reading" sessions, where the students were asked to record themselves reading an excerpt from a Spanish fairy tale, the text of which was delivered via a customized course website as a .pdf document in 18pt font. Session 2 was the "lesson" session, where each participant would record themselves interacting with a pronunciation lesson, which was modulated for the various experimental groups as will be described in more detail below. To record all sessions, students were told to use the recording function which was built into the custom website. They could do the recording either on their smartphone or another device, such as a tablet or computer. They were asked to find a place to record their sessions where they would feel comfortable to read the story out loud. Unlike a laboratory study, in this study there were no specific controls in place to maintain the quality or consistency of their recording. But having participants record on their own devices, at their own desired time and location, made it easier to get greater participation, and also maintain greater ecological validity to the study, as opposed to having participants come in for laboratory recordings.

Session 1 served as a pretest, not only to get the starting point for each participant's pronunciation of the target sound, intervocalic <d> = [ð], but also to use their pretest scores to compose three balanced groups with similar pronunciation ability. Participants were asked to read the fairy tale as if they were telling the story to someone in-person, and were encouraged to speak with a native-like pronunciation. Although there were a variety of contexts for intervocalic <d> = [ð] in the readings, only the word-medial, post-tonic context was used for this study.

Pronunciation of <d> as [ð] in word-initial (e.g., *la hermosa sonrisa de la ballena*) or word-medial tonic (e.g., *leñador*) contexts were very rare for this population and this reading task. It was found that 7 of the 32 originally identified participants pronounced the target grapheme as [ð] 50% of the time or greater at Session 1, with the remaining 25 participants doing so less than 50% of the time. Those 7 who pronounced <d> as [ð] at a higher frequency were analyzed separately from those who pronounced it at a lower frequency. This allowed us to find out how the typical, non-approximant producing student would respond to the treatment. These 25 participants were placed into 3 groups, each of which would receive slightly different treatments for sessions 2-5.

3.2.2 Materials

The readings at sessions 1, 2, 5, and 6 were selected as the crucial ones for the experiment, constituting a pretest (session 1), an in-lesson sample (session 2), and a two-pronged post-test (sessions 5 and 6). The control group would provide an example of what the typical students from this population would pronounce for <d> over the course of 6 weeks when no specific pronunciation instruction for the target sound is given. The comparison group should provide evidence for any effect of the pronunciation lesson in session 2. The experimental group should provide evidence for any potential effect of including diacritic-enhanced orthography along with the pronunciation lesson, with session 5 giving an example for how this group pronounces <d> while the diacritics are present, and session 6 showing us how this group pronounces <d> after the diacritics, which they had seen for sessions 2-5, are taken away.

The modified orthography that was presented to the experimental group in sessions 2-5 included a diacritic for all instances of the letter <d>. When it was either utterance-initial, or preceded by /n/ or /l/, it had an underlying asterisk (e.g., <Jack, ¡d̰espierta!> “Jack, wake up!”;

<vend̰e> “sell”). In all other instances, it had an underlying tie (e.g., <mercad̰o> “market”; <la leche d̰e su vaca> “its cow’s milk”). Examples of the paragraph-reading task with either authentic or diacritic-enhanced orthography are provided in Appendix B.

The diacritic <◌̰> was used for the approximant pronunciation. It is a unicode character with a unicode of “U+032E”, and it is referred to as the “Combining Breve Below.” The diacritic <◌̪> was used for the stop (plosive) pronunciation. It has a unicode of “U+0359” and is referred to as “Combining Asterisk Below.” “Combining” refers to the fact that when typing out your text, whenever you input this character it combines with, overlays itself upon, the previous character typed. So, if you type: l, a, d, (insert U+032E), o, you will get the output “lad̰o”. There are various ways to insert unicode characters that can be searched for on whatever platform or software the educator wishes to work with. At the time of writing, the website symbl.cc can provide the character for copying and pasting into a document (e.g., https://symbl.cc/en/032E/). Google Docs also has the ability to insert the character via the Insert > Special characters function, where the user can simply type the code “032E” into the search field.

An underlying diacritic was chosen in this study in order to avoid any confusion with the diacritics used in authentic Spanish orthography above the letters, such as <ñ>, <ó>, and <ü>. Furthermore, the shape of the diacritics was chosen to exemplify features of the articulatory gesture, such that the curved diacritic in <d̰> signifies an articulation that avoids the burst of a [d] sound, whereas the asterisk under <d̪> signifies the strong burst of the typical [d]. This specific interpretation of the shape of the diacritics was presented in the 5-minute pronunciation lesson in order to help embed the articulatory gestures in the visual information.

3.2.3 Word lists and targets

Sessions 1 and 3-6 each comprised a portion of a fairy tale written in simple Spanish. They varied in length from 250 - 400 words. For the sessions of interest, namely 1, 2, 5, and 6, all words that contained an intervocalic, post-tonic <d> were identified for analysis. Some of these words that were especially long, complex, or proved difficult for these students to pronounce were not analyzed for this study. Session 1 contained 8 target words, Session 5 contained 11, and Session 6 contained 8. Table 3.2 shows the target words for each session.

Table 3.2 Target word list at each critical session.

Session 1 (paragraph reading)	Session 2 (pronunciation lesson: word list)	Session 5 (paragraph reading)	Session 6 (paragraph reading)
todo	*mide*	*malvada*	*toda*
todas	*lado*	*asado*	*todas*
pescado	*boda*	*cada*	*todas*
pudo	*nada*	*comida*	*vestidos*
cada	*pide*	*todos*	*vestido*
pueda		*cansada*	*comodas*
cada		*malvada*	*vestidos*
puedes		*ustedes*	*vestidos*
		ustedes	
		comida	
		todos	

3.2.4 Explicit Pronunciation Lesson

Session 2 was the only session in which the participants did not simply read a story. Instead, they watched and interacted with a pre-recorded, 5-minute Youtube video (see www.tinyurl.com/spanish-d) that presented, in their L1 English, an explicit pronunciation lesson. For the Comparison and Experimental groups, the lesson provided a metalinguistic explanation for the target sound, whereas the Control group watched an almost identical lesson but for a

different letter/sound. The lesson included an explanation about the similar alphabets of English and Spanish, and then the problem that <d> is not a letter that behaves in the same way between the two languages. Students are introduced to the two main sounds of Spanish <d>, and then led in practicing a few repetitions of a more native-like, approximant sound for intervocalic <d>, borrowing the already familiar sound of English <th> = [ð]. The lesson is highly orthographic in nature, drawing the learner's attention to the orthographic form. See Figures 3.1 - 3.4 for some of the distinct slides used in the Comparison and Experimental group videos.

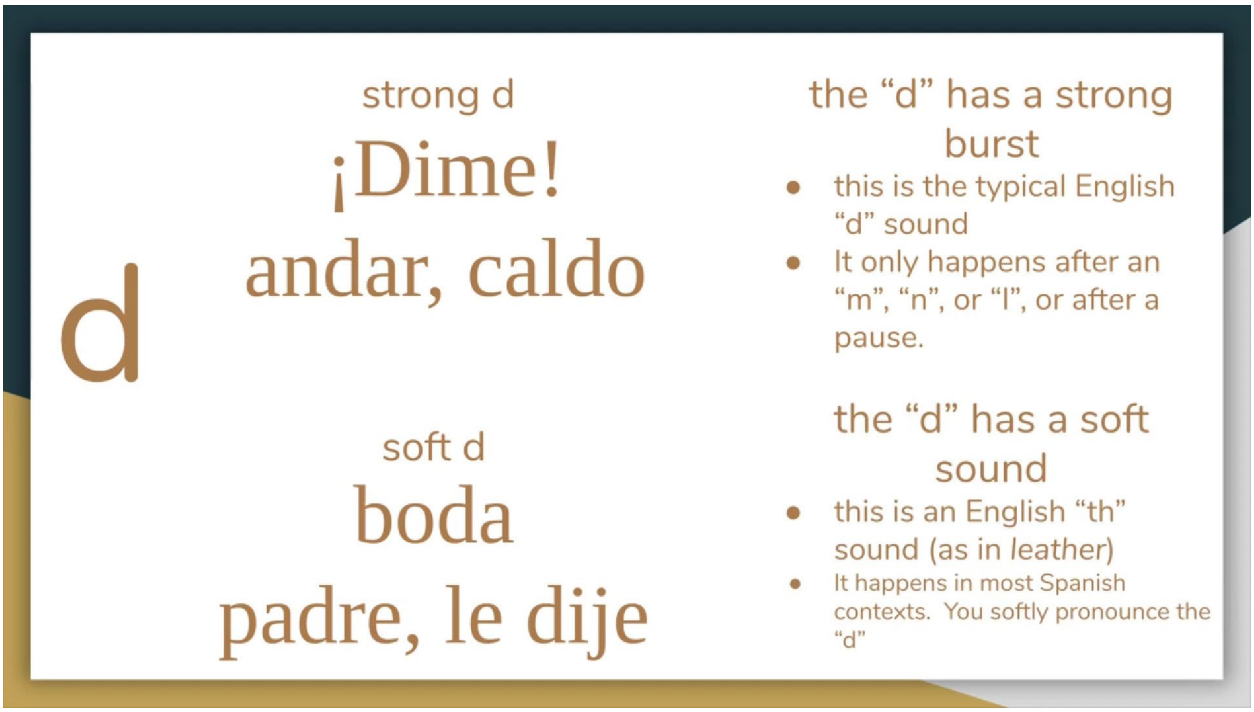


Figure 3.1 Screenshot from comparison group video, introducing the two sounds of Spanish <d>

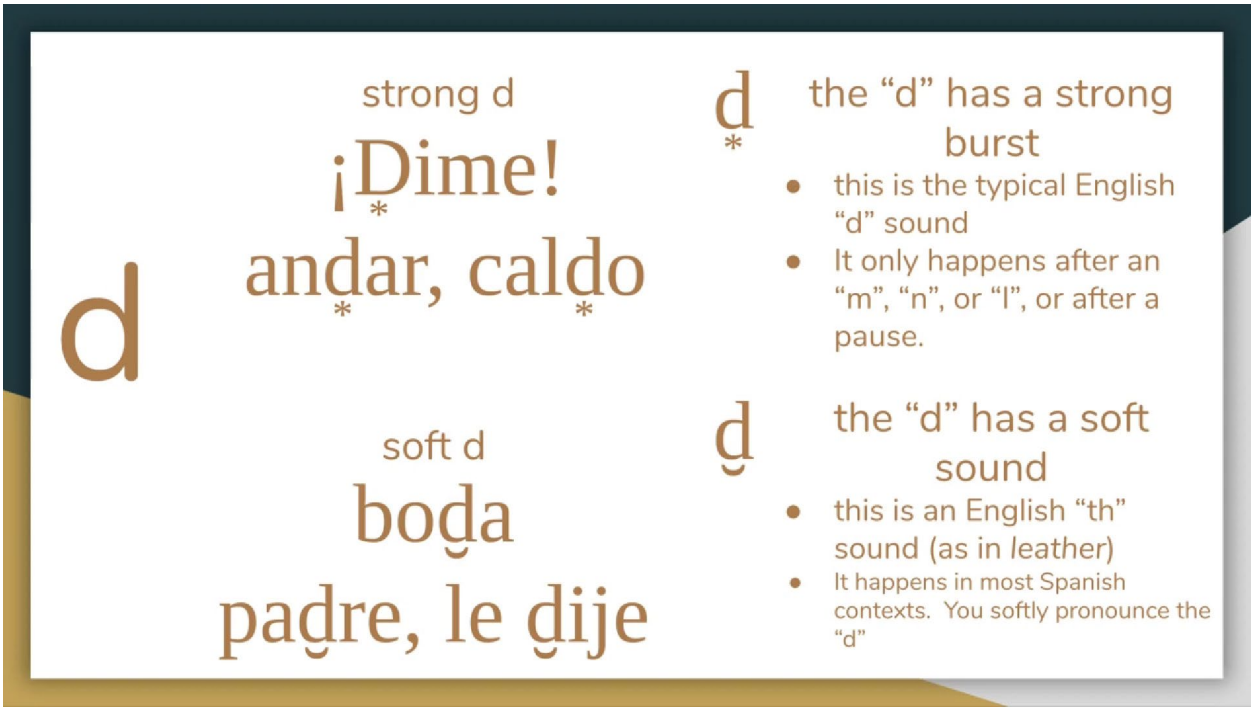


Figure 3.2 Screenshot from experimental group video, introducing the two sounds of Spanish <d> with diacritics to distinguish the sounds.

Figure 3.3 Screenshot from comparison group video, practicing the soft <d>.

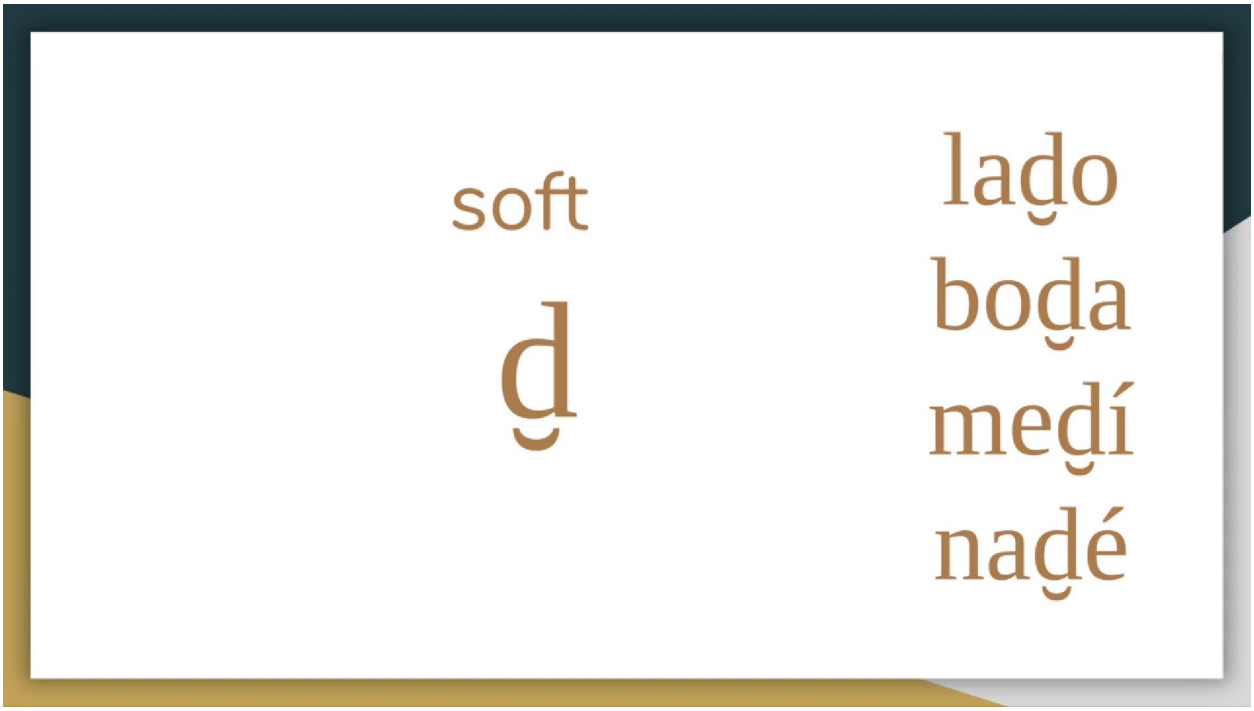


Figure 3.4 Screenshot from experimental group video, practicing the soft <d> with the diacritic.

The students were asked to record their spoken interactions with the video, with many "repeat after me" type exercises. The video concludes with a series of 10 slides, each with a single, two-syllable Spanish word with a word-medial <d> (5 being a tonic /d/, 5 being post-tonic). These words are *mide*, *lado*, *boda*, *nada*, *pide*, *medí*, *nadé*, *cedí*, *nadó*, and *sedé.* The students are asked to read that word, using the newly practiced pronunciation of <d> as [ð]. As can be seen in Figures 3.1 - 3.4, the crucial experimental distinction was that the Experimental group viewed all of the Spanish <d>s in this lesson with a diacritic - either <ḍ> for the [d] sound (though this symbol is only mentioned briefly in the lesson, and the sound was not practiced or tested at all), and <ḏ> for the [ð] sounds. The video for the Comparison group, on the other hand, just used the standard <d> in all instances. A summary of the six sessions can be found in Table 3.3.

Table 3.3 Breakdown of session contents for each group. Each session, whether the story reading or the pronunciation lesson, took about 5 minutes to complete.

	Group		
Session	Experimental	Comparison	Control
1	Read story without diacritic <d>		
2	/d/ pronunciation lesson with diacritics <d̰> / <d̤>	/d/ pronunciation lesson without diacritics <d>	/b/ pronunciation lesson
3	Read story with diacritics <d̰> / <d̤>	Read story without diacritics <d>	
4	Read story with diacritics <d̰> / <d̤>	Read story without diacritics <d>	
5	Read story with diacritics <d̰> / <d̤>	Read story without diacritics <d>	
6	Read story without diacritics <d>		

3.3 Analysis

3.3.1 Acoustic Analysis

For each of the critical sessions, sessions 1, 2, 5, and 6, the target sounds (that is, the word-medial, post-tonic intervocalic /d/ sounds) were identified and each participant's production of the host word was isolated and extracted using the audio editing software,

Audacity (version 2.4.2). These individual sound files were then opened up in Praat (Boersma & Weenink, 2022) for acoustic analysis.

Each segment waveform and spectrogram was visually inspected in order to measure the overall segment duration, which was especially critical in identifying tap productions. The segment onset of the segment was identified based on both waveform and spectrographic evidence, either by a noticeable loss of periodicity in the waveform, an abrupt decline in intensity, or a noticeable disturbance in formant patterns. The segment offset was identified much in the converse way, where the waveform became periodic again, intensity abruptly increased, or formants became solid again. When visual data was inconclusive, auditory evidence was also consulted. The boundaries of the consonantal segment were saved using the annotation function in Praat, and the duration of the segment was recorded for each token.

In this study, we purposely set out to find a more quantitative acoustic procedure for categorizing the various L2 Spanish /d/ productions. Many previous studies have relied on the more qualitative approach of using human listener assessments, which are not only less objective in their analysis, but also require multiple, trained listeners. For other L2 Spanish sounds, such as occlusive /p, t, k/ and b, d, g/, a simple acoustic measurement of the productions, namely VOT, can be very useful in measuring whether learners are approaching a more target-like pronunciation. In order to distinguish L2 productions of intervocalic Spanish /d/, however, a number of acoustic features were needed. In previous research on Spanish spirantization, such as for /b/ in Hualde et al. (2010), one of the main acoustic features used was relative intensity (RI; target consonant versus following vowel). Approximant /b/ has a very small RI, because the articulation of the consonant is voiced and the vocal tract is only slightly constricted. Thus, the change in intensity from the consonant to the subsequent vowel is minimal. For occlusives the

RI will be much higher, since the vocal tract is completely closed temporarily during the consonantal phase, and then there is a sudden release of energy as the subsequent vowel is pronounced. Thus, distinguishing between an occlusive /b/ and an approximant /b/ can be accomplished by simply measuring RI. But for intervocalic L2 Spanish /d/, learners may produce L1 influenced [d] or [ɾ], where [d] is high in RI and [ɾ] is low in RI. Similarly, they may produce target L2 sounds either as an approximant [ð̞] or a fricative [ð̞], where approximant [ð̞] is low in RI and fricative [ð̞] can be higher in RI. Thus, RI could not serve as a single acoustic dimension for discerning target L2 productions of Spanish /d/. While RI could serve as a descriptive feature, two other acoustic properties were found to be the most helpful in discerning L1 from L2 pronunciations: duration and spectral pattern.

This durational measurement was simply the time in milliseconds from the consonant onset to its offset, as described earlier.

To quantify the spectral pattern, each segment was placed into one of three categories by examining salient characteristics of the spectrogram. First, segments in which all of the formants (F1-F4) remained darkened throughout its duration were labeled as "continuous." Second, those segments that included either a vertical black stripe of energy across all frequencies, or else manifested an unusual burst of energy at any frequency were labeled as "stripe/burst." Finally, those segments that demonstrated considerable lightening across all frequencies or whitened spots at any frequencies were labeled as "weakened." Categorization was carried out in that order, such that a spectral pattern of a segment that saw both a lightened phase as well as a stripe or burst would be labeled as "stripe/burst." Examples of spectral pattern categories are given in Figures 3.5 - 3.9.

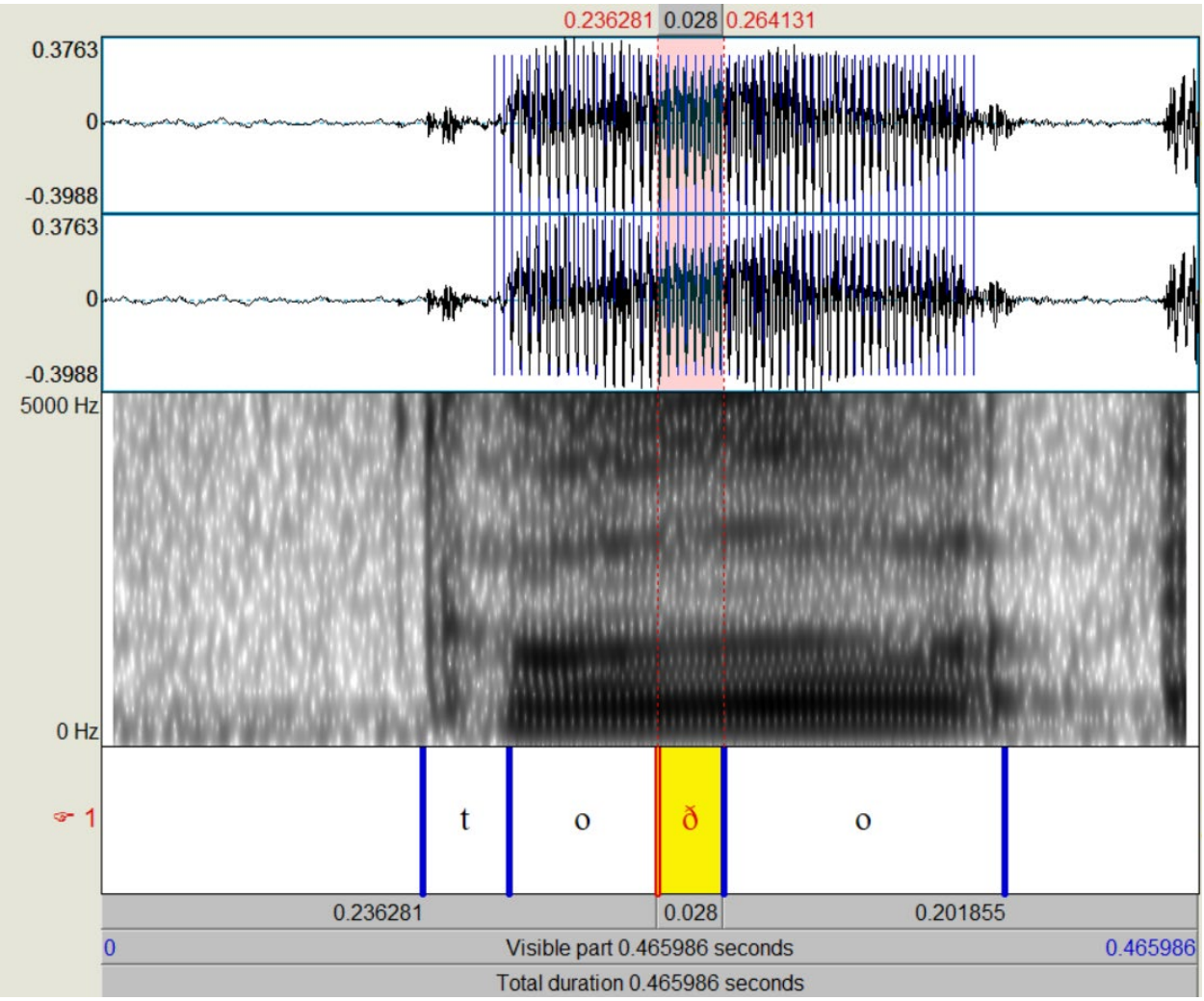


Figure 3.5 Example of a "continuous" spectral pattern for approximant [ð̞] in *todo*, /todo/. Energy at all formants remains strong throughout the segment, approximating the vowels on either side of it.

