The L2 Production of Spanish Taps and Trills by L1 Romanian and L1 French Speakers

by

Malina Radu

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy

Department of Spanish & Portuguese
University of Toronto

The L2 Production of Spanish Taps and Trills by L1 Romanian and L1 French Speakers

Malina Radu

Doctor of Philosophy

Department of Spanish & Portuguese
University of Toronto

2020

Abstract

This dissertation explores two learning problems in L2 phonology: (1) acquiring a new L2 distribution when both sounds exist in the L1 as allophones; (2) acquiring the same L2 distribution when neither sound has a direct L1 counterpart. Previous literature has investigated these problems separately, focusing more on the first situation than the second. To explore these scenarios, I analysed the L2 production of Spanish taps and trills, which are contrastive word-medially intervocalically, by 10 L1 Romanian and 10 French speakers. 13 native Spanish controls were also included. These language pairings were chosen because Romanian has both rhotics as allophones (albeit in free variation), with the tap being the most common, and the trill sometimes occurring word-initially, while French has neither. Data was elicited using two tasks: a sentence reading (more formal, orthography) and a picture description task (less formal, no orthography). The stimuli included nonce and real words containing word-initial and word-medial rhotics. Because a secondary objective was to provide empirical evidence and acoustic analyses for rhotics in Romanian, participants also completed similar tasks in their L1.

The results confirm previous observations that the most common Romanian rhotic is the tap. The trill mainly appears word-initially. The L2 data showed that the tap was less difficult than the trill, with both learner groups producing taps at high rates. However, both learner groups had low trilling rates. Though the L1 Romanian group produced more trills than the L1 French group overall, when the position of trills was considered, there were no differences between groups. Specifically, both learner groups had significantly higher trilling rates word-medially than initially, corroborating previous research, and there were no group differences. Furthermore, speakers produced more of the expected rhotics in the sentence reading than in the picture description task. Overall, given the low trilling rates, these findings suggest that recategorizing L1 allophones may be just as challenging as learning a new L2 distribution that does not have L1 counterparts, though the difficulties may be due to different reasons. Finally, this study also provides acoustic data supporting previous observations regarding Romanian rhotics.

Acknowledgments

Being a graduate student has been a part of my identity for many years, and so, this PhD means much more to me than its culmination point. I learned about far more than Linguistics and academia. I learned about me – my limits and my fears. And then, most importantly, I learned how to overcome them.

This thesis would not have been possible without the support and encouragement, both academic and personal, of so many important people.

My deepest gratitude goes to my committee members, or the "dream team", as someone once called them: Laura Colantoni, Ana Teresa Pérez-Leroux, and Jeffrey Steele. To my supervisor, Laura, thank you for everything you have taught me, years before you became my thesis supervisor. Thank you for encouraging me to be confident in voicing my opinions. I owe a large part of my personal and professional development to your guidance for the past nine years. Ana Teresa, thank you for always being supportive and encouraging of my work. I am grateful to have had the chance to learn from you. Jeff, I am grateful for the time you took to offer careful, thought-provoking feedback and comments so that I could inspect things with a more critical eye, and, ultimately, become a better scholar.

To my external committee members, Anabela Rato and Christine Shea: Thank you for your kind words on my thesis. I am grateful for your thoughtful feedback and questions, and I very much appreciated your warmth and positivity at the defence.

I would also like to thank Anna Limanni and Yadira Álvarez López for being my introduction into the world of Linguistics and inspiring me so early in my academic career. It is thanks to your encouragement and passion for this field that I became interested in it.

During my time as a student, I met several people who quickly became my friends, and made graduate school a memorable experience. Olivia and Olga, thank you for always being there to listen to all things thesis, teaching, and life in general. My heartfelt gratitude to a number of other people for their help, support, and friendship over the years: Gaby, Matt, Ruth, Ailén, Petre, Lulu and Meï-Lan.

I would also like to thank the Department of Spanish & Portuguese and SGS. To all of the wonderful faculty and staff in the Department, thank you for your support, guidance and words of encouragement at various points throughout this process. I am also grateful for the financial assistance I have received, which allowed me to collect data, as well as travel for conferences, both within Canada and internationally. This financial support includes a SSHRC Doctoral, a Joseph Gulsoy Dissertation Scholarship, a Doctoral Completion Award, SGS Travel and Conference grants and a Milton A. Buchanan Fellowship.

I extend my sincerest gratitude to all of my participants, and the people who were instrumental in helping me recruit them (Professor Laura Badescu, Professor Laura Spinu, and Ruth Martínez, among others). This thesis truly would not have been possible without my participants' willingness to take part in the experiments.

Importantly, I am also grateful for all of my family and friends outside of academia. Thank you for checking on me, for being in my life, for allowing me to maintain balance, and for reminding me of who I am.

Last but not least, thank you to those who have been in my life for much, much longer than the PhD. To my mom and dad: thank you for sacrificing so much to give me opportunities I would not have had otherwise. Thank you for encouraging me when I felt like I could not push forward and thank you for teaching me to believe in myself. *Vă iubesc de la cer până la pământ.* To Matthew: I am grateful for far more than I could ever write here. Thank you for always making me laugh and for reminding me that I can do it (whatever "it" may be).

Table of Contents

Contents

List of Tables

List of Figures

List of Appendices

Chapter 1

Introduction

1.1 Main objectives

The primary objective of the present dissertation is to contribute to our understanding of two different learning problems in second language (L2) phonology via the investigation of novel language pairings, namely Romanian-Spanish and French-Spanish. The two learning scenarios are as follows: (1) acquiring two (quasi)phonemes when both sounds exist in the L1, but as allophones instead of phonemes; (2) acquiring the same two (quasi)phonemes when neither sound exists in the L1. Specifically, this dissertation examines the