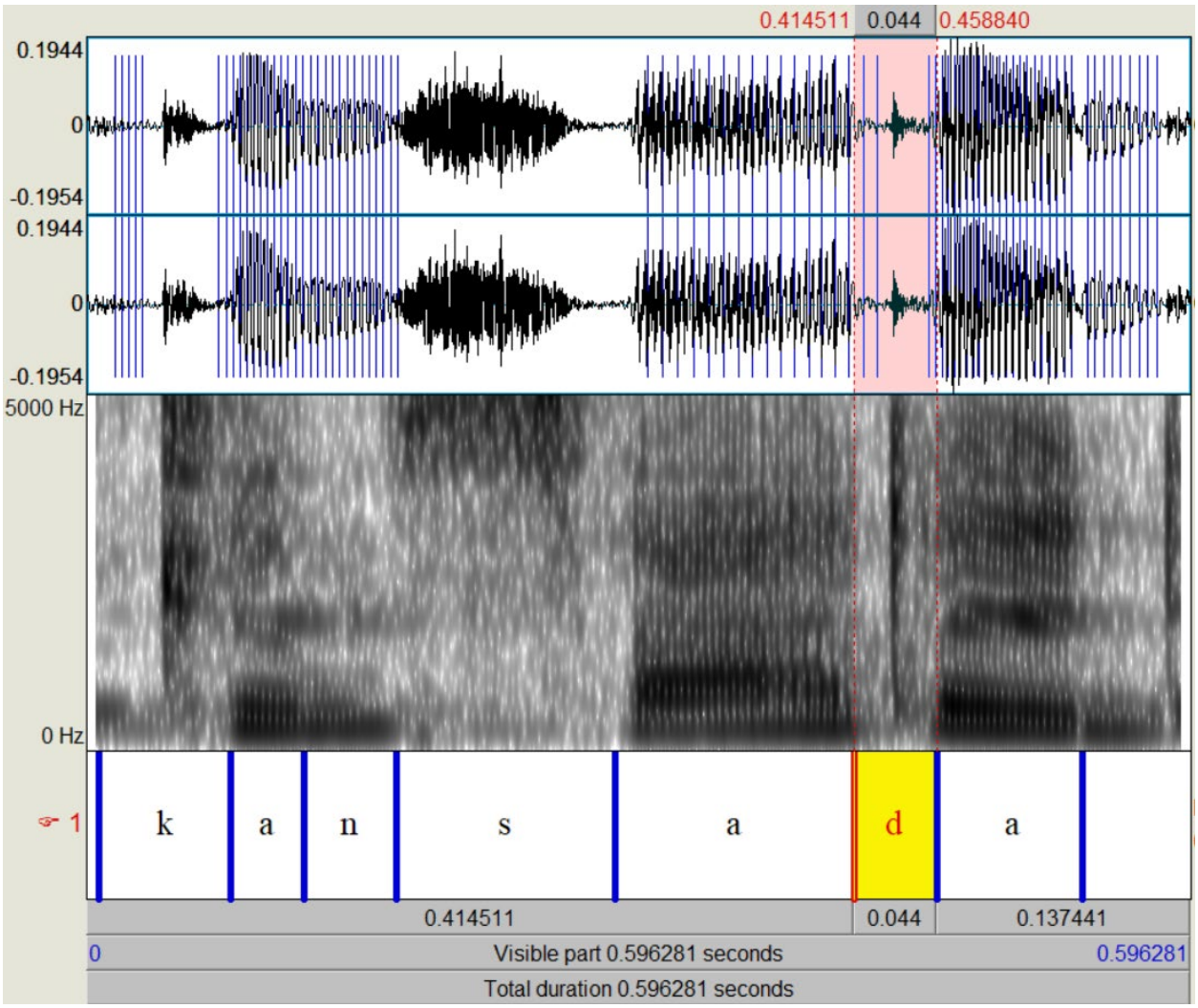


Figure 3.6 Example of a "stripe/burst" spectral pattern for [d] in *cansada*, /kansada/. Energy is greatly reduced until a sudden stripe appears across all frequencies.

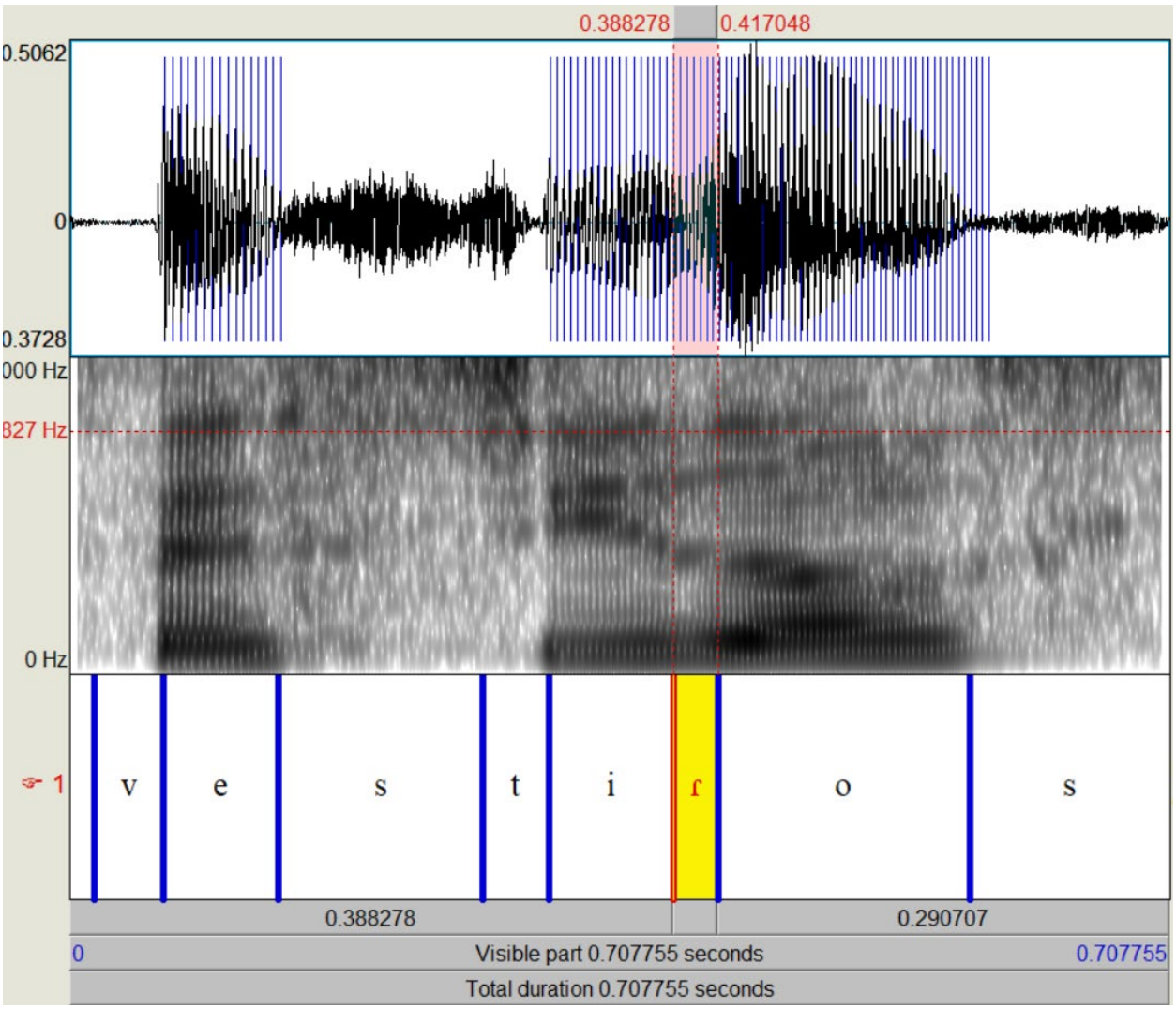


Figure 3.7 Example of a short "weakened" spectral pattern for [ɾ] in *vestido*, /bestido/. While some formants remain visible, there are breaks and significant weakening of energy (especially at F4) in this segment of short duration, distinguishing it from the vowels on either side.

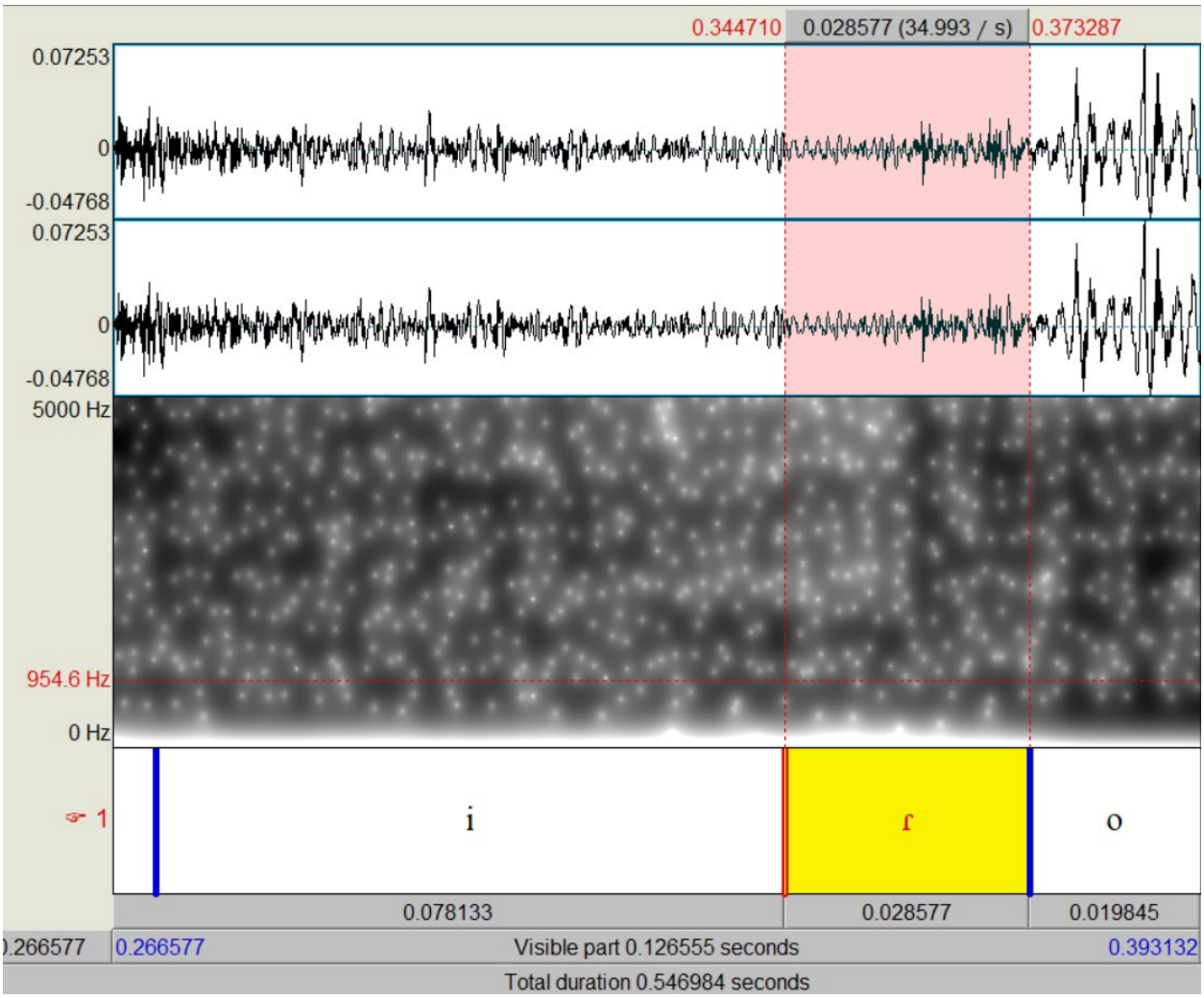


Figure 3.8 Example of a short "stripe/burst" spectral pattern for [ɾ] in *vestido*, /bestido/. There is a sudden, partial stripe in the middle of this segment of short duration, indicating a strong, brief tap before transitioning to the vowel.

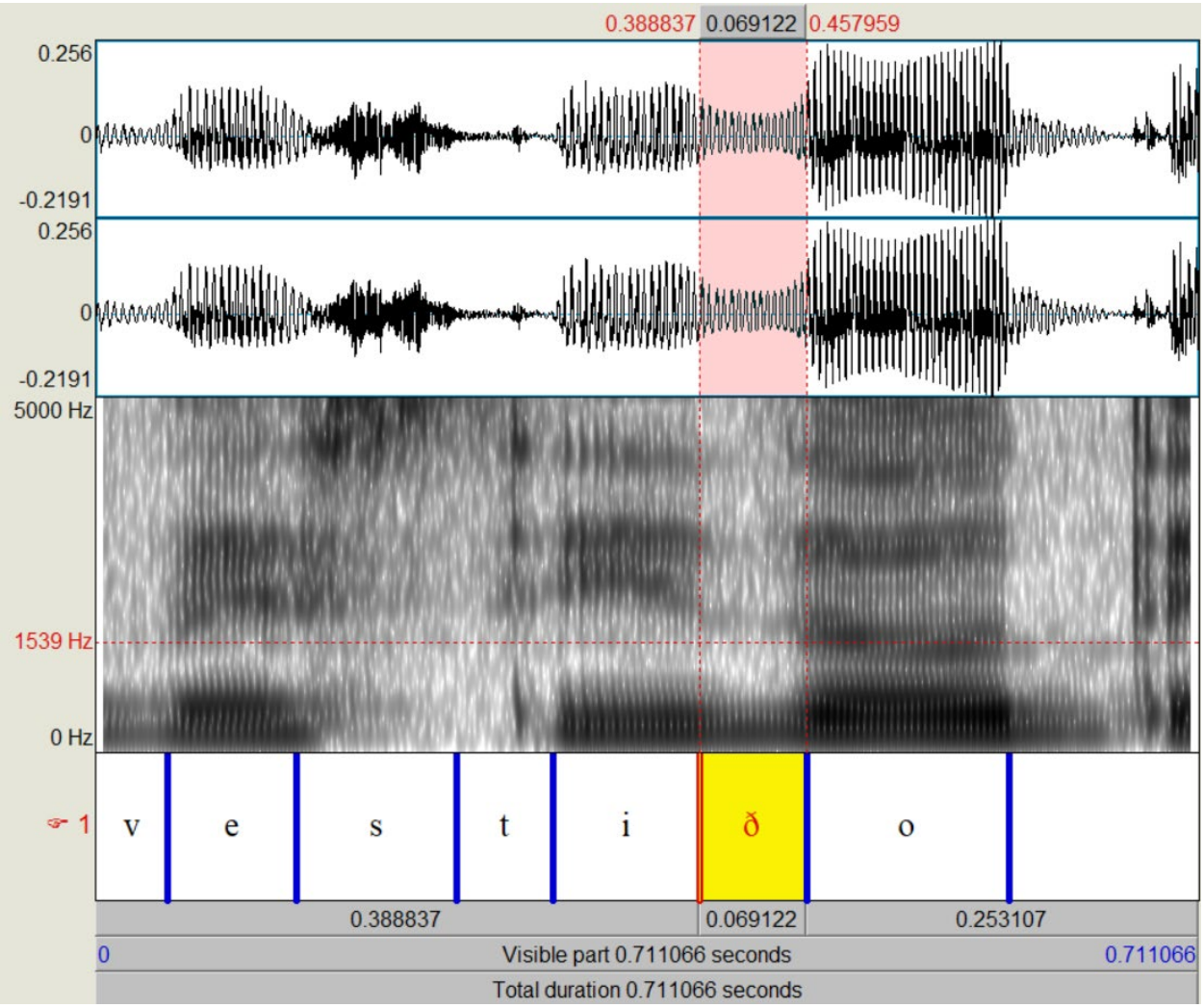


Figure 3.9 Example of a long "weakened" spectral pattern for fricative [ð̞] in *vestido*, /bestido/. Some formants remain visible, though there is a significant weakening of energy, distinguishing the segment from the vowels on either side.

In addition to the boundaries of the consonant segment, the time at the peak of the subsequent vowel was also saved using the annotation function, as indicated by the high point of the intensity line. This gave us a second interval, going from the transition from the consonant to the vowel and ending at the point of the vowel's peak intensity. Once these two intervals were saved for each token, the consonantal interval and the interval of the rising portion of the following vowel, the intensity readings at the low point of the consonantal interval and the high point of the vowel interval were recorded. Rather than use the raw change in intensity (as

measured in dB), a ratio was more desirable due to the great differences in equipment and settings in which recordings were made, as these were not recorded in a lab but by participants on their own devices and in their choice of location. This meant there was no researcher present to ensure that all participants spoke at a similar volume, at a similar distance from the microphone, or even on the same microphone.

The percent increase in intensity (RI%) was derived by taking the quotient of the high point in intensity divided by the low point, then subtracting 1 (e.g., if the minimum intensity during the consonant was 50 dB, and the high point of the subsequent vowel was 75 dB, the percent increase in intensity would be 75/50 - 1 = 0.5, i.e., 50%). This measurement gave us another variable to characterize the /d/ pronunciations in terms of how much constriction took place during the pronunciation of the consonant. A waveform and spectrogram of a student's pronunciation of /mo.do/ can be found in Figure 3.10, with an illustration of how RI and RI% are derived from the intensity measurements.

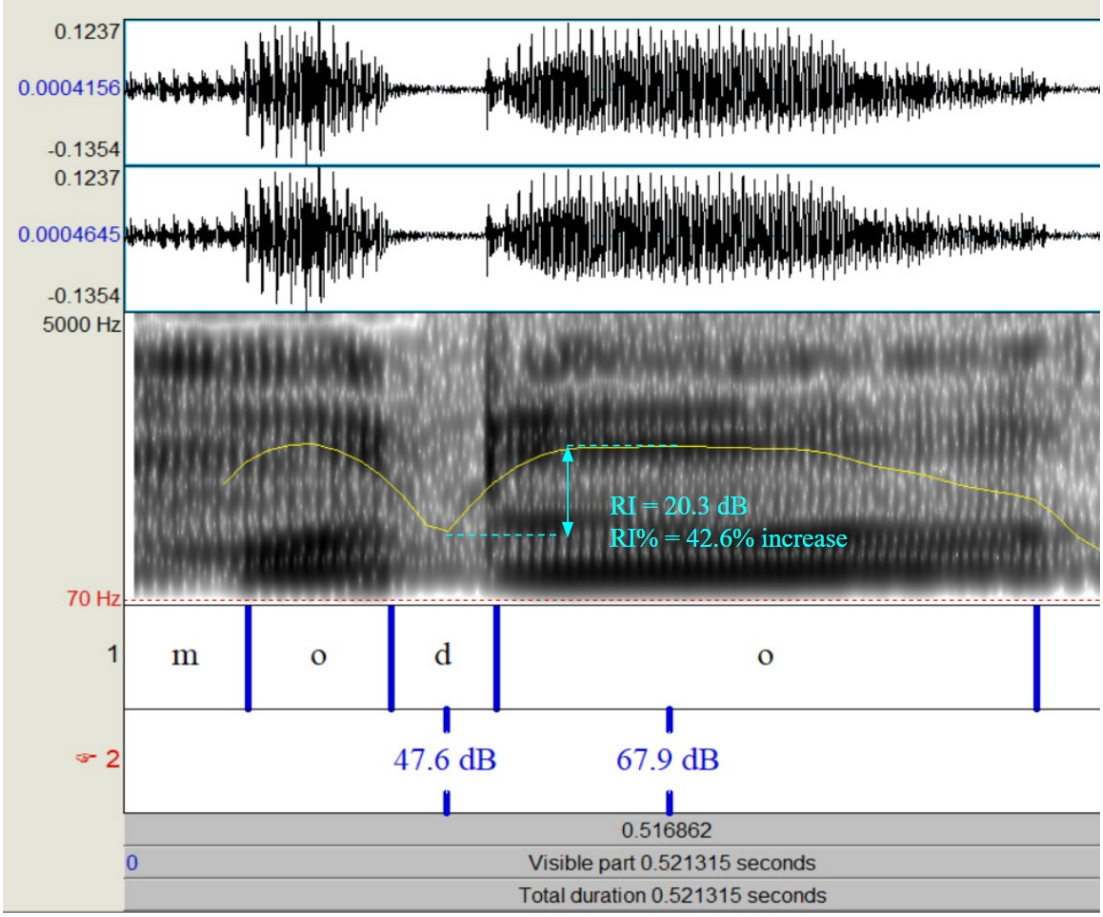


Figure 3.10 Relative Intensity (RI) measurement from a student's pronunciation of /d/ as occlusive [d]. The intensity line is traced throughout the word, and the RI measurement, between the lowest intensity in /d/ to the highest in the proceeding /o/, is shown with the arrows.

Another measurement has been used in the characterization of voiced occlusives, which is a measurement of the maximum rising velocity of change in intensity (Hualde et al., 2010), that is, the greatest positive change in intensity recorded over very brief, 1ms intervals during the release burst (the interval from the low point of intensity in the consonantal segment to the peak intensity in the following vowel). So, comparing this to a measurement of physical movement, if the total change in intensity, from the minimum intensity in the consonantal phase to the peak intensity in the following vowel, were the total distance a car traveled, this maximum velocity is a measure of the "top speed" that was reached at any point during that trip. A Praat script was

designed to measure the change in intensity at each 1ms interval, and the maximum velocity achieved during the release burst was recorded.

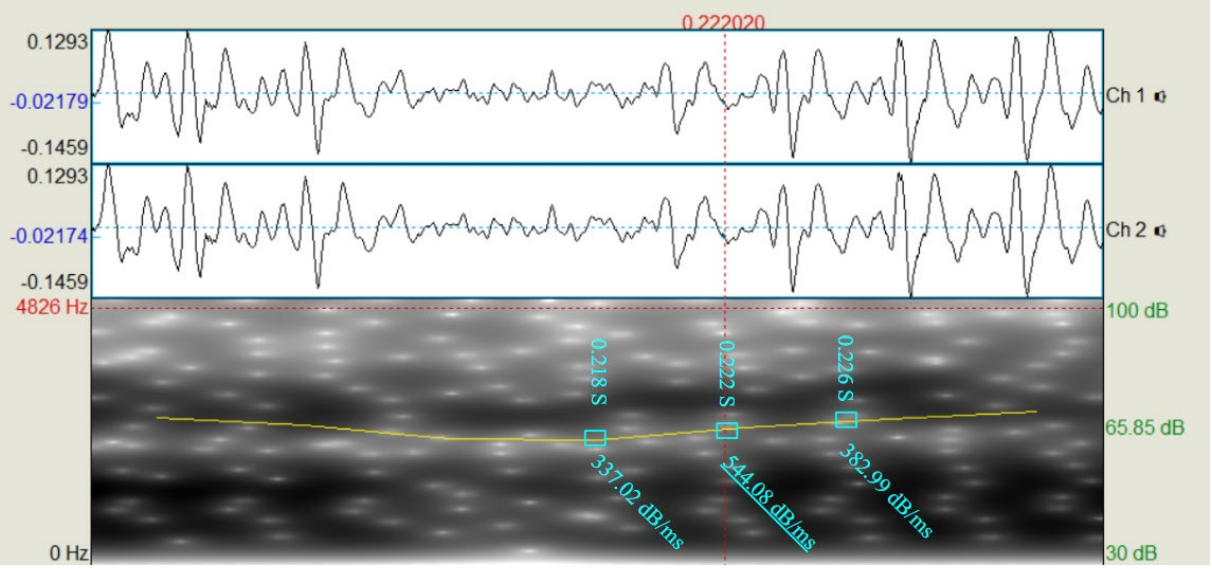


Figure 3.11 Maximum rising velocity measurement. A student's pronunciation of /d/ as tap [ɾ]. The intensity line is traced throughout the word, and three representative 1 ms windows are shown at 0.218, 0.222, and 0.226 seconds within the selection. The velocity, that is the change in intensity per ms, is shown for each window, and the maximum velocity of 544.08 dB/ms is underlined, indicating the greatest "burst" of intensity during the release of this tap [ɾ].

This maximum rising velocity measurement can be helpful in characterizing how much burst (rapid release of energy) was in the pronunciation of /d/, and can be helpful in separating occlusive [d] from approximant [ð̞], as occlusive pronunciations tend to have a large burst of energy whereas approximant pronunciations generally have a very gradual increase in intensity. But because [ɾ] can be pronounced either with a single, clear burst (as captured by short pronunciations appearing with a stripe/burst on the spectrogram) or with a very brief, weakened articulation (as captured by short pronunciations appearing with weakened formants and/or white sections of the spectrogram), the maximum velocity was not effective in discerning taps from approximant or fricative pronunciations of /d/. Fricative [ð̞] pronunciations could likewise vary in the amount of burst exhibited, with some registering lower maximum velocities for more

approximant-like pronunciations, and others with high maximum velocities, approaching similar readings as stops. Because participants' speaking volume and microphone sensitivities were not consistent (whether between participants, or between sessions), the velocity measurement was weighted by the maximum intensity of the particular production.

Table 3.4 provides a summary of the various acoustic measurements that were taken for each production, including their description and how they function to distinguish between various phonetic categories. As indicated above, only the combination of duration and spectral category were able to reliably and effectively distinguish the three needed phonetic categories of approximant [ð], tap [ɾ], and stop [d].

Table 3.4 Summary of acoustic measurements taken in this study

Acoustic measurement	Description	Distinguishing ability
Duration	Duration (ms) of the consonantal segment, measured from consonant onset to offset.	Distinguishes taps and brief approximants (duration < 30ms) from stops, fricatives, and long approximants (duration > 30ms)
Spectral Category	Categorization based on spectrogram of each production, either as continuous (C) , weakened (W), or stripe/burst (S)	Stripe/burst pattern indicates either a stop or tap, weakened pattern indicates either a fricative or tap, and a continuous pattern indicates an approximant
Relative Intensity (RI), and Percent Increase in Relative Intensity (RI%)	RI: Difference in intensity (dB) from low point in consonantal segment (min) to high point in subsequent vowel (max). RI%: RI divided by min, in order to normalize data from a variety of speakers, recording conditions, etc.	Characterized segments with higher changes in intensity, usually due to obstruction or longer articulation (stops, strong fricatives) and segments with lower changes in intensity due to minimal obstruction or brevity (approximants, taps)

Maximum Rising Velocity - weighted (WMRV)	The greatest 1-millisecond increase in intensity (dB/ms), weighted by the maximum intensity of the subsequent vowel.	Characterized segments with higher bursts of intensity, usually due to release from some obstruction (stops, strong taps) and segments with gradual changes in intensity due to minimal obstruction or brevity (fricatives, approximants, perceptual taps)

3.3.2 Data Analysis

Participants' production of all intervocalic, word-medial, post-tonic <d> were analyzed in the three crucial sessions (1, 5, and 6). Praat was used to generate waveforms for each sound. Following a similar study by Hualde et al. (2010), a filter (pass-Hann band from 500 to 10,000 Hz) was applied to each sound, in order to reduce the effects of the f0 band on the intensity readings, thus increasing the sensitivity to slight changes in intensity due to articulation. Based on Warner and Tucker (2011), Praat pitch settings were also adjusted, with the minimum pitch being set to 150 Hz in order to make the intensity curve more fine-grained and thus prevent obscuring the very brief intensity changes for [ɾ].

A subset of 100 tokens was hand-coded to categorize sounds as either stops, taps, or approximants. Within these 100 tokens, specific patterns of duration and spectral pattern were found to reliably correlate with the hand-coded categorizations. Tokens that were shorter than 30 ms and had either a "stripe/burst" or "weakened" spectral pattern were found to consistently have been hand-coded as taps. For an example of a "weakened" tap, see Figure 3.7, and for a "stripe/burst" tap, see Figure 3.8. Tokens that were longer than 30 ms and had a "stripe/burst" had been categorized as stops. Tokens that had either a "continuous" spectral pattern, or were 30 ms or longer and had a "weakened" spectral pattern, were found to have been categorized as

approximants. This formula for categorization, which was over 90% accurate according to the hand-coded sample, was thus applied to the remainder of the data, giving each token an automatically generated categorization. The few tokens where the formula gave a categorization that were not consistent with the hand-coded rating were among the extremely brief pronunciations (less than 20 ms) that showed a "continuous" spectral pattern, but sounded like a tap. These would be very light, very brief taps, that left no evidence in the spectral pattern, which some have labeled as "approximant taps." Because they were not all that common, and because there were no objective methods found to accurately tease them out, they were left in the "approximant" category, but will still be addressed in the Results and Discussion chapters.

3.4 Summary

In summary, this chapter has provided an explanation of the methodology that was set up to test the research questions of the study. By breaking the participants into three groups, we are able to control for various L2 experiences and isolate as much as possible the specific effects of explicit pronunciation instruction and the use of diacritic-enhanced orthography. By using a paragraph-reading task instead of isolated words on a word list, we can test L2 speech that approaches a more natural style where learners are processing the meaning of the words and sentences. The targeted item, Spanish <d>, gives us a window into cases where L2 learners have to reassociate familiar sounds and graphemes. If the experimental features uncover potential helps for L2 phonetic acquisition related to this target, there could be implications of improved teaching methodologies going forward for similar learning tasks for learners from this population.

In the following chapter, the results of the experiment will be presented and comparisons will be possible between the experimental, comparison, and control groups at each of the target

sessions. Session 1 will serve as a pretest for all three groups. Session 2 will show us how the experimental and comparison groups pronounced intervocalic <d> during the pronunciation lesson itself (the control group was not given a pronunciation lesson on the target phoneme). Session 5 will allow us to compare how the <d> pronunciation of the experimental group, viewing the diacritics, differed from the other groups. Session 6 will allow us to compare not only between groups at that point in time, but especially if there was any change in the experimental group once the diacritics were removed.

CHAPTER 4: RESULTS

4.0 Overview

In the preceding chapter, I described the methods employed in the present experiment designed to test (a) the effect of explicit instruction related to grapheme-phoneme mapping, (b) the effect of diacritics on the pronunciation of the target grapheme, and (c) whether pronunciation associated with the diacritic is maintained after the diacritic is taken away. In this chapter, I will report the results and evaluate the hypotheses outlined in Chapter 2. Results are reported mainly concerning approximant production, derived from the duration and spectral category data, as described in 3.3.1. See Appendix C for statistics related to other variables such as Relative Intensity percent increase (RI%) and weighted maximum rising velocity of intensity, which were not useful for reliably discerning approximant production.

The remainder of this chapter is structured as follows. In Section 4.1, I will report on the results of the pronunciation of intervocalic <d> by group, highlighting the effects of explicit instruction, by comparing the control group with the comparison group through the four sessions of interest, as well as the effects of diacritic-enhancement by comparing the experimental group with the other two groups. Both a descriptive analysis of the data as well as a Bayesian analysis will be presented. In Section 4.2, I will report on the individual results in the same manner. Section 4.3 provides a summary for this results section.

4.1 Group results

4.1.1 Descriptive Analysis

The raw data for the pronunciation of intervocalic <d> by group and session can be found in Table 4.1, which is also visualized in Figure 4.1. Sessions 1, 5, and 6 were paragraph-reading

tasks and Session 2 was the treatment session, in which a word-list reading task was carried out within the <d> pronunciation lesson. This lesson was only administered to the comparison and experimental groups, with exactly 5 word-medial, post-tonic, intervocalic <d> targets (the control group received a lesson on a distractor <b> target).

As can be seen there, all three groups show a strong tendency for producing taps at Session 1 (comparison: 86%, control: 83%, experimental: 90%), with only a few stops and approximants produced. In Session 2, which was a word-list task administered right after the pronunciation lesson (the control group was given an alternative lesson at Session 2 and did not attempt <d> pronunciations), the comparison and experimental groups each produced mostly approximants (comparison: 66%, experimental: 57%), although more stops also emerged for both groups (comparison: 18%, experimental: 30%). At Sessions 5 and 6, the control group had pronunciation patterns almost identical to their Session 1 rates, producing almost exclusively taps (Session 5: 86%, Session 6: 91%). From Session 1 to Session 5, the comparison group saw a sharp increase in stops (Session 1: 7%, Session 5: 25%), with a very modest increase in approximants (Session 1: 8%, Session 5: 13%). In Session 6, the comparison group saw approximant production drop even lower than Session 1 rates (Session 1: 8%, Session 6: 5%), but maintained the high level of stop production that was seen in Session 5 (Session 6: 25%). The experimental group, like the comparison group, saw an increase in stop production from Session 1 to 5 (Session 1: 0%, Session 5: 16%), but unlike the comparison group, the experimental group saw a sharp increase in approximant production (Session 1: 10%, Session 5: 34%). As a reminder, the experimental group paragraph at Session 5 included diacritics, which distinguished stop and approximant /d/, while the comparison group paragraph had unaltered orthography. Interestingly, at Session 6, when diacritics were removed from the experimental

group's reading, they still maintained an elevated level of approximant production (Session 6: 37%) which was much higher than that of the control group (5%). A Cohen's *d* effect size of 1.19 was calculated comparing the experimental and control groups at Session 6.

group	participants	session	category	percent	n
Comparison	10	1	approximant	8%	6
			stop	7%	5
			tap	86%	65
		2	approximant	66%	33
			stop	18%	9
			tap	16%	8
		5	approximant	13%	14
			stop	25%	27
			tap	62%	67
		6	approximant	5%	4
			stop	25%	19
			tap	70%	54
Control	8	1	approximant	8%	5
			stop	8%	5
			tap	83%	49
		5	approximant	8%	7
			stop	6%	5
			tap	86%	72
		6	approximant	5%	3
			stop	4%	2
			tap	91%	51
Experimental	7	1	approximant	10%	5
			stop	0%	0
			tap	90%	47
		2	approximant	57%	17
			stop	30%	9
			tap	13%	4
		5	approximant	34%	24
			stop	16%	11
			tap	50%	35
		6	approximant	37%	19
			stop	15%	8
			tap	48%	25

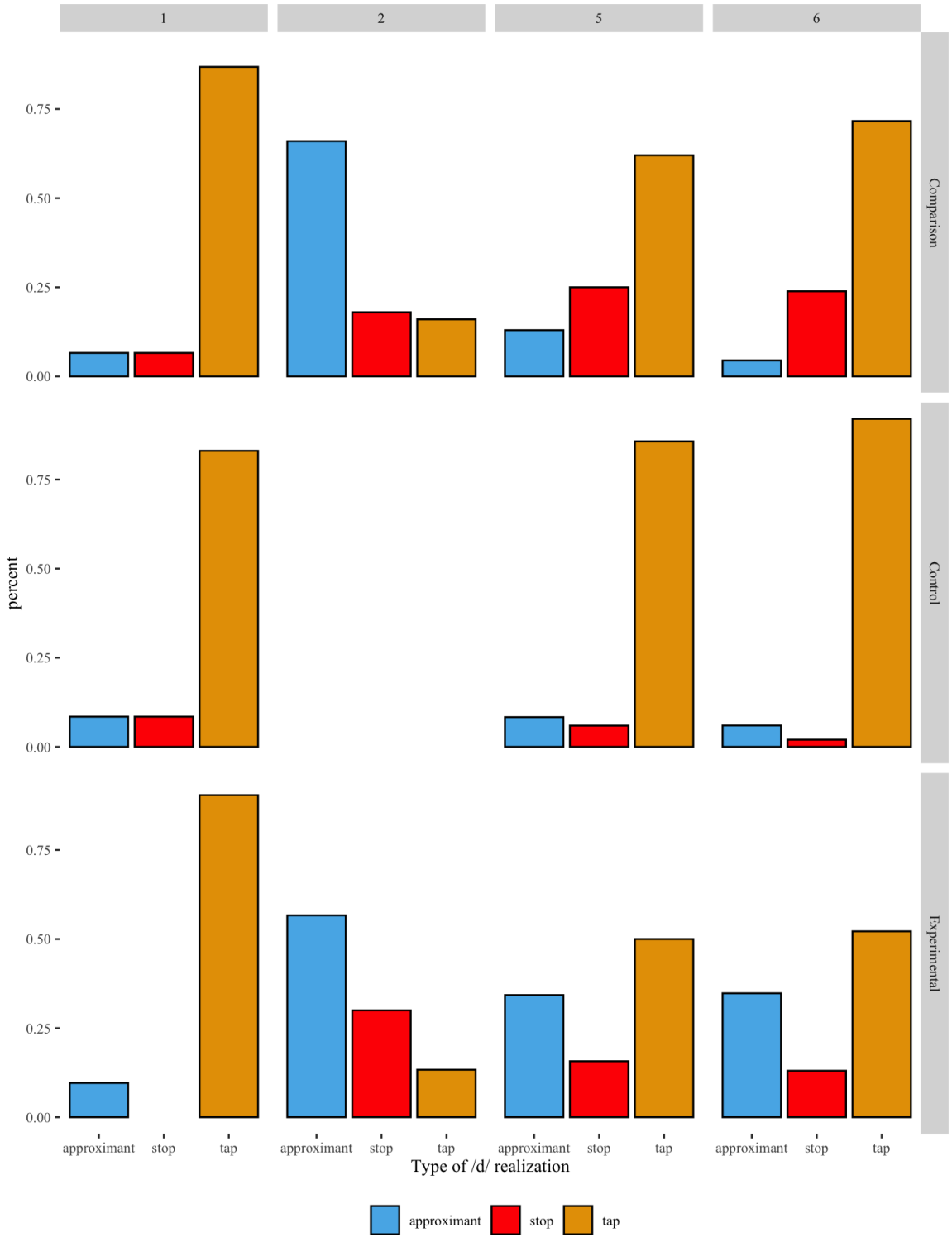


Figure 4.1 Rates of pronunciation by group, session and category.

The patterns of pronunciation can also be viewed in the scatter-plot in Figure 4.2, which shows each token's realization on a graph by group and session. For each group and session, the productions are graphed by one of 3 possible x-axis positions for spectral pattern, either continuous (C), weakened (W), or stripe/burst (S). As a reminder to what was presented in section 3.3, a continuous spectral pattern was one in which the vowel formants were not broken or weakened throughout the consonantal segment, a weakened pattern was one in which one or more of the vowel formants saw significant weakening of energy, indicated by white or faded bands in the formant, and a stripe/burst pattern was one in which the formants saw a either a black stripe of released energy across all formants, or else a sudden burst of energy at any formant. The y-axis of the scatterplot indicates the duration of the segment. In addition to visualizing these acoustic features on the scatter-plot, each production is color-coded to show its categorization as either an approximant (blue), stop (red), or tap (mustard).

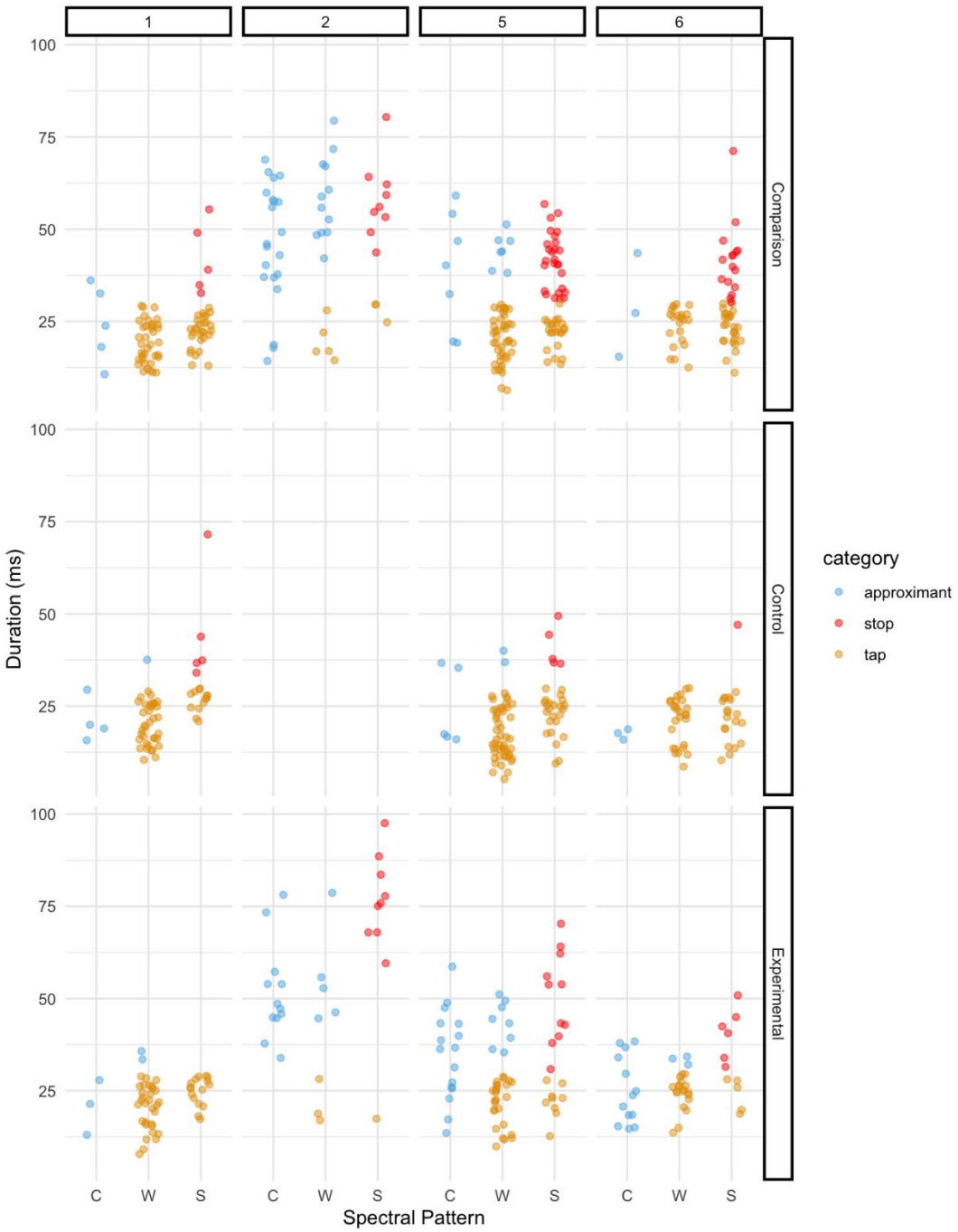


Figure 4.2 Scatter-plot of all target <d> productions by group and session. The x-axis indicates spectral pattern of either continuous (C), weakened (W), or stripe/burst (S). The y-axis indicates the duration. Each production is color-coded to show its categorization as either an approximant (blue), stop (red), or tap (mustard). Due to differences in group size, the comparison group has many more data points per session. Session 5 also had more targets than Sessions 1 and 6.

Some key takeaways from the descriptive analysis by group are that, in general, this group of students was producing almost only taps on the paragraph-reading task at Session 1. This is despite the fact that all participants in this analysis were asked to read in as native-like pronunciation as possible. Both groups that received a pronunciation lesson were able to shift to producing mostly approximants on a word-list task at Session 2, with much longer durations probably due to the slowed down and enunciated pronunciations of a word-list task. For each group, in Figure 4.2 it is easy to visualize their pronunciation patterns across Sessions 1, 5, and 6. The control group remains mostly unchanged across the sessions. The comparison group shows an increase mostly in stop pronunciations, though maintaining a high concentration of taps. The experimental group shows some increase in stops, but also a considerable increase in approximants, alongside a diminished concentration of tap pronunciations. In the next section, a Bayesian analysis of the group data will be presented.

4.1.2 Bayesian analysis

The raw data were analyzed using a Bayesian logistic regression. This particular method is preferable when there is not an especially high volume of data, and when there are unequal numbers of tokens and participants, as it is still able to form a statistical model around such datasets from which probabilistic approximations can be derived. A Bayesian approach is also more robust to small sample size than a frequentist analysis.

A regression model for mixed effects was fit to the data where the outcome variable was the log-odds of realizing the <d> as an approximant, which we converted to probability for the sake of interpretation. Details of the model's parameters can be found in Appendix D. The model's predicted probability for each group to pronounce an approximant at each session can be found in Table 4.2, which lists the estimated probability as well as the 95% credible intervals in

brackets. As can be seen, the model predicts that all three groups had a relatively low probability of producing an approximant at Session 1 (comparison: 4.1% [95% HDI 0.009-0.128], control: 6% [0.013-0.194], experimental: 6.8% [0.012-0.229]). For the comparison and control groups, there was little change in their probability of producing approximants in Session 5 (comparison: 6.4% [0.014-0.208], control: 4.7% [0.008-0.183]) and Session 6 (comparison: 6.4% [0.001-0.084] , control: 4.7% [0.003-0.134]). The experimental group, however, saw an increased probability of producing an approximant at Session 5 (28.2% [0.069-0.626]) and Session 6 (34% [0.094-0.692]). These data can also be viewed graphically in Figure 4.3. The large credible intervals for the experimental group at Sessions 5 and 6 indicate that there is a large degree of uncertainty of the model's estimated value, though the values towards the extremes of the intervals are far less likely than values at the center.

Table 4.2 Probability of an approximant realization by group and session: mean estimate and 95% credible interval

group	Session 1	Session 5	Session 6
Comparison	0.041 [0.009-0.128]	0.064 [0.014-0.208]	0.019 [0.001-0.084]
Control	0.06 [0.013-0.194]	0.047 [0.008-0.183]	0.03 [0.003-0.134]
Experimental	0.068 [0.012-0.229]	0.282 [0.069-0.626]	0.34 [0.094-0.692]

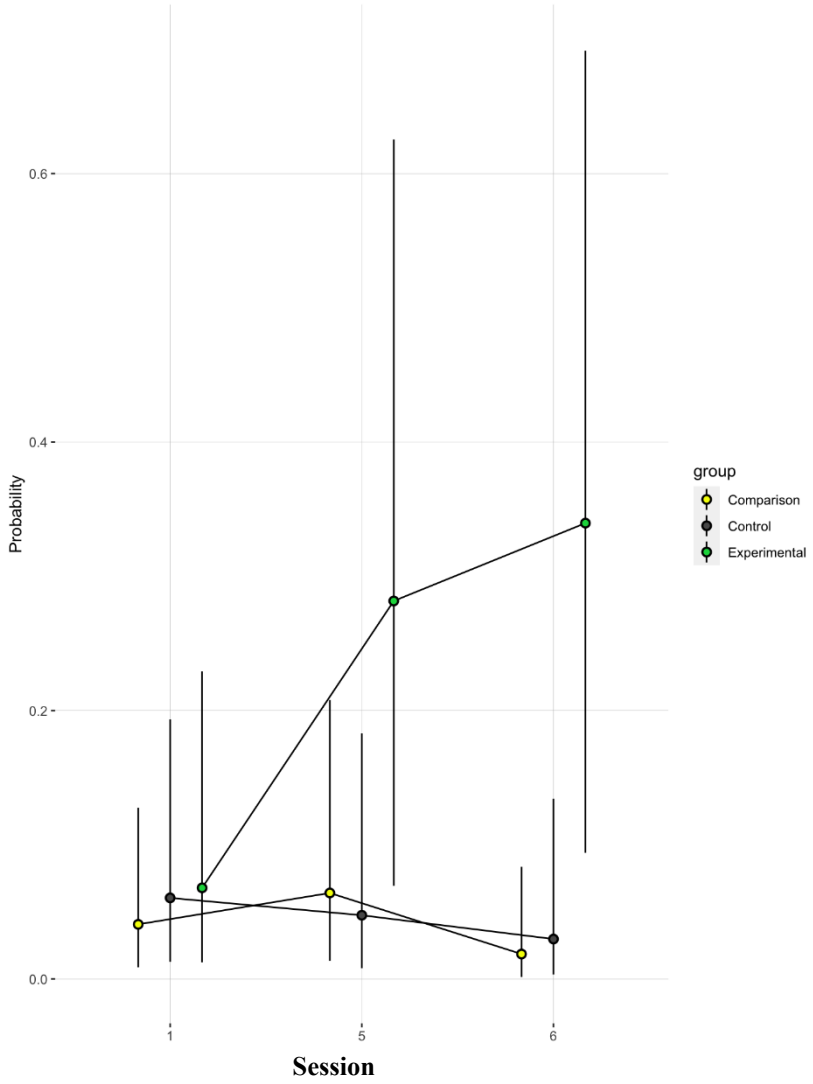


Figure 4.3 Probability of an approximant realization by group and session. The whiskers show the 95% credible intervals for each approximation.

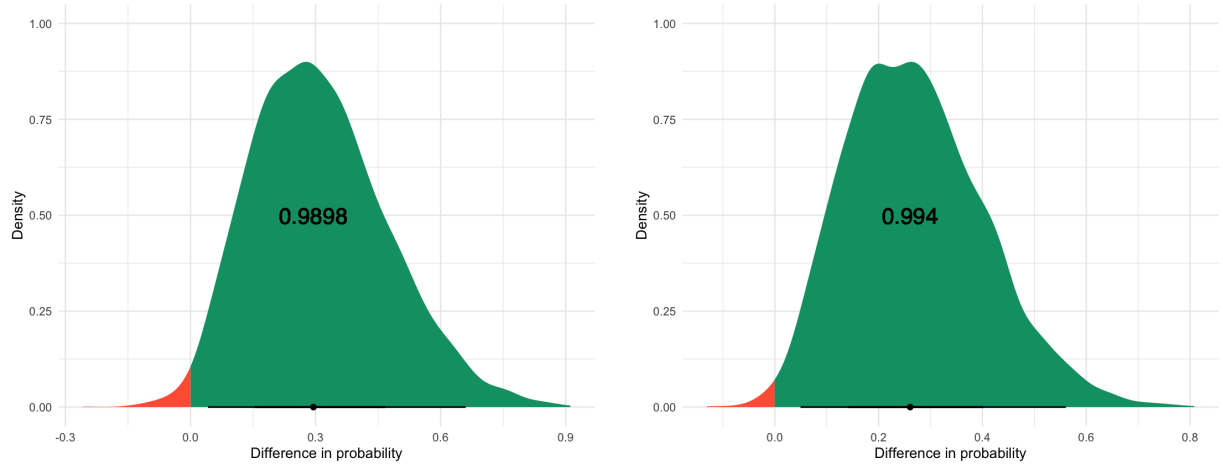
One of the main strengths of using a Bayesian analysis, as opposed to the traditional frequentist approach, when doing either a between-group or within-group comparison is that it allows us to interpret this comparison within a richer framework of probability. The frequentist analysis is relatively binary, where tests for a significant *p* value are carried out to test against the null hypothesis and provide evidence (or not) for the null hypothesis. A Bayesian approach, on the other hand, allows for a more continuous interpretation of the results. While both approaches

seek to give us a way of interpreting whether two data sets are indeed describing two different groups or behaviors, the Bayesian analysis can give us a better conceptualization of how similar or dissimilar the statistical model estimates that the two data sets are. As a result, it better quantifies uncertainty surrounding our conclusions.

For example, the Bayesian logistic regression analysis of the comparison group at Session 1 estimates that the group will have an average performance of producing an approximant 4.1% of the time. The model doesn't know the exact number, this is just its best estimate. By reading the 95% credible intervals at this session, which go from 0.9% to 12.8 %, we can interpret that out of 100 iterations of the model's estimate of the comparison group's ability to produce an approximant at Session 1, 95 of those estimates will be between 0.9% and 12.8%. Although the distribution of the estimates may not be a perfect Bell curve, you can picture a histogram, centered on 4.1% as the mean estimate with the highest likelihood, but with two tails that extend down to 0.9% on the lower end and up to 12.8% on the upper end, thus capturing 95% of all possible outcomes according to the model. Thus, the estimates at the extremes of the credible interval are far less probable than the mean.

When comparing two different groups or sessions, if two credible intervals do not overlap at all, that means that the model predicts with over 95% certainty that those two groups or sessions are different. In a frequentist view, this would be similar to rejecting the null hypothesis because $p < 0.05$. For example, in Figure 4.3, if you compare the Session 6 estimates for the experimental group and the comparison group you can see that they do not overlap at all, so the model is extremely confident that these two groups are different at Session 6. And although there is some overlap in the tails of the credible intervals of the experimental and control groups, the likelihood of the control group outperforming the experimental group at Session 6 according

to the model is only 1.02% (See Figure 4.4A). Similarly, although there is overlap in the credible intervals of the experimental group at Sessions 1 and 6, the model predicts that there is a 99.4% chance that the Experimental group performs higher at Session 6 than at Session 1 (See Figure 4.4B). This explanation helps to interpret what an overlap of two tails means.



A. Session 6 - experimental group over control group

B. Experimental group - Session 6 over Session 1

Figure 4.4 The regression model's measured difference in probability of producing an approximant, based on 4000 iterations, comparing (A) experimental group and control group at Session 6, and comparing (B) the experimental group at Session 6 and Session 1.

4.2 Individual results

4.2.1 Descriptive analysis

The results were also analyzed at the individual level, giving a window into possible individual pronunciation patterns and responses to the treatment. In this chapter, only the results themselves will be presented, but in chapter 5 the results of various individuals will be further analyzed and discussed.

For the comparison group (see scatter plot in Figure 4.5), some interesting observations are first that almost all of the Session 1 stop pronunciations come from Participant 3. But in

Sessions 5 and 6, stop productions are distributed more evenly among many participants, with 5 participants at each session producing more than one stop. Concerning the approximant productions, it is interesting that during Session 2, Participant 3 did not pronounce any at all; this is the only participant in the group who did not succeed in pronouncing approximants during the pronunciation lesson. Most participants (7 out of 10) produced at least 3 of the 5 targets as approximants in Session 2. In Session 5, Participant 41 stands out with a high rate of approximant pronunciation (7 out of 11). This participant was also the only one in the group who had more than one approximant at Session 1 (3 out of 7), and also at Session 6 (3 out of 7). The only other participant to see a sizable increase in approximant production from Session 1 (0 out of 8) to Session 5 (3 out of 10) is Participant 7, though this participant did not produce any at Session 6 (0 out of 8).

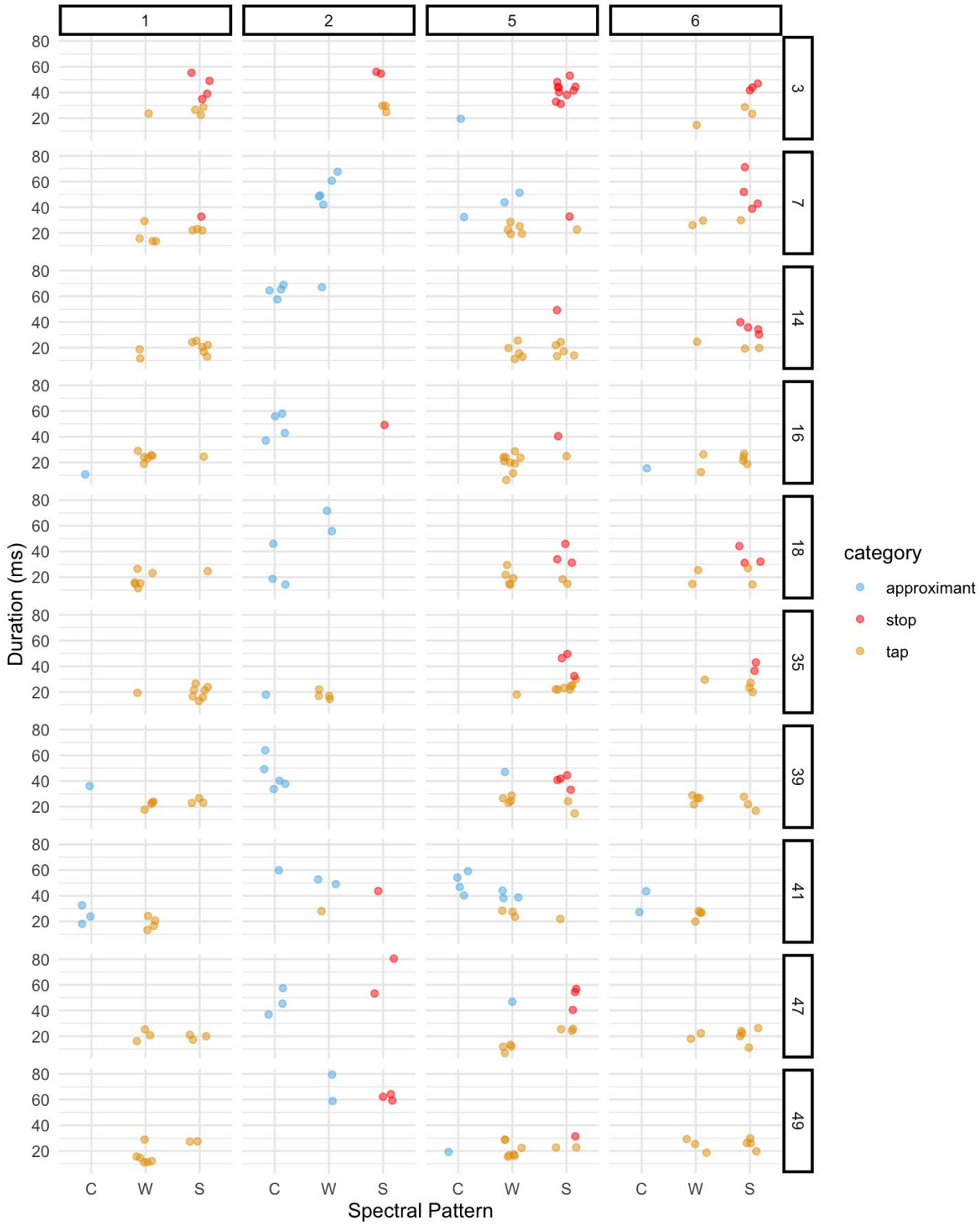


Figure 4.5 Comparison group: individual scatter-plots

Overall, some general trends are seen in this group:

1. Stop production increased substantially across many participants in the comparison group from Sessions 1 to 6;
2. Most participants in the comparison group produced high rates of approximants during the pronunciation lesson at Session 2; and
3. Very few participants in the comparison group saw a meaningful increase in approximant production from Session 1 to Session 5, and none saw a lasting approximant production increase in the paragraph-reading task at Session 6.

For the control group (see scatter plot in Figure 4.6), there were no noticeable changes in stop or approximant production between Sessions 1, 5 and 6. Tap productions dominate every participant's pronunciation at every session.

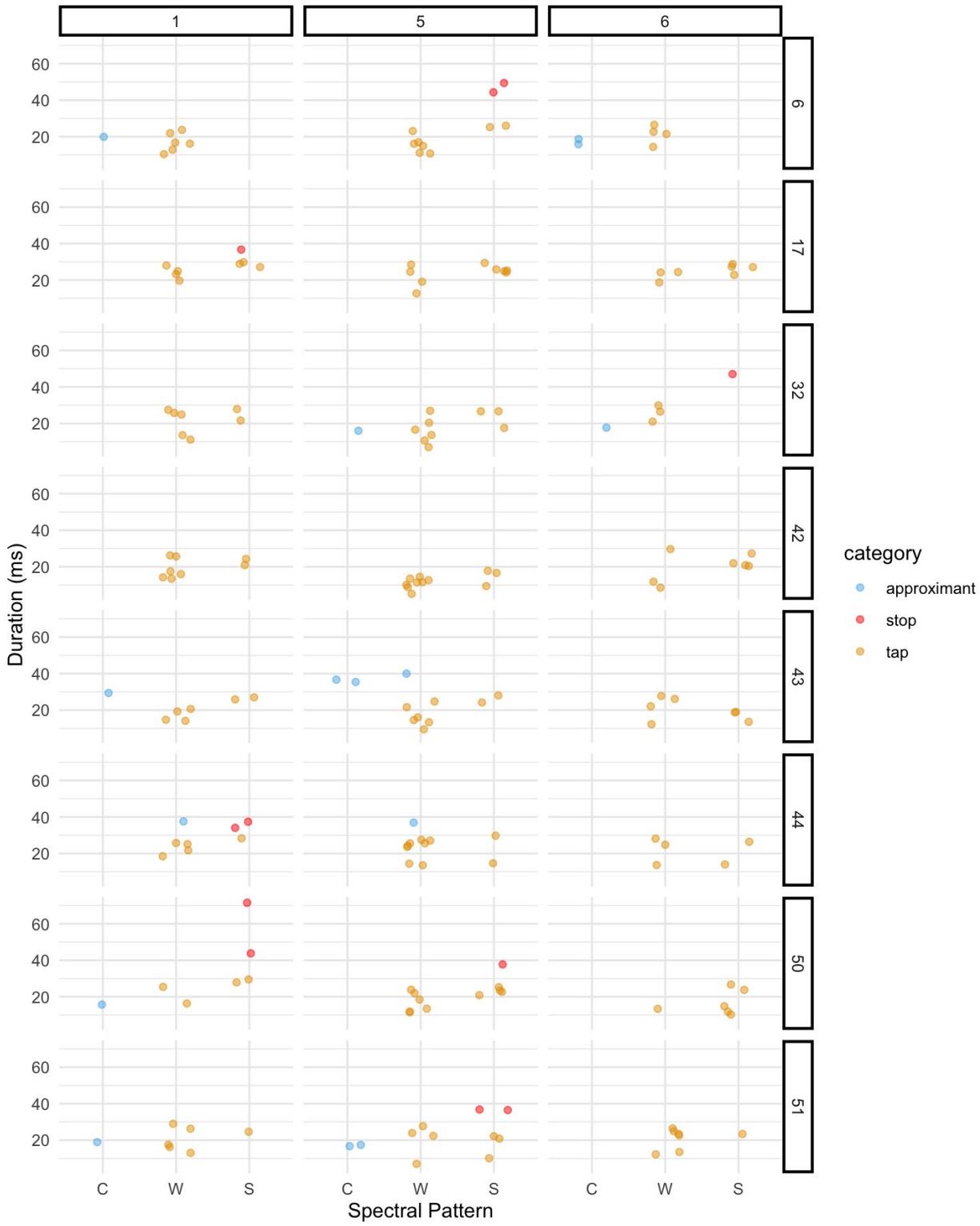


Figure 4.6 Control group: individual scatter-plots

Thus, the general trend in the control group was:

4. Tap productions are predominant throughout, with no substantial change for any participants in the control group from Session 1 to Sessions 5 or 6.

For the experimental group (see scatter plot in Figure 4.7), Participant 12 stands out at Session 2, in which all of the productions during the pronunciation lesson were stops. This participant is joined by two others who produced less than 3 approximants out of the 5 targets at Session 2. Participant 11 had a technical error at Session 2 and their data was lost. At Session 5, 4 of the 7 participants in this group produced multiple approximants, with two participants producing 8 approximants at this session. Participant 12 again stands out at Session 5, producing mostly stops in the paragraph-reading task. Two other participants (Participants 8 and 13) increased in stop production from Session 1 to Session 5, though they only produced two stops each. The change from Session 5 to Session 6 is of particular interest for the experimental group, as the diacritics they had seen at Session 5 were no longer present at Session 6. Nonetheless, at Session 6, 6 out of the 7 participants in this group produced multiple approximants. One interesting individual change was with Participant 31, who produced 8 out of 9 targets as approximants at Session 5, but decreased to 4 out of 8 at Session 6. Participant 12 also showed an interesting change at Session 6, where stop production (1 out of 7) diminished significantly from Session 5 (7 out of 11).

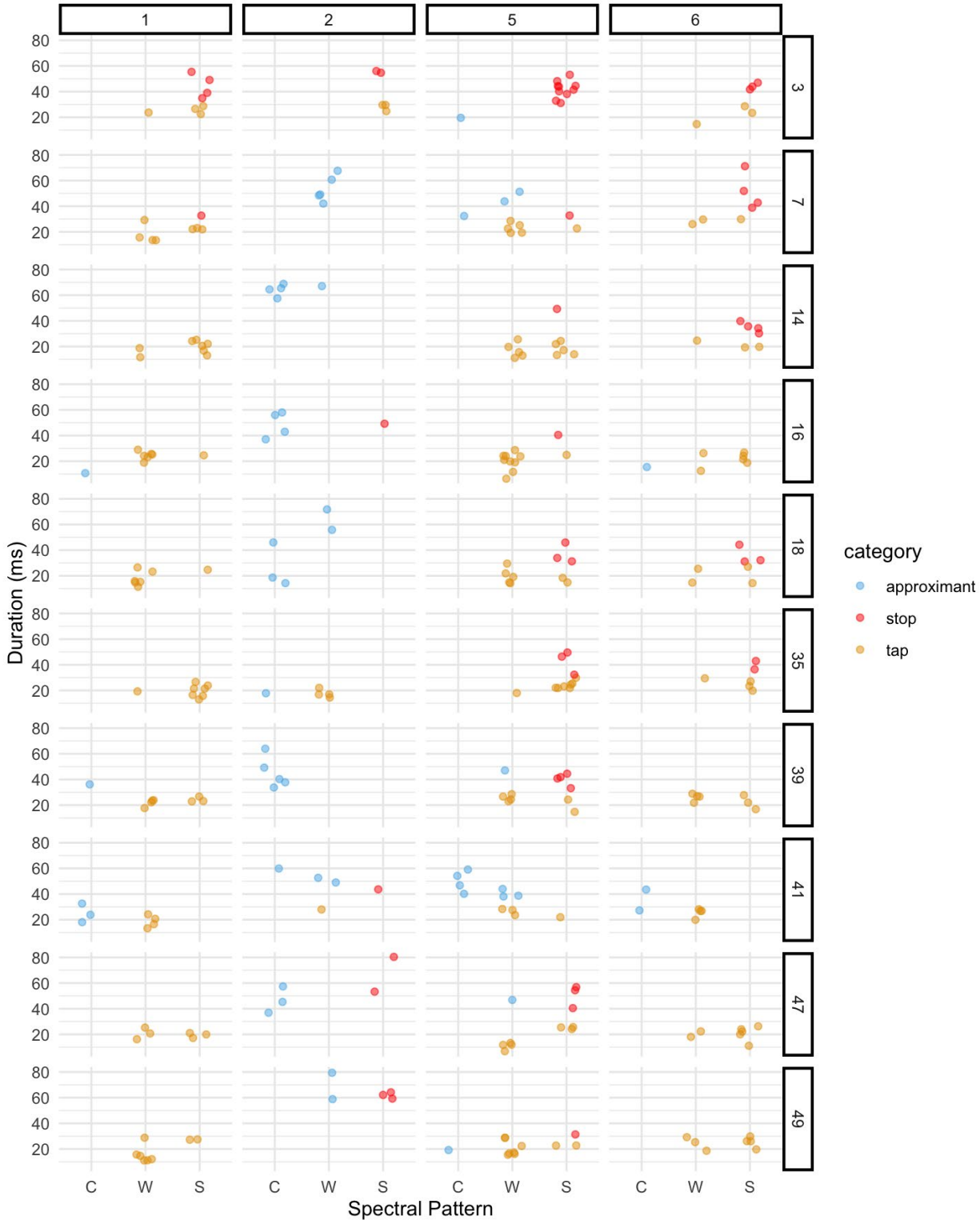


Figure 4.7 Experimental group: individual scatter-plots

Overall, the following general trends were found among participants in the experimental group:

5. Stop production among participants in the experimental group did see moderate increase at Session 5, but this diminished significantly at Session 6;
6. Participants in the experimental group saw substantial approximant production at Session 2, though these were not quite as robust and widespread as the increases by the comparison group.
7. Most participants in the experimental group saw notable increases in approximant production, both at Sessions 5 and 6.
8. Only a few participants in the experimental group showed noticeable changes between Sessions 5 and 6.

In this section, the descriptive analysis of individual pronunciation was presented in order to highlight patterns of learner behavior that a group analysis may not capture. Some key takeaways from this analysis are that for the control group not a single participant saw noticeable changes in pronunciation across the sessions. Individual analysis of Session 2 for both the comparison and experimental groups revealed that while approximant production increased, a few students in each group did not produce any approximants - even directly after a pronunciation lesson. In Session 5, while the group analysis showed moderate increase in approximant realization for the comparison group, the individual analysis showed that it was mainly just 2 of the 10 participants who fueled that increase, and in Session 6, those individuals decreased in approximant production significantly. Many individuals saw noticeable increases in stop production at both Sessions 5 and 6. For the experimental group, there was widespread

incremental improvement in approximant production at Sessions 5 and 6. In the following section, individual results will be further considered, using a Bayesian analysis.

4.2.2 Bayesian analysis

Individual data was also analyzed using the same Bayesian logistic regression approach as was done at the group level. The model's predicted log-odds of individual participants' realizing the <d> as an approximant at each session, converted to probabilities, can be found in Figures 4.8-4.10. The points on the graph show the estimated probability at that session, with error bars showing the 95% credible interval. As can be seen, the model predicts that participants in all three groups had a relatively low probability of producing an approximant at Session 1, with 23 of the 25 participants having less than 20% likelihood of pronouncing an approximant, and the remaining 2 between 20-40% likelihood.

For the 10 participants in the comparison group, at Session 1, 9 of them had less than a 20% likelihood of producing an approximant, and 1 (Participant 41) had a 35.1% of doing so. Advancing to Session 5 we see that two participants (7 and 41) made greater than 10% increases from Session 1 in likelihood of producing an approximant (increasing by 14% and 28%, respectively). All other participants in the comparison group saw negligible (between 0 and 5%) increases in likelihood of producing an approximant. Interestingly, in Session 6, every participant in the comparison group declined in likelihood of pronouncing an approximant, with all participants (including the two who saw increases at Session 5) ending at Session 6 with a slightly lower likelihood of producing an approximant as they had in Session 1.

For the 8 participants in the control group, all of them had less than a 20% likelihood of producing an approximant at Session 1, and there was very little improvement between sessions. Participant 43 is the only one who saw any increase greater than 5% (increasing by 5.6% at

Session 5). Otherwise, the model describes a systematic gradual decline in their probability of producing approximants, both from Session 1 to Session 5, and also from Session 5 to Session 6. Again, all of the participants had a lower likelihood of pronouncing an approximant at Session 6 than they had at Session 1, all remaining well below 20%.

For the 7 participants in the experimental group, 6 of them had less than a 20% likelihood of producing an approximant, and 1 (Participant 11) had a 28.8% of doing so. Individuals in this group saw systematic and often sizable increases in likelihood of producing an approximant at Session 5. Of the 7 participants, 6 of them increased by at least 5%. Participants 1 and 8 increased by 23.2% and 24.0%, respectively, and Participants 11 and 31 increased by 44.7% and 59.4%, respectively. Interestingly, 4 of the 7 participants increased even further from Session 5 to Session 6 by between 5.8% and 10.1%. Participant 31 stands out in that their probability of producing an approximant increased dramatically at Session 5 (+59.4%), but then came back down in Session 6 (-18.1%), though still finishing with a considerably higher likelihood than they had been at Session 1. This was the most dramatic change for any participant between Sessions 5 and 6, and may point to the effect of removing the diacritic between those sessions. By Session 6, all but one of the participants in the experimental group had a greater than 20% likelihood of producing an approximant, and 2 (Participants 11 and 31) ended with a greater than 50% chance of doing so.

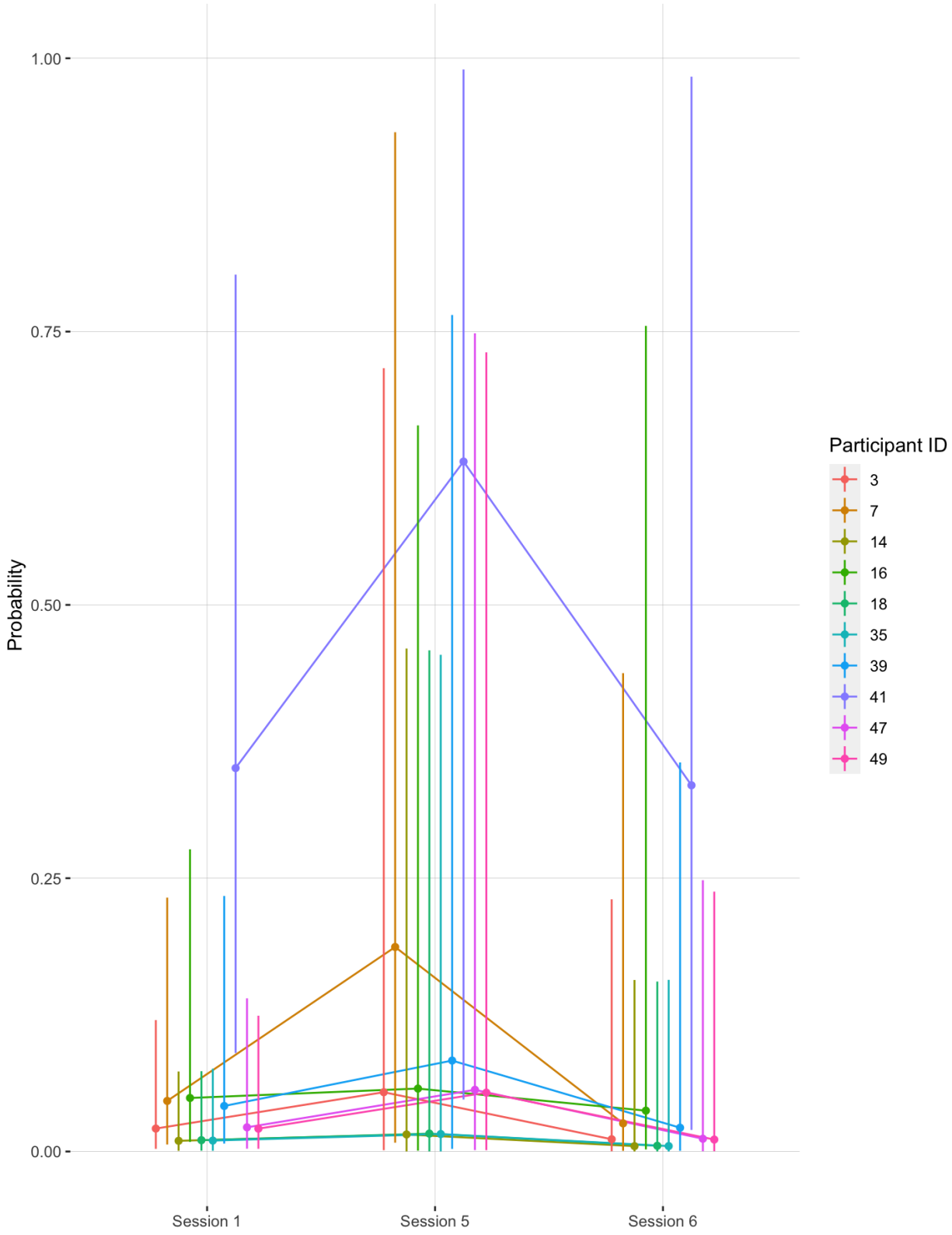


Figure 4.8 Comparison group: participants' probability of producing an approximant at each session.

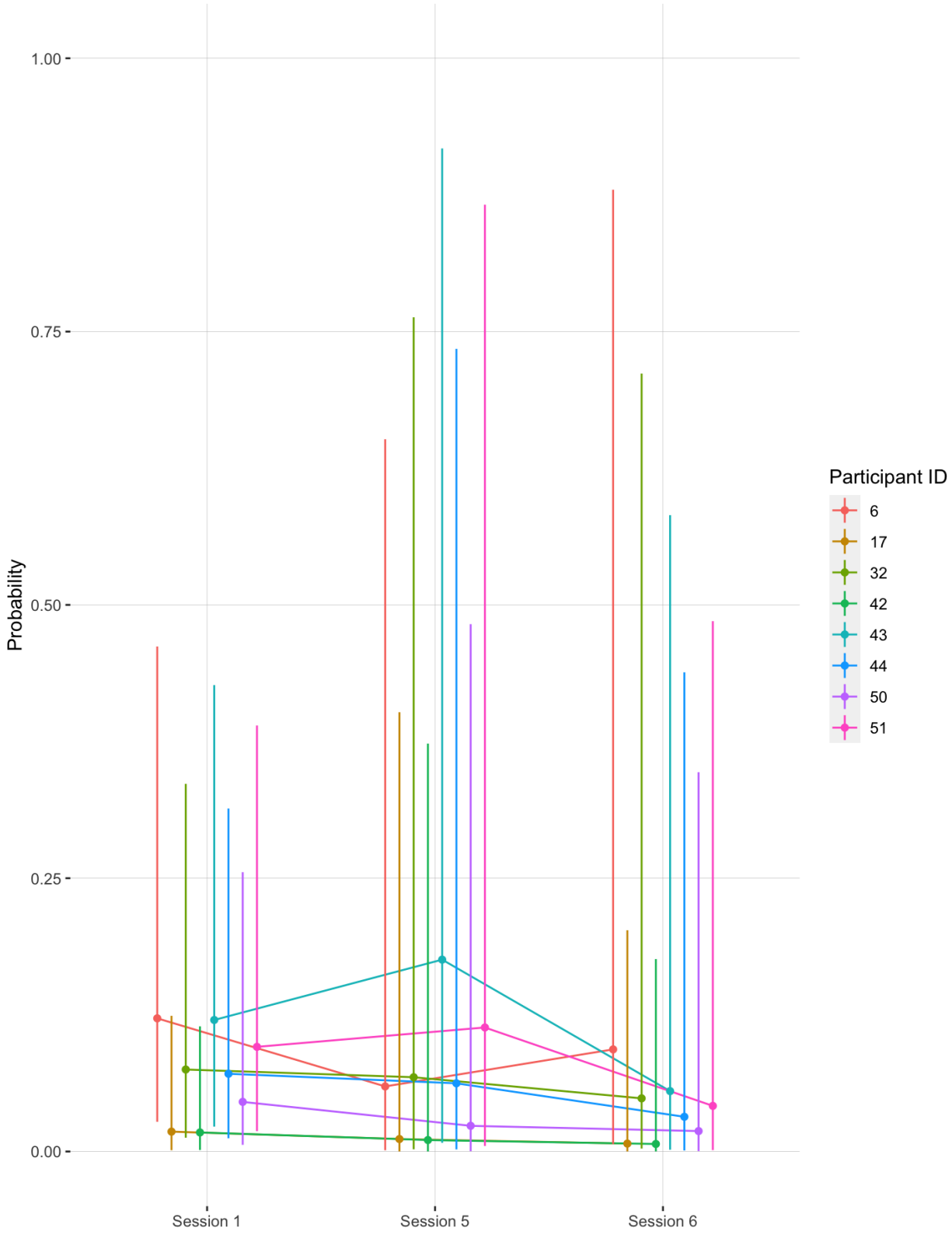


Figure 4.9 Control group: participants' probability of producing an approximant at each session.

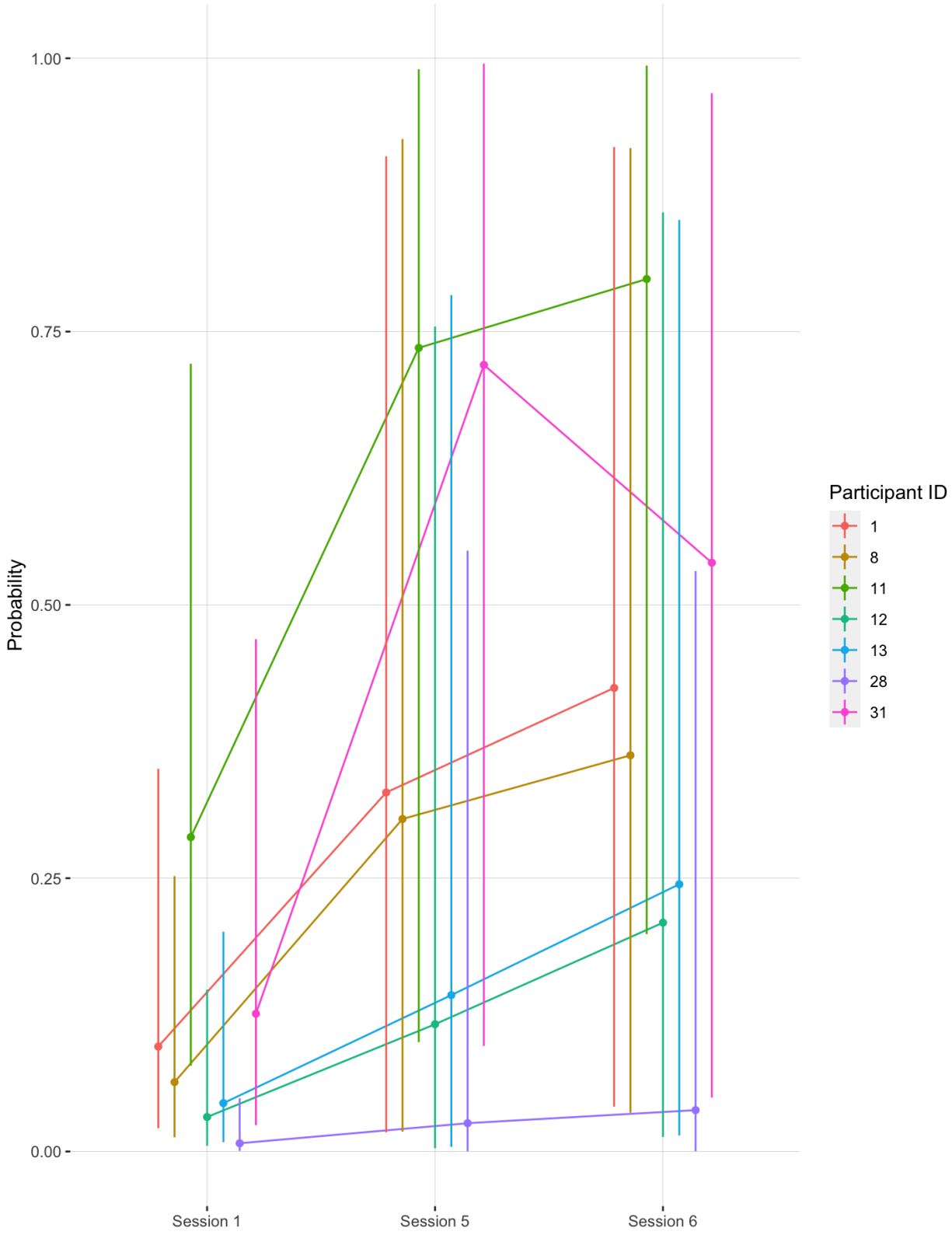


Figure 4.10 Experimental group: participants' probability of producing an approximant at each session.

Overall, the Bayesian analysis at the individual level shows a universal downward trend in likelihood for approximant production in both the comparison and control group going from Session 1 to Session 6. There were 2 participants in the comparison group that did see a noticeable increase in such likelihood from Session 1 to Session 5, indicating a positive response to the pronunciation lesson in Session 2, though the 8 other participants in the group saw negligible change from Session 1. Individuals in the experimental group, however, saw a systematic increase in likelihood of approximant production, with 6 out of the 7 participants increasing in such likelihood by at least 17%.

4.3 Summary

In this chapter, the results from the experiment have been presented both at the group level and at the individual level. At both of these levels of analysis, the experimental group stands out in terms of increasing approximant production across sessions. Both the comparison group and the control group, however, seem to manifest a decrease in such production. This is reflected in the descriptive data as well as the probabilistic Bayesian logistic regression analysis. The transition from Session 5 to Session 6 saw a widespread decrease in approximant production for the comparison and control group, but a mix of increase and decrease for the experimental group. In the following chapter, these results will be discussed in relation to the research questions and hypotheses, with considerations in light of relevant theoretical models and previous studies on pronunciation instruction and the use of diacritic-enhanced orthography.

CHAPTER 5: DISCUSSION AND CONCLUSION

5.0 Overview

In Chapter 4, I presented the results showing how explicit pronunciation instruction and diacritic-enhanced orthography affected pronunciation of L2 Spanish intervocalic <d> during a paragraph-reading task. Additionally, results were presented showing how the removal of diacritics might affect those who previously had been reading with them. Results were analyzed both at the group and individual level.

In this chapter I will discuss those results in detail, considering both how this research interacts with the connectionist theories upon which it was based, and also how it might inform pedagogical application. Finally, the limitations and some future directions of research in this area will be presented. In section 5.1 the pronunciation lesson itself will be considered. In section 5.2 the effects of explicit instruction alone are discussed, and in 5.3 the effect of including diacritic-enhanced orthography is considered. In 5.4 some theoretical implications are presented, and in 5.5 the pedagogical implications are discussed at length. Section 5.6 discusses the limitations in this study, and section 5.7 presents a number of directions for future research in this area. Finally, section 5.8 provides a brief conclusion.

5.1 The pronunciation lesson

The first research hypothesis was concerning the pronunciation lesson itself, and whether the inclusion of the diacritic in the lesson would help participants produce target-like [ð] during the pronunciation lesson more frequently. As a reminder, the pronunciation lesson on intervocalic <d> was very brief, with only about 4 minutes of instruction and practice, followed by the short word-list-reading task. The lesson was unidirectional, with students watching and

interacting with a pre-recorded video on YouTube. Students recorded themselves during this lesson using whatever voice recorder app was convenient on either their computer or cell phone. No in-lesson feedback was able to be provided to students; they could only self-correct according to examples provided in the lesson.

Results showed that all participants who received explicit pronunciation instruction on intervocalic <d>, with or without the diacritic, produced relatively high rates of approximants on a word-list task directly following the pronunciation lesson. The comparison group, who only interacted with authentic orthography, produced approximants 66% of the time at Session 2, and the experimental group, who interacted with diacritic-enhanced orthography, produced an approximate 57% of the time. Compare that to the Session 1 paragraph-reading task, in which the comparison group only produced the approximant 8% of the time, and the experimental group 10% of the time. Thomson and Derwing (2015) pointed out that reading from a word list is quite different from other types of tasks, allowing participants to focus much more on pronunciation, as vocabulary retrieval, grammar, and other linguistic factors which are necessary in natural speech are not needed when simply reading an isolated word from a list. Thus, the improved rate of [ð] pronunciation is not entirely surprising. The purpose of collecting data at Session 2 was to see if participants were able to immediately apply the explicit instruction concerning grapheme-phoneme mapping and overcome the L1 [ɾ] pronunciation. Indeed, the comparison group decreased tap production from 86% at Session 1 to 16% at Session 2, and the experimental group decreased tap production from 90% at Session 1 to 13% at Session 2. Thus, participants generally showed an ability to avoid the typical L1 pronunciation of <d> at Session 2.

The prediction that the diacritic-enhanced text would help learners pronounce target [ð] more consistently, however, was not borne out. The comparison group and the experimental group showed very similar rates of pronunciation at Session 2, indicating that the presence of the diacritic in this lesson did not have a strong effect on pronunciation.

Turning to the DRC model (Coltheart et al., 2001) presented in Chapter 2, the above results indicate that these learners had, to some degree, associated [ð] with Spanish <d> in their L2 Grapheme-Phoneme Rule System, such that they were able to pronounce the approximant in a very focused word-list-reading task. While the experimental group, whose word list made use of diacritic-enhanced <d̰>, did not exceed the comparison group in terms of [ð] pronunciation, the fact that their error rates were similar to the comparison group (who viewed authentic <d>) suggests that they are to some degree still processing the <d> grapheme. This could be contrasted with Jackson's (2016) use of a completely artificial grapheme <ƱI> for an Arabic sound, where any grapheme-phoneme correspondence would be tied to the artificial grapheme. Because L2 learners will eventually be confronted with authentic L2 orthography, it is likely a positive development that learners interacting with the diacritic-enhanced <d̰> in the present study show signs of processing the authentic <d> grapheme.

Note that the data from Session 2 are not meant to tell us whether or not participants improved in their pronunciation. Rather, Session 2 data simply give us a window into whether or not participants internalized and could apply the pronunciation lesson they had just received. In order to see the effects of the pronunciation lesson on these learners' speech, we must compare the results of their paragraph-reading pronunciation from Session 1, before the lesson, to Session 5, after the lesson.

5.2 Explicit instruction alone

The second research hypothesis was that because explicit instruction has been shown to generally help learners' pronunciation, participants would produce more target-like [ð] on paragraph-reading tasks after the lesson. The descriptive analysis of group data in Chapter 4 showed that the comparison group, who had received a pronunciation lesson but only interacted with authentic L2 orthography, showed signs of very modest improvement in approximant production from Session 1 (8%) to Session 5 (13%). The control group, however, produced approximants at the same rate at both sessions (8%), with the individual-level data also attesting to there being no substantial change for any individual in the control group.

The Bayesian analysis of group data was very static for both the comparison group and the control group. The probability of producing an approximant increased from Session 1 to Session 5 just slightly for the comparison group (4% to 6%) while slightly decreasing for the control group (6% to 5%). For each of these groups, the considerable overlap of the 95% credibility intervals indicate that the model doesn't see much difference between Session 1 and Session 5 for either of these groups.

The results of the control group give us an indication of what happens over the course of four weeks (the approximate amount of time from Session 1 to Session 5) with this group of L2 Spanish learners who do not receive any pronunciation instruction. Their pronunciation of intervocalic <d> in the reading task as target [ð] did not improve, and might have declined. This trend with the control group continued at Session 6, with descriptively measured approximant pronunciation declining to 5% and the regression model estimating that their probability of approximant pronunciation had dropped to 3%. Looking at the individual level data, the participants in the control group were rather homogenous, with the only notable, though

temporary, improvement being one participant showing about a 5% increase in probability of pronouncing an approximant at Session 5, only to fall back again at Session 6. Every single participant in the control group saw a decreased probability of pronouncing an approximant at Session 6 compared to Session 1.

This lack of improvement, and possible decline, in accurate L2 pronunciation of Spanish /d/ over a 5-week period is also what was found in the control group of Bajuniemi (2013), which likewise produced fewer target-like pronunciations after a 5-week interval. Such a decline over time is also reminiscent of what Zampini (1994) found in her study of L2 Spanish pronunciation of <v>, where students in the 4th semester produced non-target [v] at an even higher rate than the 2nd semester students. The control group in Elliott (1997) also saw no improvement in L2 Spanish pronunciation for any of the test items in that study over the course of one full semester.

Collectively, the results of the present study alongside previous similar studies suggest that students in a Spanish course which does not give explicit instruction concerning problematic graphemes will not improve in their pronunciation of those sounds, at least not in a measurable way within a semester, and perhaps not even after multiple semesters of instruction. In fact, perhaps due to continued contact with L2 orthography, which reinforces L1 GPCs, students' pronunciation of these items may be pulled further towards an L1 pronunciation. This is exactly what Young-Scholten (1995) described in her depiction of the problematic input that learners typically receive in the L2 classroom, where she states that "premature exposure to orthographic input (at or near the initial stages of L2 phonological development) can be expected to impede progression to native phonological competence in the L2" (Young-Scholten, 1995, p. 112). In the present study, there was indeed a notable impedance for learners to progress toward native

phonological competence for the target /d/, as revealed in the results for both the control and comparison group, and it seems that orthographic effects were a crucial factor.

Returning to the results of the present study, the comparison group did show signs of moderate improvement in target pronunciation of intervocalic <d> from Session 1 to Session 5. This may indicate that a few learners were applying the pronunciation lesson when reading the authentic orthography with intervocalic <d> at Session 5. Looking at the individual results, it seems that the moderate increase for the comparison group at Session 5 was driven by just 2 of the 10 participants in this group (Participants 7 and 41), which is probably why the statistical model did not predict much improvement had gone on for the group overall. (Reviewing the Bilingual Language Profile results did not reveal anything exceptional about these 2 participants.) But at Session 6, the comparison group dropped all the way down to a [ð] pronunciation rate of 5%, which was even lower than their Session 1 rate. Similarly, the regression model predicts that their probability of pronouncing an approximant had dropped to 2% at Session 6.

The initial increase in [ð] pronunciation by the comparison group may demonstrate some positive effect for the pronunciation lesson alone for at least a couple of learners. The relative recency of the lesson may be a factor for why these students showed some improvement at Session 5. But the drop at Session 6 may demonstrate that the memory of the lesson is starting to wane. Viewed from within a connectionist framework, this could mean that whatever improvement was present at Session 5 was relying mostly on the declarative knowledge learned in Session 2 and a heightened awareness due to the pronunciation lesson. But in the absence of considerable statistical frequency of application, the pronunciation of intervocalic <d> as [ð] was not automatized. Thus, as attentional resources waned over time, and without a sufficiently

automatized target L2 pronunciation built up, the previously automatized L2 pronunciation of intervocalic <d> as [ɾ], which was the prevailing pronunciation at Session 1, resurfaced and became the prevailing pronunciation again at Session 6.

But a separate and interesting distinction between the comparison group and the control group is in relation to their production of <d> as [d]. It should be recalled that in native English speakers, Herd et al. (2013) reported that based on various studies, flapping of English /d/ occurs between 94% and 99% of the time in post-tonic, intervocalic position. Indeed, even in their L2 Spanish, at Session 1 both the comparison and the control group produced taps 86% and 83% of the time, respectively, with relatively few stops. But at Session 5, the comparison group saw a significant increase in stop production, going from 7% at Session 1 all the way to 25% at Session 5. The control group, however, produced the stop at a rate of 8% at Session 1 and 6% at Session 5. This robust increase in stop production for the comparison group is striking, and persisted into Session 6 (25%), whereas the control group's stop production remained quite small (4%).

There are at least a few possible explanations for this change in pronunciation in the comparison group. One possibility is that these participants may simply be more conscious of <d> in the reading due to the pronunciation lesson. The heightened consciousness could cause them to speak more emphatically, thus avoiding the tap pronunciation. But even if learners are more conscious of <d> due to the lesson, the processing demand of applying the newly learned [ð] pronunciation to <d> is too high, leaving them with a stop production. The fact that participants must process lexical and grammatical information during a paragraph-reading task may be a reason for why the very high [ð] pronunciation rate which was attained by the comparison group in Session 2 (66%) was not applied in Sessions 5 and 6.

Another possibility, somewhat related to the previous explanation, is that the participants in the comparison group may be missapplying what was taught in the pronunciation lesson. In the lesson, participants were taught that there are two Spanish pronunciations of <d>, either a stop after a pause, nasal, or lateral, or an approximant in all other contexts. Although a few examples were given, it's likely that the students did not form strong conceptualizations of the two contexts. Indeed, if given a list of Spanish words containing <d>, they may not even be able to identify whether a given <d> should be pronounced one way or another. This being the case, it could be that some in the comparison group are trying to apply the pronunciation lesson by producing <d> as [d], not realizing that these <d>s are intervocalic.

To summarize, this study found no lasting improvement in approximant production based solely on the brief explicit pronunciation instruction provided to the comparison group in Session 2. Just the same, the very moderate, temporary improvement seen by some participants in Session 5, as well as the increase in stop production, may indicate that some of these learners did have some memory of the lesson and were to some degree processing the GPC rules acquired in the lesson, but were unable to apply them in the paragraph-reading task. The results found here are similar to those reported by Elliot (1997) and Herd et al. (2013), in which students failed to show improvement for the spirantization of /d/, despite focused training/instruction on the target item.

But the results of this study do not line up with those of Bajuniemi (2013), in which students who received pronunciation instruction improved their percent of “target-like” pronunciations of intervocalic /d/ by 24.1%. A crucial difference between the present study and Bajuniemi (2013) is the amount of class time dedicated to the item, as the present study only used a one-time, 4-minute training video watched by students outside of class, whereas

Bajuniemi (2013) provided students a 10-minute lesson on pronunciation of approximant /d/ twice a week for 5 weeks (totalling about 100 minutes of pronunciation instruction). The learning activities in Bajuniemi (2013) also included repetition exercises as a class and in pairs, whereas in the present study the learner merely repeated the sounds from the video, with no interpersonal activities. It is very likely that the difference in nature of these approaches to pronunciation instruction led to very different results.

One framework to pronunciation instruction that has been suggested is that of Mellado (2012), who proposes a 6-step approach: 1. Sensitization; 2. Explanation; 3. Re-educating the Ear; 4. Imitation; 5. Controlled Practice; 6. Free Practice. In her proposal, before learners receive explanation concerning the target sound in step 2, they are given auditory activities in step 1 to help sensitize them to the important contrast that is going to be taught. Indeed, auditory training figures heavily into her approach, as in step 3 they are given further training to re-educate the ear. In Flege's (1995) seminal Speech Learning Model, he proposed that the ability to produce an L2 sound was limited by the learner's ability to first perceive that sound. In his and Bohn's revised model (Flege and Bohn, 2022), this particular proposal was modified and the notion of perception's precedence was removed, although they maintain that "a strong bidirectional connection exists between production and perception" (Flege and Bohn, 2022, p. 29). The 4-minute lesson in the present study did indeed provide examples of the two allophonic realizations of /d/, but there was no particular training to ensure development in the learner's ability to perceive those sounds distinctively. It is possible that due to a lack of discernment between these realizations, learners were unable to properly execute their pronunciations.

5.3 Explicit instruction with diacritic enhancement

5.3.1 On paragraph reading

The third research hypothesis predicted that participants who received the pronunciation lesson with diacritics and then carry out paragraph-reading tasks with diacritics would improve [ð] production rates due to the help of the visual cue, reminding participants to apply the pronunciation lesson held in their long-term memory. Results from the descriptive analysis showed that the experimental group saw a much more robust improvement from Session 1 (10%) to Session 5 (34%) compared to the comparison group (8% to 13%). Similarly, the Bayesian analysis showed that the experimental group had increased from a 7% likelihood of producing an approximant at Session 1 to a 28% likelihood at Session 5, far more than the comparison group (4% to 6%). While the large range of the 95% credibility intervals for the experimental group indicates there is uncertainty about the exact prediction (the model estimates that at Session 5, the group has a 95% chance of producing anywhere from 7% and 63% approximants), the model's estimates are densely centered around the 28% likelihood, with performance towards the extremes of the credible interval growing increasingly unlikely.

Looking at the descriptive individual results, Participants 8, 11, and 31 made substantial increases in approximant production from Session 1 to Session 5, increasing their raw rates by 33%, 33%, and 89%, respectively, while most others also saw modest increases in approximant production. Similarly, the Bayesian regression analysis showed that 6 out of the 7 participants saw increases of 5% or more in likelihood of producing an approximant, with 4 participants increasing by over 20%. While the 95% credibility intervals for each individual still indicate a good deal of uncertainty, it seems likely that the use of diacritics had a meaningfully positive effect on [ð] production rates.

A review of the Bilingual Language Profile results for the more outstanding members of this group indicated that Participant 11 did have a somewhat more balanced Dominance score compared to the average of the other members of the group (101 vs. 126), indicating a higher degree of Spanish dominance. Also, Participants 8, 11, and 31 all responded with the highest option (6, on a scale of 0-6) for a question asking if they want others to think they are a native speaker of Spanish. This was not altogether unusual among the 25 participants included in the study: 3 others from the comparison group and 1 from the control group also did so - however none of those 4 saw any notable increase in approximant production at Sessions 5 or 6. The fact that these 3 in the experimental group saw a substantial increase in approximant production at Session 5 may indicate some interaction of attitudes toward L2 pronunciation and the combination of the pronunciation lesson and diacritic-enhanced materials.

There may be various explanations for the apparent effect of diacritics on pronunciation. When reading authentic L2 orthography, the shared <d> grapheme brings with it a strong association with the L1 pronunciation. A brief pronunciation lesson, even directed specifically at that grapheme, cannot easily overcome that L1 association (as evidenced by the comparison group results). By modifying the grapheme with a diacritic, a slight visual transformation of the authentic grapheme, it is possible that learners can more easily associate a new L2 sound with that grapheme.

But why would such a small visual modification make such a difference? It may be helpful to return to a simple quote from chapter 2 of this dissertation: “orthographic input is the input that is available to learners in written form” (Hayes-Harb & Barrios, 2021). By adding a diacritic, and then explicitly teaching a pronunciation lesson giving meaning to that diacritic, we make a little more linguistic input available to learners every time they simply read.

Solier et al. (2019), pointed out that the visual orthographic trace of the word provides "a permanent external memory" which learners can refer to (Solier et al., 2019, p. 1460). The comparison group had received the same pronunciation lesson and showed excellent ability to produce the approximant while reading from a word-list during that lesson, even though they were reading unenhanced <d>. This shows us that they had generally received the correct phonetic knowledge into their memory. In this regard, at Session 2 the experimental and comparison groups seemed to be exactly the same. But as the weeks went by and the comparison group went back to a paragraph-reading task, they performed not much better than before having received the lesson. It may be that whatever association had been learned and applied during the lesson did not have anything permanent or concrete to link it to, and these learners were relying on a fading, internal memory. Furthermore, the new association of <d> = [ð] was competing with an association of <d> = /d/ already cemented with its own permanent external memory. For the experimental group, the novel diacritic <◌̆> below the <d> may have given the learner something permanent and external (because it is permanently fixed on the page, not depending on anything of the learner) which is dedicated exclusively to an association with pronunciation knowledge that was internalized at the lesson.

5.3.2 On paragraph reading after diacritic removal

The fourth and final research hypothesis predicted that for those participants who were reading with diacritics, once they are removed there would be some diminishing of [ð] production, though not all the way to pretest levels, since due to some statistical frequency, some participants may have started to internalize a more accurate L2 allophonic distribution for /d/. Results from the descriptive analysis showed that the experimental group in fact did not diminish in [ð] production, pronouncing the target allophone a bit more frequently at Session 6 (37%) than

they had in Session 5 (34%). Similarly, the Bayesian analysis showed that the experimental group had increased from a 28% likelihood of producing an approximant at Session 5 to a 34% likelihood at Session 6, though there is a considerable overlap of the credibility intervals of the experimental group at these two sessions, there is still a 99% probability that the experimental group is performing higher than the control group at Session 6.

Looking at the descriptive individual results, most participants pronounced a similar or slightly higher rate of approximants at Session 6 compared to Session 5, with the notable exception of Participant 31, who had pronounced 8/9 (89%) targets as an approximant at Session 5, but declined to 4/8 (50%) at Session 6. Similarly, the Bayesian regression analysis showed that 6 out of the 7 participants saw increases of between 1.2% and 10.1% in likelihood of producing an approximant, with the exception of Participant 31 who declined in likelihood, from 72% at Session 5 to 53.9% at Session 6. The 95% credibility intervals for each individual overlapped significantly between sessions 5 and 6, indicating a great deal of uncertainty concerning any difference between these two sessions, though there are no apparent indications from the model that individuals in the experimental group decreased in their ability to pronounce the approximant at Session 6, except for Participant 31.

The case of Participant 31 may give some indication to the effect of diacritic-enhancement. The sharp increase to near perfect approximant production while reading with diacritics is striking. It may indicate that for this particular student, the combination of the pronunciation lesson with the support of diacritics was particularly effective. It's possible that this student may have had some degree of ability to pronounce the approximant, and the Session 1 task didn't manifest that. It should be noted that students had been instructed to speak with as native-like pronunciation as possible even at Session 1. But it's possible that this student simply

didn't actively attempt to pronounce approximants until after the lesson. Just the same, no other student out of the 17 who received a pronunciation lesson on <d> made such a dramatic increase. The fact that this student dropped quite a bit once the diacritics were removed also points to their possible effectiveness.

5.3.3 Possible lexical and phonotactic effects

One of the interesting divergences in the results is between the comparison group and the experimental group between Session 5 and Session 6. The comparison group saw a noticeable decrease in approximant production, whereas the experimental group saw a slight increase. While a diminishing effect for the pronunciation lesson over time is one explanation for the comparison group's decline, another possible explanation could be related to the test items. Returning to the target word list in section 3.2, a number of targets can be identified as high-frequency, two-syllable (HF2S) functional words, namely *cada* and the various forms of *todo*. Frequency has indeed been found to be a factor influencing lenition in L1 Spanish /d/ (Eddington, 2011). In Sessions 1, 5, and 6, the proportion of these HF2S items was 4 out of 8 (50%), 3 out of 11 (27%), and 3 out of 8 (38%), respectively. Interestingly, at every session, each group saw higher rates of tap pronunciation for <d> in HF2S words than for the other targets. Overall, HF2S words saw their <d> pronounced as a tap 85.1% of the time, whereas the rest of the target words only saw <d> pronounced as a tap 67.4% of the time. These words also saw the lowest rate of stop pronunciations, being pronounced as [d] only 5.5% of the time, compared to 16.7% for the other words. A very likely explanation for this tendency is that these intermediate learners have read (and pronounced) these high-frequency items so many times in their years of L2 experience that the lexical item is recognized without analysis of the component

graphemes, and the entrenched phonological form and tap pronunciation of that form are immediately activated.

Thus, the comparison group's decrease in approximant production from Session 5 to Session 6 may be a function of the increased proportion of HF2S words in Session 6. The experimental group, however, maintained their approximant production between these two sessions. If lexical effects are a reason for this difference in behavior, it may indicate that any benefit that the comparison group received from the pronunciation lesson at Session 2 may not have penetrated the HF2S words, whereas the diacritic-enhanced learning for the experimental group did indeed make inroads into how learners process and pronounce these lexicalized words. To test these hypotheses, a future iteration of this study would need to include a balance of these classes of lexical items, and should consider investigating learners at the initial stages of learning, who would not have built up strong lexical representations of the high-frequency words.

Another item-related phenomena that was uncovered is related to the vowels surrounding the /d/ segment. Looking only at the lower frequency words, when the vowel preceding the /d/ was a low vowel (i.e., /a/) or a high vowel (/i, u/), the /d/ was pronounced as an approximant more frequently (19.3% and 17.5% of the time, respectively), than if it was a mid vowel (9.2%). And if the vowel following the /d/ was an /a/, it was pronounced more frequently as an approximant (20.9% of the time) than if it was an /o/ (14.5%) or an /e/ (9.6%). In a study on L1 Spanish spirantization of /d/, Simonet et al. (2012) also found vocalic context to be a factor in the degree of weakening, with preceding low vowels showing a greater degree of weakening than high vowels. Blair and Lease's (2021) likewise found more weakened /d/ pronunciations after a low vowel in their study of Heritage speakers of Spanish.

With these tendencies in view, we can return to the target words of each session and we find that Session 5 had a higher proportion of targets with a preceding low vowel (3/11), and/or a following /a/ (5/11), compared to Session 6 (0/11 and 0/11, respectively). It may be that the surrounding vowels create articulatory contexts that either favor or disfavor the pronunciation of an approximant. Thus, in addition to the increased proportion of HF2S words at Session 6, the less favorable vocalic context of the target words at Sessions 6 may have also contributed to decreased approximant production for the comparison group. Because the number of test items representing these various vocalic contexts was not balanced at each session, this effect can only be hypothesized, and a future iteration of this type of study should design a procedure and target items that take this possible effect into account.

Overall, when viewed through the task of paragraph reading, this study demonstrated that, by themselves, short, explicit pronunciation lessons for problematic grapheme-phoneme mappings may be ineffective. A lack of improvement in paragraph reading was found despite the fact that learners in the comparison group had shown target-like pronunciation during the lesson itself. Conversely, this study does give evidence in support of a possibly strong effect of explicit pronunciation instruction when combined with diacritic-enhanced orthography. It also gives evidence supporting a maintenance of that improvement even after diacritics are removed, indicating that learners' L2 mappings for the authentic grapheme have been positively impacted. In the next section, some theoretical implications will be considered.

5.4 Theoretical implications

When we turn to the DRC model, the results of the present study bring up some interesting questions about what systems might be involved when these participants read diacritic-enhanced texts. The first is regarding the Grapheme-Phoneme Rule System. In this

study of intermediate-level Spanish students, participants would already have an entry for Spanish <d> = /d/. So, what happens when they are taught to associate a different sound with the diacritically enhanced <d̪>? Does the mapping for <d> = /d/ get activated, or is a new <d̪> mapping created?

In Jackson (2016), students were taught a contrast between <k> = /k/ and <k̟> = /q/, where <k̟> was a new grapheme associated with a new L2 phoneme /q/. Crucially, those phonemes were contrastive, and learners were even tested on items in which they were contrasted. Presumably, in that case a brand new entry for <k̟> = /q/ in the Grapheme-Phoneme Rule System. It's possible that participants in the present study would similarly generate a new L2 Spanish phoneme /ð/, coexisting with L2 Spanish /d/, with a new entry in the Grapheme-Phoneme Rule System: <d̪> = /ð/. The fact that /ð/ is an L1 English phoneme would also facilitate such a development. Additionally, these intermediate learners have already created a mapping for an authentic diacritic-enhanced L2 grapheme: <ñ> = /ɲ/. Thus, it is conceivable that these learners would create an entry for <d̪> = /ð/. Such a possibility is represented as (A) in Figure 5.1.

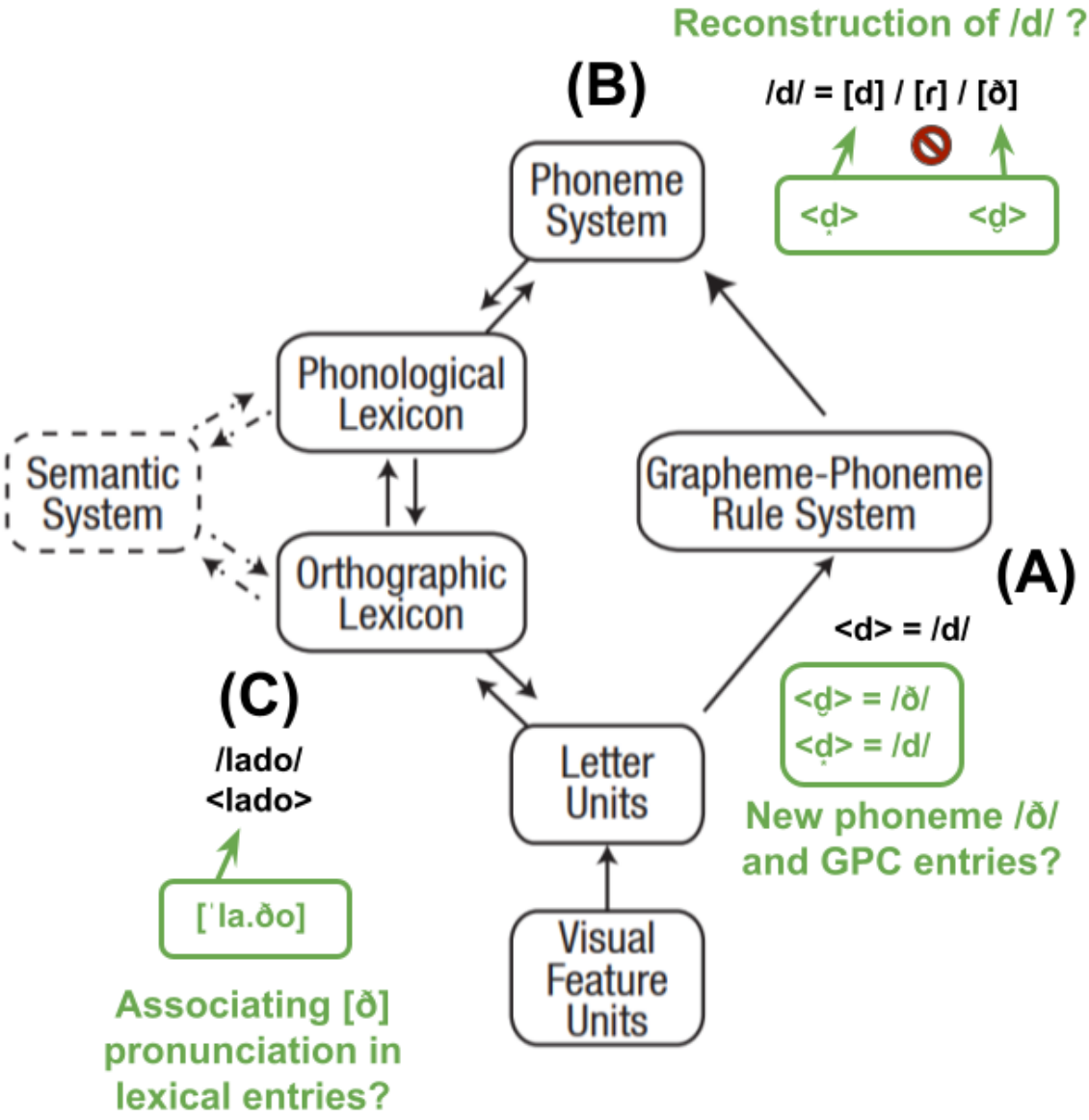


Figure 5.1 Simplified DRC model, with possibilities of new developments due to diacritic-enhanced learning for Spanish <d>: (A) a new /ð/ phoneme and new Grapheme-Phoneme Rule System entries; (B) a reconstruction of the Phonological System, and (C) new phonetic associations for the Phonological and Orthographic Lexicons.

But in the present study, the learners are not truly learning a new L2 phoneme. Instead they are learning that the letter <d> has more than one pronunciation and to pronounce the proper allophone [ð] when they read <d> in an intervocalic context. These sounds were not presented as contrastive segments, but rather the complementary distribution of the [d] and [ð]

realizations was described, and only the [ð] was practiced. Also, the "root" of the diacritic-enhanced <ḏ> is undeniably <d>. Thus, it may be possible that the learners are not creating a new phoneme, but are just learning to become sensitive to two different pronunciations of the one phoneme. The fact that the experimental group maintained the [ð] pronunciation of <d> after the diacritic was taken away strongly indicates that they associated the [ð] pronunciation with the grapheme <d>. This implies that they did not create a new entry in the Grapheme to Phoneme Rule System for <ḏ> = /ð/, otherwise, they would have completely abandoned the [ð] pronunciation once <ḏ> was absent. Because these were intermediate learners, even though they did not pronounce [ð] very much at Session 1, they had surely built up some phonological awareness of [ð] as an acceptable form of /d/, having received some accurate auditory input over the prior years of their L2 experience.

With all of that in mind, it seems very possible that these learners, when learning and then practicing to read <ḏ> as [ð], were actually doing some reconstruction of their L2 Phoneme System, developing and strengthening the allophonic distribution of /d/, a possibility represented as (B) in Figure 5.1. Or perhaps, more specifically, within the Phoneme System for L2 Spanish /d/, there was development and strengthening of the pronunciation of the [ð] allophone as it relates to the articulatory (speech motor) realization. The diacritic-enhanced <ḏ>, with its permanent, external sign tied to a known L1 pronunciation (by means of the explicit instruction), seems to have functioned as a bridge to help these learners to bring the [ð] articulatory gestures to L2 /d/ and <d>, increasing statistical frequency of a more accurate pronunciation of this phoneme and grapheme. Thus, whereas Jackson (2016) gave evidence for diacritic-enhancement helping create a new L2 phoneme (<ḳ> = Arabic /q/), the present study suggests that diacritic-

enhancement may be useful in reconstructing L2 phonemes with a more native-like allophonic pronunciation.

Furthermore, according to the DRC there are two routes by which we retrieve and pronounce words: the Grapheme-Phoneme Rule System (as considered above) and the Phonological Lexicon. This lexicon is where the phonological forms of entire words are stored. As learners increase statistical frequency of the accurate pronunciation of L2 words via the sublexical, GPC-Phonological System route, the resting activation level of the proper pronunciation of that word in the lexical route rises. Because the present study only tested the speech during a paragraph-reading task, which leans mostly on speech issuing from the orthography, there is no evidence for whether or not the learners were internalizing the pronunciation of the entire words. It seems very likely that such associations would be developing, such that after reading and pronouncing a certain word many times (e.g., <lado> = [ˈla.ðo]), that in their natural speech they would also pronounce [ˈla.ðo]. This possible development is represented as (C) in Figure 5.1. Future study that elicits the natural speech apart from a reading task would be needed in order to ascertain whether or not developments are taking place in the Phonological Lexicon of these L2 learners.

5.5 Pedagogical implications

5.5.1 As applied in this study

Perhaps the most poignant implications of the present study are related to L2 teaching. If the results of this study indeed point to the effectiveness of including diacritic-enhancement in the teaching of L2 phonology and pronunciation, there could be the potential to help learners acquire a more accurate L2 pronunciation more quickly and efficiently.

It may be helpful to reiterate the relatively low pedagogical investment that was made in bringing about the results in this study, both on the side of the educator and the student. From the educator's side, a simple 5-minute video was created that highlighted the target grapheme <d> and its native-like pronunciation between vowels. Materials were enhanced by simply adding the diacritic to <d> wherever it was presented for teaching/speaking purposes, as described in section 3.2.2.

Because the lesson was recorded as a video, it was easily disseminated to many students, it could be viewed asynchronously, and it could be reused in future classes. And because the lesson was recorded in English, it could be applied to any level of learner without any concern for comprehension of key concepts.

After watching and interacting with the pronunciation video, students were asked to carry out one reading per week for the next 4 weeks. Readings in this case were gleaned from available online source material, namely, Spanish fairy tales. The readings were relatively brief, about 400-500 words in length. Once the material was loaded into a word processor (in this case, Google Docs), I used the "Find and replace" function, searching simply for the character <d>. Whenever it was in the phonetic context for a [ð] pronunciation, I replaced the character with <ḏ>, and whenever it was in the context for [d], I replaced the character with <d̰>.

The above process describes all that was necessary on the side of the educator. In this case, no feedback was ever provided for the student. It was simply a one-directional pedagogical approach, with the student taking upon themselves the application of the lesson. The student watched the 5-minute video one week, and then recorded themselves reading one story per week over the next four weeks. Each recording of a story only lasted about 3-4 minutes. Thus, on the student's side, the entire lesson and follow up practice required only about 21 minutes, total.

Seeing as the total investment of time for both the educator and student was very low and that the results in this case were quite noticeable, incorporating such a diacritic-enhanced pronunciation lesson and practice materials into Spanish instruction could potentially be very profitable.

5.5.2 Suggested developments

One way that the above framework could likely be improved would be with oral feedback. In their meta-analysis of the effectiveness of L2 pronunciation instruction, Lee et al. (2015) found that providing feedback was associated with larger effects. Lyster et al. (2013) provide a state-of-the-art review on the topic of corrective feedback in L2 classrooms, and for L2 pronunciation they point out that short pronunciation-focused oral feedback can play an important role in L2 pronunciation development. Another form of feedback was provided in Lord (2005), in which students were able to use voice analysis software, where students could view spectrographs that gave visual feedback concerning their speech. The students in her class saw significant improvement across a range of L2 Spanish phonological items. Based on the findings from these other studies, a future iteration of the framework used in the present study would likely see further improvement if students were provided with some form (or various forms) of feedback.

While in-person, classroom feedback may be the most common approach to providing pronunciation feedback, some other options are available online. Using iSpraak (https://www.ispraak.net/) teachers can choose their target language and upload short passages in text form, which they can then send out to their students. Students can open the assignment, read the passage, and listen to a computer-generated audio of the passage before recording their own pronunciation of it. The web app will give the student an immediate score on how well they pronounced it (which score the teacher can also view). Another option is the website

www.forvo.com which has a database of many user-generated recordings of words in many languages. While this does not provide a way for students to record and grade their own pronunciation, it allows them to explore a variety of speakers' pronunciations of a given word or words. Hearing input from a variety of speakers can be helpful, "as it has been argued that speaker variability may be essential to the creation of robust perceptual categories" (Simon et al., 2010, p. 383). Finally, previous studies such as Lord (2005) and Herd at al. (2013) allowed students to view waveforms and spectrograms of their own speech, providing a visual feedback that could allow them to "see" things in their pronunciation that they may not hear on their own. Although the use of such software takes additional computer resources and some training for the students, it may provide detailed feedback that students could carry out and receive on their own.

Another intriguing study that may point to an improvement on the present approach is Solier et al. (2019). Their ground-breaking research with Moroccan learners of L2 French showed that pronunciation accuracy, perhaps contrary to natural intuitions, was improved more by a training centered on rewriting the letters of words than a training that only involved listening and repeating words. For beginning L1 Arabic learners of French, it was known that the students made considerable pronunciation errors in French vowels. Thus, the authors designed a study focused on the pronunciation of four word-final vowels. After a word-repetition pretest, the participants were divided into different training conditions. One training group heard and repeated back a list of words containing the target sounds, receiving corrective feedback (if there were any mispronunciations). Another group was asked to dictate (transcribe in written form) words that were heard auditorily. And another training group simply read and copied down in written form the orthographic forms of the words, not hearing any auditory forms. Somewhat surprisingly, in a posttest word-repetition task (which was the same as the

pretest), although each group saw improvement in pronunciation, the group which had passed through the training of copying down in written form the words actually saw greater improvement in their pronunciation. To reiterate, it was the pronunciations from the spoken word-repetition task which was used for the pretest and posttest data. What differed was the training they received in between those tasks. The results from Solier et al. (2019) indicate that copying down the written forms of the words containing the target word-final vowels seemingly did more to help their pronunciation of those sounds in a subsequent word-repetition task.

In the discussion of their findings, the authors elaborate on the particular reconstruction that goes on in the phonological system of an L2 learner when they process the visual forms of words in an alphabetic orthography. For languages in which the orthography contains patterns of grapheme to phoneme correspondence, if the learner engages in a higher degree of processing that orthography, making the correct orthographic form of the word more available in their speech processing, then the pronunciation of words can be more greatly improved.

The results of Solier et al. (2019) do suggest that the approach used in the present study may be improved with the addition of writing tasks. Specifically, it seems that if learners not only read the diacritic-enhanced words during the pronunciation lesson, but even rewrote the words with the diacritic, there would have been a higher degree of processing of the link between the target sound and the enhanced grapheme. According to Solier et al. (2019), this increased processing of the visual information may do even more to help learners to reconstruct their phonological system.

Because the process involved in the present study is related to phonotactics, such that the pronunciation of a letter/phoneme is dependent on the sounds around it, another natural development of this approach would be to teach the students to add diacritics themselves based

on those phonotactic rules. They could be trained to locate all contexts where the <d> is phrase-initial, or follows a nasal or <l>, marking it with a burst diacritic underneath. For all other contexts, they would mark it with the approximant diacritic. This would likely be a very fruitful exercise, as it would train them to be conscious of the phonetic environment, which is the critical factor in determining the correct /d/ pronunciation. After applying the phonotactic rule and drawing in the diacritics, they may progress to a stage where they can read unenhanced orthography and visualize those diacritics as a help in activating the proper pronunciation. Eventually, the phonotactic environment itself could be enough to activate the correct pronunciation directly.

In this section I have presented some pedagogical applications of the findings from this study. The possibility of giving students substantial help in their L2 phonological development for a very low investment of teacher and student resources makes the approach presented here very intriguing. Furthermore, a similar approach could be developed for many other L1/L2 target language items. Based on other current research in the area, some possible improvements were also presented, especially the additions of verbal feedback and hand-written activities involving the diacritic-enhanced graphemes. In the following section, I will present some limitations to the present study.

5.6 Limitations

While the present study does shed some light on the interaction between diacritic-enhanced text and L2 pronunciation, there were some limitations that both compel us to take some caution in interpreting these results, and also point the field to the need of further research focused on this matter.

The primary limitation in this study is a relatively low volume of data. The study controlled for a very homogenous pool of participants, namely, Spanish students in a particular, intermediate level, all with little to no native Spanish-speaking experiences, and all being taught, at the time of the study, by a non-native Spanish-speaking teacher. Among these, only participants who did not pronounce approximants more than 50% of the time on a paragraph-reading task were selected for analysis. This left us with a relatively small (n=25) group to work with. Recruiting from a larger range of class levels, or perhaps collecting data over a number of semesters from various classes, could help bring the group sizes up to a higher level, which would allow for a much more powerful statistical analysis. Concerning a control for the type of input that students would receive from their teacher, it would also be better to obtain some speech data from the instructors in order to verify their pronunciation of the target item.

Another limitation came about due to the study's goal of carrying out research with the highest possible ecological validity, incorporating study interventions that would fit seamlessly into the educator's curriculum and would be very natural on the student's side. The study materials were authentic fairy tales written in Spanish, which the learners used in order to practice verb forms (preterite and imperfect) that were a main focus of study in the course curriculum. The number of intervocalic, word-medial, post-tonic <d> targets was not very robust in some of the stories, thus limiting the number of tokens acquired at each session. Furthermore, many of these <d> targets happened to fall in contexts that caused suboptimal acoustic data or awkward pronunciations. Many <d> targets were in the final word of a sentence, where prosodic processes often lead to changes in pitch, volume, and voice quality (creakiness), which rendered the recordings of many targets unusable. Also, some <d> targets were found in words that were so long and difficult for this level of learner (e.g., *acantilados*)

that recordings of the word were often characterized by mispronunciations (such as metathesis, moving of the stress, etc.). All of these factors again limited the volume of data that was usable for this study. A future iteration of this study should take into account these factors and either lengthen the readings, or alter them somewhat to avoid any dependence on sentence final words or difficult words that would cause unusual effects in the reading task.

Also, because students would record the materials on their own devices, in whatever setting they chose (whether a noisy dorm building or a quiet study area), the microphone quality, the amount of background noise, and the type of voice that the student read in varied greatly, sometimes making the recording difficult to acoustically analyze. Obviously, all of these factors could be better controlled in a laboratory study. A metaanalysis of the effects of pronunciation instruction by Lee et al. (2015) presented the differences in laboratory-based versus classroom-based research. The effect size found in research done in laboratory-based pronunciation instruction can be higher, but it's possible that this could be a measure of participants' focused performance for the study as opposed to how they would perform if carrying out typical classwork. In the case of the present study, this limitation was a known one and the decision to use ecologically valid data collection means was intentional. Having some holes in the data is not necessarily problematic if there is a large enough volume of data and if methods of data analysis are used which allow for some variation in the number of tokens per subject, the number of subjects per group, etc. A future iteration of the present study may want to build in some additional scaffolds and supports to ensure thorough, high-quality data collection from the participants.

As is the case with many pronunciation studies, the one carried out here was not longitudinal in nature, and only gives us data points within about a 6-week window. While

seeing the effects of a treatment within this timeframe are interesting, the goal of language teaching is its providing lasting L2 ability to the learner. Of course, maintaining the ability to collect data over longer periods of time with students creates substantial challenges to researchers. Nonetheless, a future study should include at least a delayed posttest that measures student speech well after the treatment. This would provide evidence for whether or not the associations built up during the practice stage with the diacritics persists as the students interact with more and more authentic L2 orthography.

Finally, while the pronunciation of students during a paragraph-reading task may shine light on certain areas of phonological proficiency, a big question is whether or not students are learning to perceive and pronounce the target allophone in natural speech production. It can be notoriously difficult to elicit desired target segments in natural speech, however with intermediate-level students, tasks such as picture naming or telling a story based on a series of pictures could be used as part of a pre-posttest design. A future study on this subject should consider incorporating a test for this level of speech performance.

5.7 Further study

The findings of the present study do open up many interesting possibilities for research on other L2 pronunciation items for language pairs such as Spanish/English where shared graphemes have diverging phonetic realizations. In Spanish, one of the most obvious follow-up set of graphemes would be the other approximant allophones of the voiced stops <b> = /b/ and <g> = /g/. These allophones are known to be difficult for L1 English learners of Spanish, and their mispronunciation has been reported as a marker for foreign accentedness (Gonzalez-Bueno, 2019). A very similar pronunciation lesson could be designed showing the allophonic split of these phonemes, where the intervocalic context calls for the approximant realizations. It might

make sense to use the same diacritics to differentiate these allophones, where the burst realization takes on the combining asterisk below (<b̪> and <g̪>) and the allophones the combining breve below (<b̯> and <g̯>). In this way, the same conceptualization of manner of articulation is applied to all three voiced stop consonants.

There are a wide array of unicode diacritics available, and Spanish teachers of English-background students may want to consider other applications for problematic graphemes. Some examples are the tap /ɾ/ pronunciation of <r>, the /s/ pronunciation of <z>, the non-pronunciation of <h>, as well as of <u> in <qu> and <gu>. These are, of course, many of the target items in the typical introductory Spanish phonetics classroom. But since all of these sounds are actually fully pronounceable by even the novice learner, it seems there could be much benefit to the L2 student's learning and language experience if they could develop a more accurate phonological system from the very start of their learning experience. While the results of this study are limited in scope, it does indicate the possibility for potentially broad improvements in L2 Spanish pronunciation.

A number of language pairs were already brought up in prior studies on the effects of orthographic input on L2 phonology, especially L1/L2 language-learning scenarios that involve English. These include:

Table 5.1 - Language pairs with documented OI effects

L1/L2	OI effects
English/Spanish	fricative <v> in Zampini (1994)
Spanish/English	flapping of <d, dd, t, tt> in Vokic (2011)
Italian/English	pronouncing silent letters in Basetti and Atkinson (2015)
English/German	devoicing of final stops, Young-Scholten and Langer (2015)

English/Chinese pinyin	lexical tone, Showalter and Hayes Harb (2013)

It is quite possible that a combination of brief explicit pronunciation instruction and diacritic-enhanced reading activities could help L2 learners overcome many of these phenomena. There are, of course, a myriad of other similar orthography-related pronunciation difficulties that should arise in L1/L2 scenarios between languages that share the Latin alphabet (including English, French, Spanish, Italian, Dutch, Norwegian, German, Portuguese, Czech, Slovak, Hungarian, Polish, Danish, Welsh, Swedish, Icelandic, Finnish, and Turkish). There is no shortage of possibilities for future research on possible benefits of incorporating a methodology similar to what is presented in this study, but tailored to problematic issues in the specific L1/L2 language pair.

Another development that could be considered for future study is related to the varieties of possible Spanish realizations of the intervocalic /d/. In this study, for the sake of simplicity, only the [ð] allophone was taught in the pronunciation lesson, and it was the only target allophone that was analyzed. In the greater Spanish-speaking world, there are in fact other realizations of intervocalic /d/. One of the most common alternative pronunciations for intervocalic /d/ would be a deleted /d/, which can be frequent in the Spanish spoken in the Caribbean and Southern Spain. Solon and Kanwit (2022) provide a summary of this pronunciation and then explored L2 Spanish learners' sociophonetic competence related to /d/ deletion. In Central American and Highland Spanish varieties, there is also a documented maintanance of stop [d] in many contexts where a spirant would be pronounced in other varietes (Carrasco et al., 2012). While at the initial stages of Spanish learning, the simple, canonical [ð] may be a good starting point for instruction, as their competence level increases and they are introduced to a greater variety of Spanishes that are spoken, it would be important to teach

students about these varieties. It could even be possible to incorporate diacritics to help guide students into other possible pronunciation patterns.

5.8 Conclusion

In conclusion, the present study sought to explore the possible benefits of manipulating the visual information provided to L2 learners in their orthographic input in order to assist them in their acquisition of more accurate L2 pronunciation. Because most L2 curricula lean heavily on orthography from the outset of learning, it is important to understand the effects of such input on L2 phonological development, especially if modified pedagogical approaches would help learners acquire a more effective representation of L2 phonological patterns. Relevant theoretical constructs based in a connectionist framework were presented as possible underlying linguistic structures that may be involved, especially as conceptualized through the Dual-Route Cascading model (Coltheart et al., 2001). An approach of incorporating diacritic-enhancement, that up until the time of writing had only been researched in a handful of L1/L2 settings, in combination with explicit phonetic instruction was used for English learners of Spanish in order to reconstruct cognitive associations related to pronunciation of L2 Spanish <d>. While a brief pronunciation lesson on the target item was alone unable to make lasting differences in students' pronunciation (as measured by paragraph-reading tasks), when that same lesson incorporated diacritic-enhanced text for target <d>, there was evidence for substantial improvement in pronunciation of the item when the diacritic enhancement was used in the subsequent paragraph-reading tasks. Furthermore, these improvements were found to still be intact in a paragraph-reading task after the diacritics were removed, suggesting that learners had associated the desired pronunciation with the authentic <d> itself.

Because the results of the present study were obtained using a relatively small sample size of students, some caution should be exercised in the interpretation of these results. Further study of this particular effect must be carried out in order to corroborate what was found here. Nonetheless, there appears to be the potential for real benefits to incorporating diacritic-enhancement in L1/L2 learning situations where the two languages share an orthographic system but have diverging phonetic realizations of graphemes.

BIBLIOGRAPHY

Alsadoon, R., & Heift, T. (2015). Textual input enhancement for vowel blindness: A study with Arabic ESL learners. *The Modern Language Journal*, *99*(1), 57-79.

Bajuniemi, A. (2013). Teaching intervention on the pronunciation of Spanish intervocalic/d. In *Selected Proceedings of the 15th Hispanic Linguistics Symposium*, 39-50. Somerville, MA: Cascadilla Proceedings Project.

Bassetti, B. (2006). Orthographic input and phonological representations in learners of Chinese as a foreign language. *Written Language & Literacy*, *9*(1), 95-114.

Bassetti, B. (2007). Effects of hanyu pinyin on pronunciation in learners of Chinese as a foreign language. *Birkbeck ePrints*. https://eprints.bbk.ac.uk/id/eprint/528/1/528.pdf

Bassetti, B., & Atkinson, N. (2015). Effects of orthographic forms on pronunciation in experienced instructed second language learners. *Applied Psycholinguistics*, *36*(1), 67-91.

Bassetti, B., Escudero, P., & Hayes-Harb, R. (2015). Second language phonology at the interface between acoustic and orthographic input. *Applied Psycholinguistics*, *36*(1), 1-6.

Birdsong, D., Gertken, L. M., & Amengual, M. (2012). Bilingual language profile: An easy-to-use instrument to assess bilingualism. *COERLL, University of Texas at Austin*.

Blair, K., & Lease, S. (2021). An examination of social, phonetic, and lexical variables on the lenition of intervocalic voiced stops by spanish heritage speakers. *Languages*, *6*(2), 108. https://doi.org/10.3390/languages6020108

Boersma, P., & Weenink, D. (2022). Praat: A system for doing phonetics by computer, 2000. https://www.fon.hum.uva.nl/praat/

Bürki, A., Spinelli, E., & Gaskell, M. G. (2012). A written word is worth a thousand spoken words: The influence of spelling on spoken-word production. *Journal of Memory and Language*, *67*(4), 449-467.

Carrasco, P., Hualde, J. I., & Simonet, M. (2012). Dialectal differences in Spanish voiced obstruent allophony: Costa Rican versus Iberian Spanish. *Phonetica*, *69*(3), 149-179.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*(1), 5-51.

Colantoni, L., Escudero, P., Marrero-Aguiar, V., & Steele, J. (2021). Evidence-based design principles for Spanish pronunciation teaching. *Frontiers in Communication*, *6*, 639889.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*(4), 589-608.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204-256.

Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics*, *34*(2), 269-284.

DeKeyser, R. M. (2020). Skill acquisition theory. In B. VanPatten, G.D. Keating, & S. Wulff (Eds.), *Theories in Second Language Acquisition. An Introduction* (3rd ed., pp. 83-104). New York, NY: Routledge.

Eddington, D. (2011). What are the contextual phonetic variants of /β, ð, ɣ/ in colloquial Spanish?. *Probus*, *23*(1), 1-19.

Ellis, N. C. (2006). Cognitive perspectives on SLA: The associative-cognitive CREED. *Aila Review*, *19*(1), 100-121.

Ellis, N. C. (2011). Implicit and explicit SLA and their interface. In C. Sanz and R. Leow (Eds.), *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA and bilingualism* (pp. 35-47). Georgetown University Press.

Elliott, A. R. (1997). On the teaching and acquisition of pronunciation within a communicative approach. *Hispania*, 95-108.

Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, *36*(2), 345-360.

Face, T. L., & Menke, M. R. (2009). Acquisition of the Spanish voiced spirants by second language learners. In *Selected proceedings of the 11th Hispanic linguistics symposium*, 39-52. Somerville, MA: Cascadilla Proceedings Project.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*(1), 47-65.

Flege, J. E., & Bohn, O. S. (2021). The revised speech learning model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3-83). Cambridge University Press.

Glushko, R. J. (1981). Principles for pronouncing print: The psychology of phonography. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading* (pp. 61-84). Hillsdale, NJ: Erlbaum.

González-Bueno, M. (2019). Suggestions for teaching Spanish voiced stops /b, d, g/ and their lenited allophones [β̞, ð̞, ɣ̞]. In R. Rao (Ed.) *Key issues in the teaching of Spanish pronunciation: From description to pedagogy* (pp. 60-83). Routledge.

Han, J. I., & Kim, J. Y. (2017). The influence of orthography on the production of alphabetic, second-language allophones by speakers of a non-alphabetic language. *Journal of Psycholinguistic Research*, *46*, 963-982.

Hayes-Harb, R., Nicol, J., & Barker, J. (2010). Learning the phonological forms of new words: Effects of orthographic and auditory input. *Language and Speech*, *53*(3), 367-381.

Hayes-Harb, R., & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. *Language Teaching*, *54*(3), 297-326.

Herd, W., Jongman, A., & Sereno, J. (2013). Perceptual and production training of intervocalic /d, ɾ, r/ in American English learners of Spanish. *The Journal of the Acoustical Society of America*, *133*(6), 4247-4255.

Hualde, J. I. (2005). *The sounds of Spanish with audio CD*. Cambridge University Press.

Hualde, J. I., Nadeu, M., & Simonet, M. (2010). Lenition and phonemic contrast in Majorcan Catalan. In Colina, S., Olarrea, A., & Carvalho, A. M. (Eds.) *Romance Linguistics 2009: Selected Papers from the 39th Linguistic Symposium on Romance Languages (LSRL), Tucson, Arizona, March 2009*. John Benjamins Publishing.

Jackson, J. L. (2016). *The effects of novel orthographic input and phonetic instruction in second language phonological acquisition* (Order No. 10159125). Available from Linguistics and Language Behavior Abstracts (LLBA); ProQuest Dissertations & Theses Global. (1840191290).

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: the orthographic depth hypothesis. *Advances in Psychology*, *94*, 67-84.

Koda, K. (1992). The effects of lower-level processing skills on FL reading performance: Implications for instruction. *The Modern Language Journal*, *76*(4), 502-512.

Koda, K. (2007). Reading and Language Learning: Crosslinguistic Constraints on Second Language Reading Development. *Language Learning*, *57*, 1-44.

Lado R. (1957). *Language teaching: A scientific approach*. New York, NY: McGraw Hill.

Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, *36*(3), 345-366.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*(4), 492.

Looney, D., & Lusin, N. (2019, June). Enrollments in languages other than English in United States institutions of higher education, Summer 2016 and Fall 2016. In *Modern language association*. Modern Language Association. 26 Broadway 3rd Floor, New York, NY. https://files.eric.ed.gov/fulltext/ED599007.pdf

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

Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, *46*(1), 1-40.

MacWhinney, B. (1987). Applying the competition model to bilingualism. *Applied Psycholinguistics*, *8*(4), 315-327.

MacWhinney, B. (1997). Second language acquisition and the competition model. In A. de Groot and J. Kroll (Eds.) *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 113-142). Mahwah, N.J.: Lawrence Erlbaum Associates.

Mellado, A. (2012). La pronunciación, un objetivo crucial para la enseñanza/aprendizaje de la competencia comunicativa. *Aproximación a la enseñanza de la pronunciación en el aula de español*, 11-41.

Mines, M. A., Hanson, B. F., & Shoup, J. E. (1978). Frequency of occurrence of phonemes in conversational English. *Language and Speech*, *21*(3), 221-241.

Morgan, T. (2010). *Sonidos en contexto*. New Haven, CT: Yale University Press.

Odlin, T. (1989). *Language Transfer: Cross-Linguistic Influence in Language Learning* (Cambridge Applied Linguistics). Cambridge: Cambridge University Press.

Offerman, H. M. (2020). *The effects of pronunciation instruction on L2 production and L2 perception in Spanish: A comparative analysis* [Doctoral dissertation, Purdue University]. *ProQuest.*

Pytlyk, C. (2017). Are orthographic effects language specific? The influence of second language orthography on second language phoneme awareness. *Applied Psycholinguistics*, *38*(2), 233-262.

Rafat, Y. (2011). *Orthography-induced Transfer in the Production of Novice Adult English-speaking Learners of Spanish* [Doctoral dissertation, University of Toronto]. https://central.bac-lac.gc.ca/.item?id=NR78092&op=pdf&app=Library&oclc_number=1019490380

Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition*, *19*(1), 1-30.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*(2), 151-218.

Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *72*(2), 95-129.

Showalter, C. E., & Hayes-Harb, R. (2013). Unfamiliar orthographic information and second language word learning: A novel lexicon study. *Second Language Research*, *29*(2), 185-200.

Showalter, C. E., & Hayes-Harb, R. (2015). Native English speakers learning Arabic: The influence of novel orthographic information on second language phonological acquisition. *Applied Psycholinguistics*, *36*(1), 23-42.

Simon, E., & Van Herreweghe, M. (2010). The relation between orthography and phonology from different angles: Insights from psycholinguistics and second language acquisition. *Language and Speech*, *53*(3), 303-306.

Simon, E., Chambless, D., & Alves, U. K. (2010). Understanding the role of orthography in the acquisition of a non-native vowel contrast. *Language Sciences*, *32*(3), 380-394.

Simonet, M., Hualde, J. I., & Nadeu, M. (2012). Lenition of /d/ in spontaneous Spanish and Catalan. In *13th Annual Conference of the International Speech Communication Association 2012, INTERSPEECH 2012*, 1416-1419. https://www.isca-speech.org/archive_v0/archive_papers/interspeech_2012/i12_1416.pdf

Solier, C., Perret, C., Baqué, L., & Soum-Favaro, C. (2019). Written training tasks are better than oral training tasks at improving L2 learners' speech production. *Applied Psycholinguistics*, *40*(6), 1455-1480.

Solon, M., & Kanwit, M. (2022). New methods for tracking development of sociophonetic competence: Exploring a preference task for Spanish /d/ deletion. *Applied Linguistics*, *43*(4), 805-825.

Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*, *36*(3), 326-344.

Vokic, G. (2011). When alphabets collide: Alphabetic first-language speakers' approach to speech production in an alphabetic second language. *Second Language Research*, *27*(3), 391-417.

Waltmunson, J. C. (2005). *The relative degree of difficulty of L2 Spanish /d, t/, trill, and tap by L1 English speakers: Auditory and acoustic methods of defining pronunciation accuracy*. [Doctoral Dissertation, University of Washington].

Warner, N., & Tucker, B. V. (2011). Phonetic variability of stops and flaps in spontaneous and careful speech. *The Journal of the Acoustical Society of America*, *130*(3), 1606-1617.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*(1), 1-25. https://doi.org/10.1016/S0749-596X(03)00105-0

Weinreich, U. (1979) *Languages in Contact : Findings and Problems*, De Gruyter, Inc. *ProQuest Ebook Central*, https://ebookcentral.proquest.com/lib/wisc/detail.action?docID=929128.

Young-Scholten, M. (1995). The negative effects of 'positive' evidence on L2 phonology. In L. Eubank, et al., (Eds.) *The Current State of Interlanguage: Studies in honor of William E. Rutherford*, 107-122. John Benjamins. http://ebookcentral.proquest.com/lib/wisc/detail.action?docID=680946.

Young-Scholten, M. (2002). Orthographic input in L2 phonological development. In P. Burmeister, T. Piske, and A. Rohde (Eds). *An integrated view of language development: Papers in honor of Henning Wode* (pp. 263-279). Wissenschaftlicher Verlag.

Young-Scholten, M. (2004). Prosodic constraints on allophonic distribution in adult L2 acquisition. *International Journal of Bilingualism*, *8*(1), 67-77.

Young-Scholten, M., & Langer, M. (2015). The role of orthographic input in second language German: Evidence from naturalistic adult learners' production. *Applied Psycholinguistics*, *36*(1), 93-114.

Zampini, M. L. (1994). The role of native language transfer and task formality in the acquisition of Spanish spirantization. *Hispania*, 470-481.

APPENDICES

Appendix A Bilingual Language Profile

Bilingual Language Profile: English-Spanish

We would like to ask you to help us by answering the following questions concerning your language history, use, attitudes, and proficiency. This survey was created with support from the Center for Open Educational Resources and Language Learning at the University of Texas at Austin to better understand the profiles of bilingual speakers in diverse settings with diverse backgrounds. The survey consists of 19 questions and will take less than 10 minutes to complete. This is not a test, so there are no right or wrong answers. Please answer every question and give your answers sincerely. Thank you very much for your help.

I. Biographical Information

Name _____ Today's Date ____/____/____

Age____ ☐Male / ☐Female Current place of residence: city/state_____ country_____

Highest level of formal education: ☐ Less than high school ☐ High school ☐ Some college ☐ College (B.A., B.S.) ☐ Some graduate school ☐ Masters ☐ PhD/MD/JD ☐ Other: _____

II. Language history

In this section, we would like you to answer some factual questions about your language history by placing a check in the appropriate box.

1. At what age did you **start learning** the following languages?

English

☐ Since birth ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

Spanish

☐ Since birth ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

2. At what age did you **start to feel comfortable** using the following languages?

English

☐ As early as I can remember ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+ ☐ not yet

Spanish

☐ As early as I can remember ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+ ☐ not yet

3. How many years of **classes (grammar, history, math, etc.)** have you had in the following languages (primary school through university)?

English

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

Spanish

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

4. How many years have you spent in a **country/region** where the following languages are spoken?

English

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

Spanish

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

5. How many years have you spent in a **family** where the following languages are spoken?

English

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

Spanish

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

6. How many years have you spent in a **work environment** where the following languages are spoken?

English

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

Spanish

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

III. Language use

In this section, we would like you to answer some questions about your language use by placing a check in the appropriate box. Total use for all languages in a given question should equal 100%.

7. In an average week, what percentage of the time do you use the following languages **with friends**?

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
English	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Spanish	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Other languages	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

8. In an average week, what percentage of the time do you use the following languages **with family**?

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
English	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Spanish	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Other languages	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

9. In an average week, what percentage of the time do you use the following languages **at school/work**?

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
English	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Spanish	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Other languages	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

10. When you talk to yourself, how often do you **talk to yourself** in the following languages?

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
English	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Spanish	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Other languages	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

11. When you count, how often do you **count** in the following languages?

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
English	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Spanish	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Other languages	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

IV. Language proficiency

In this section, we would like you to rate your language proficiency by giving marks from 0 to 6.

	0=not well at all						*6=very well*
12. a. How well do you speak **English**?	☐0	☐1	☐2	☐3	☐4	☐5	☐6
b. How well do you speak **Spanish**?	☐0	☐1	☐2	☐3	☐4	☐5	☐6
13. a. How well do you understand **English**?	☐0	☐1	☐2	☐3	☐4	☐5	☐6
b. How well do you understand **Spanish**?	☐0	☐1	☐2	☐3	☐4	☐5	☐6
14. a. How well do you read **English**?	☐0	☐1	☐2	☐3	☐4	☐5	☐6
b. How well do you read **Spanish**?	☐0	☐1	☐2	☐3	☐4	☐5	☐6
15. a. How well do you write **English**?	☐0	☐1	☐2	☐3	☐4	☐5	☐6
b. How well do you write **Spanish**?	☐0	☐1	☐2	☐3	☐4	☐5	☐6

V. Language attitudes

In this section, we would like you to respond to statements about language attitudes by giving marks from 0-6.

	0=disagree						*6=agree*
16. a. I feel like myself when I speak **English**.	☐0	☐1	☐2	☐3	☐4	☐5	☐6
b. I feel like myself when I speak **Spanish**.	☐0	☐1	☐2	☐3	☐4	☐5	☐6
17. a. I identify with an **English-speaking** culture.	☐0	☐1	☐2	☐3	☐4	☐5	☐6
b. I identify with a **Spanish-speaking** culture.	☐0	☐1	☐2	☐3	☐4	☐5	☐6
18. a. It is important to me to use (or eventually use) **English** like a native speaker.	☐0	☐1	☐2	☐3	☐4	☐5	☐6
b. It is important to me to use (or eventually use) **Spanish** like a native speaker.	☐0	☐1	☐2	☐3	☐4	☐5	☐6
19. a. I want others to think I am a native speaker of **English**.	☐0	☐1	☐2	☐3	☐4	☐5	☐6
b. I want others to think I am a native speaker of **Spanish**.	☐0	☐1	☐2	☐3	☐4	☐5	☐6

Appendix B Examples of session readings

The following is an extract from the comparison and control group readings (plain orthography) at Session 5. This is an example of what all groups would have read from at Sessions 1 and 6, and what the comparison and control groups would have read from at Sessions 1, 3, 4, 5, and 6.

Hace muchos años, hubo un leñador. Él vivía en lo profundo del bosque con su esposa. Ella era una mujer buena y amable. Tenían dos hijos: Hansel y Gretel.

La familia no tenía mucho dinero, pero su casa siempre estaba calientita y los niños siempre tenían sopa y pan para comer. Nunca pasaban hambre.

Un año, hubo un invierno terriblemente frío y la madre de Hansel y Gretel se enfermó y murió.

Con el tiempo, el leñador se casó con otra mujer, pero ella no era bondadosa ni amable. Era malvada y egoísta. Quería ser rica. Y quería comer alimentos caros como asado de res, puré de papas, arvejas, choclo, tortas de limón y frutas exóticas de soleadas tierras.

And the following is an extract from experimental group reading (diacritic-enhanced orthography) at Session 5. This is an example of what the experimental group would have read from at Sessions 3, 4, and 5.

Hace muchos años, hubo un leñaḡor. Él vivía en lo profunḍo ḡel bosque con su esposa. Ella era una mujer buena y amable. Tenían ḍos hijos: Hansel y Gretel.

La familia no tenía mucho ḡinero, pero su casa siempre estaba calientita y los niños siempre tenían sopa y pan para comer. Nunca pasaban hambre.

Un año, hubo un invierno terriblemente frío y la maḡre ḡe Hansel y Gretel se enfermó y murió.

Con el tiempo, el leñaḡor se casó con otra mujer, pero ella no era bonḍaḡosa ni amable. Era malvaḡa y egoísta. Quería ser rica. Y quería comer alimentos caros como asaḡo ḡe res, puré ḡe papas, arvejas, choclo, tortas ḡe limón y frutas exóticas ḡe soleaḡas tierras.

Appendix C Statistics for other acoustic variables

Acoustic variables such as relative intensity (RI) or maximum relative velocity (MRV) were suggested in the literature for characterizing production of voiced stops (see section 3.3.1). But for English speaking learners of L2 Spanish, <d> causes both L1 tap [ɾ] or stop [d] realizations as well as L2 stop [d] or approximant [ð] realizations. These approximant realizations can either be fricatives [ð̞] or true approximants [ð̞]. In this study, because there was not control for quality of microphone, speaking environment, speaking volume, etc., it was necessary to normalize the data, which was accomplished by measuring RI as a percent increase in intensity (RI%) from the low point in the consonantal segment to the high point in the following vowel. The MRV was weighted (WMRV) by dividing it by the maximum intensity of the following vowel.

There is a great deal of overlap in these variables for the 4 possible realizations as can be seen in Figures A.1 and A.2. While the taps and the stops are clearly discernible, the true approximant productions [ð̞] can sometimes be very brief and overlap with the taps (especially perceptual taps) in terms of duration and RI% / WMRV, and the fricative productions [ð̞] can sometimes have a strong, sudden change in intensity, causing an overlap with the stops. But by categorizing the productions by duration and spectral pattern instead of by duration with these variables, the allophones were reliably teased apart and thus that was the approach used, as described in 3.3.2. Further statistics related to RI% and WMRV are also provided in Table A.1.

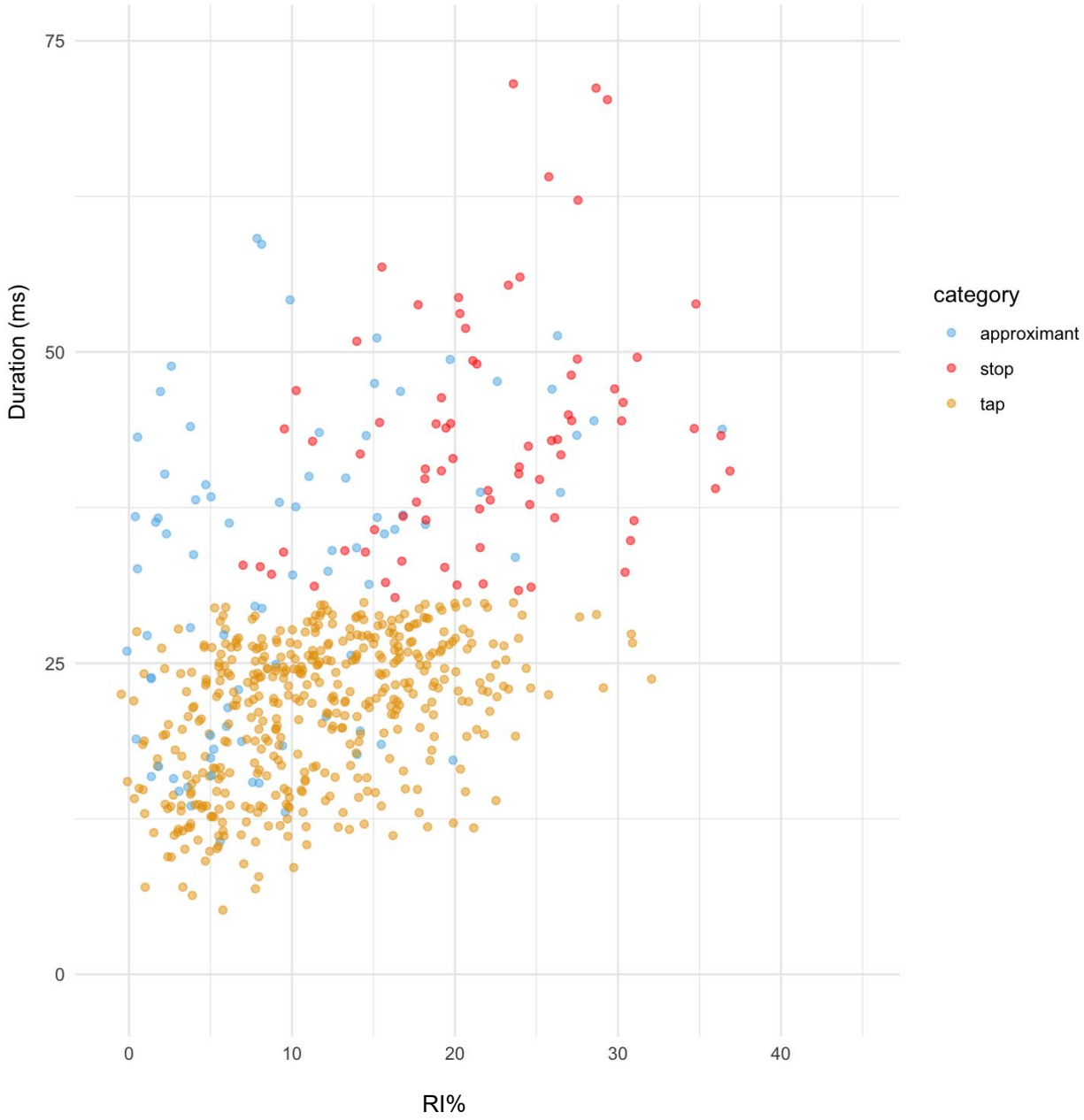


Figure A.1 Scatterplot of duration and RI% for all productions in session 1, 5, and 6. RI% is the percent increase in intensity, from the low point in the consonantal segment to the high point in the following vowel.

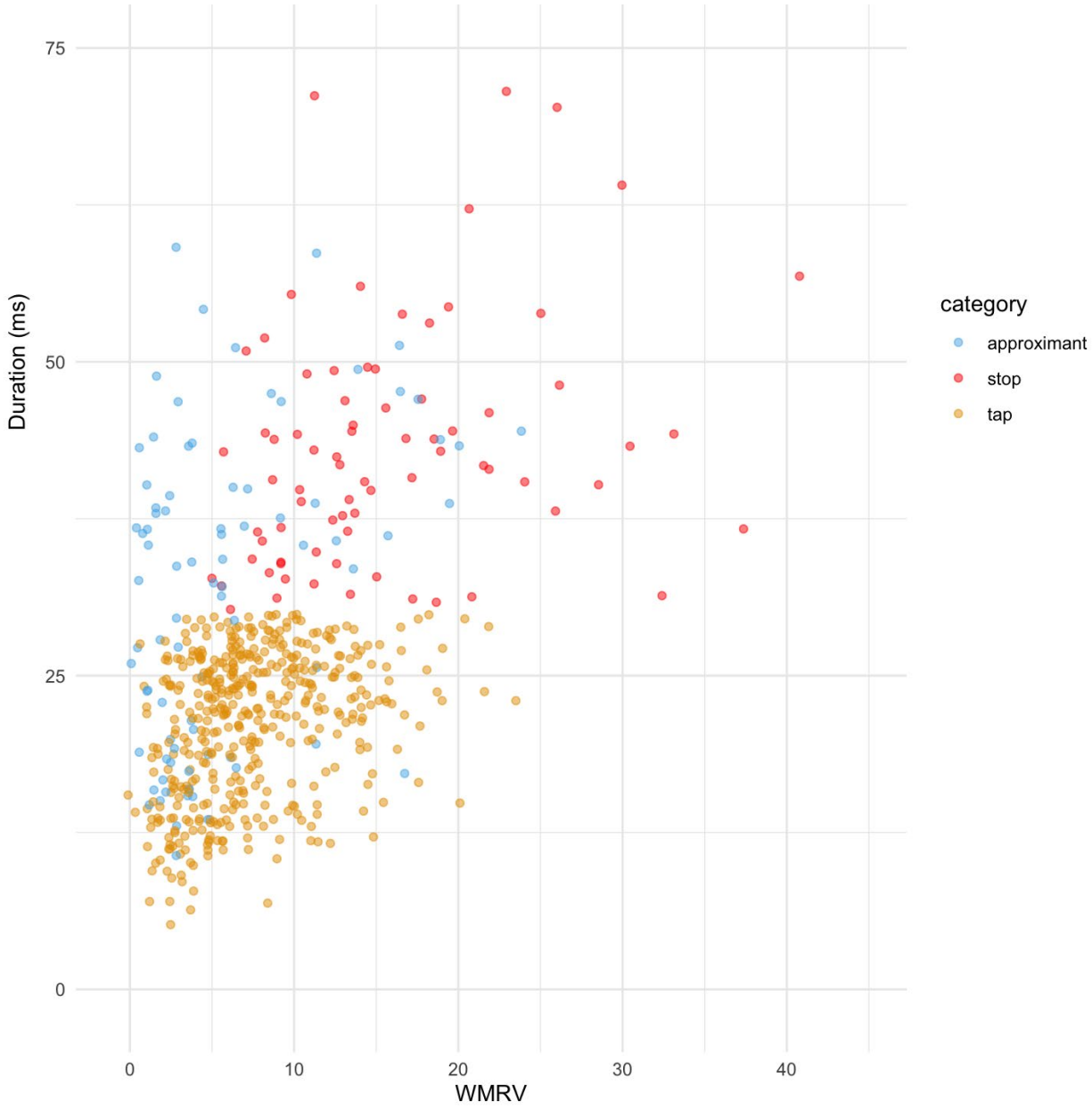


Figure A.2 Scatterplot of duration and WMRV for all productions in session 1, 5, and 6. MRV is the greatest change in intensity recorded over very brief, 1ms intervals during the release burst. This value is weighted by dividing the value by the maximum intensity of the following vowel, giving us the WMRV.

Table A.1 RI% and WMRV statistics for all groups and sessions

category	group	session	n	mean RI%	SD RI%	mean WMRV	SD WMRV	mean duration	SD duration	total	percent
approximant	Comparison	1	5	6.18	7.11	4.51	6.28	24.26	10.41	76	0.07
stop	Comparison	1	5	26.19	6.94	10.99	1.60	42.19	9.68	76	0.07
tap	Comparison	1	66	10.63	6.46	6.89	3.95	20.55	5.15	76	0.87
approximant	Control	1	5	5.57	3.89	4.10	3.58	24.30	9.00	59	0.08
stop	Control	1	5	18.74	7.03	18.03	12.18	44.69	15.43	59	0.08
tap	Control	1	49	12.56	6.39	9.34	4.97	21.65	5.68	59	0.83
approximant	Experimental	1	5	11.90	8.09	6.93	5.76	26.31	9.27	52	0.10
tap	Experimental	1	47	11.78	6.41	8.16	3.97	21.89	5.68	52	0.90
approximant	Comparison	2	33	18.39	7.52	11.77	5.49	50.73	15.58	50	0.66
stop	Comparison	2	9	25.08	7.79	17.06	5.89	58.10	10.48	50	0.18
tap	Comparison	2	8	13.41	5.95	9.69	3.72	22.77	6.13	50	0.16
approximant	Experimental	2	17	15.58	10.93	8.91	6.85	52.79	12.95	30	0.57
stop	Experimental	2	9	31.12	5.97	23.20	5.61	77.09	11.58	30	0.30
tap	Experimental	2	4	15.40	5.26	12.77	0.93	20.35	5.29	30	0.13
approximant	Comparison	5	14	13.39	10.69	7.63	6.42	41.52	11.60	108	0.13
stop	Comparison	5	27	21.48	7.13	19.22	8.09	41.58	7.52	108	0.25
tap	Comparison	5	67	13.71	6.24	10.29	4.93	20.74	5.66	108	0.62
approximant	Control	5	7	8.16	6.15	4.20	2.20	28.43	11.09	84	0.08
stop	Control	5	5	20.46	5.20	11.75	2.83	40.96	5.72	84	0.06
tap	Control	5	72	8.99	5.44	5.21	2.74	18.91	6.65	84	0.86
approximant	Experimental	5	24	11.81	9.09	8.07	7.15	37.65	11.21	70	0.34
stop	Experimental	5	11	26.15	5.86	20.77	6.18	50.44	12.39	70	0.16
tap	Experimental	5	35	12.76	7.53	9.56	4.94	21.61	5.61	70	0.50
approximant	Comparison	6	3	8.40	3.10	3.40	0.43	28.74	14.11	67	0.04
stop	Comparison	6	16	20.08	7.82	10.30	4.44	41.52	9.92	67	0.24
tap	Comparison	6	48	12.17	5.55	6.47	2.27	23.30	4.92	67	0.72
approximant	Control	6	3	7.41	6.21	4.18	2.61	17.42	1.43	50	0.06
stop	Control	6	1	29.66	NA	17.88	NA	47.03	NA	50	0.02
tap	Control	6	46	11.06	5.68	5.66	2.59	21.09	6.05	50	0.92
approximant	Experimental	6	16	7.83	4.47	3.09	1.75	26.75	8.77	46	0.35
stop	Experimental	6	6	19.04	5.49	10.77	2.73	40.71	7.15	46	0.13
tap	Experimental	6	24	10.01	4.77	5.08	2.35	24.14	4.27	46	0.52

Appendix D Details of the logistic regression model

Due to the sample size and the unequal number of tokens between participants and groups, the data were analyzed using a Bayesian logistic regression, using the bernoulli family and logit linking function in the R package brms. The model priors followed the default in brms, using the Student's T distribution with 3 degrees of freedom. A single model was fit to the data where the outcome variable was the probability of realizing each <d> as an approximant. The outcome was predicted by the population (fixed) effects of group (3 levels: labeled "control", "experimental", and "comparison" in the data), and session (3 levels: labeled "1", "5" and "6"). The random effects included a random slope by participant for session. The model was run with 2000 iterations of Hamiltonian Monte-Carlo sampling (1000 warm up), across 4 chains and 8 processing cores.

Table A.2 shows the output of the Bayesian logistic regression model. Overall, the model's conditional r squared value suggests that the model explained 29 percent of the variance observed in the data. Figure A.1 visualizes the model in a forest plot. Each density plot is a summary of the most probable parameter estimates from the Bayesian model, and the values on the right side of the plot are the 95% highest density interval in log-odds.

Table A.2 Output of Bayesian logistic regression model

	Estimate	**Est.Error**	**Q2.5**	**Q97.5**
Intercept	-3.534676	0.8330690	-5.377699	-2.122251
session5	0.8238395	0.8224129	-0.775344	2.445888
session6	-0.638588	1.0724133	-3.106340	1.204214
groupControl	0.7301176	1.0885412	-1.338189	2.962385
groupExperimental	0.8172550	1.1415364	-1.299490	3.155540
session5:groupControl	-1.070082	1.1012484	-3.288601	1.056761
session6:groupControl	-0.138721	1.3137497	-2.608269	2.576936
session5:groupExperimental	0.9359264	1.0866427	-1.176458	3.124677

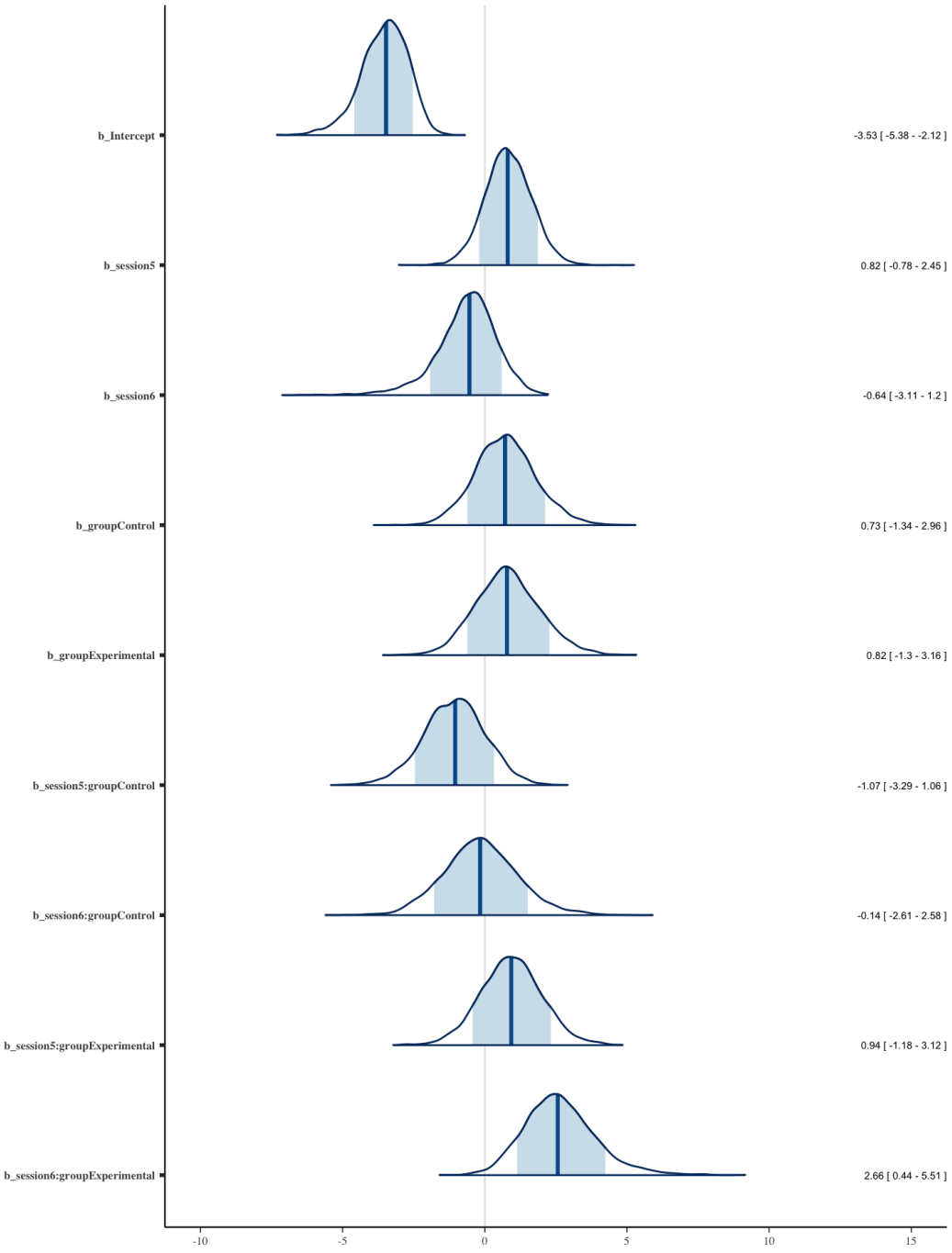


Figure A.3 Density plots of the most probable parameter estimates from the Bayesian model.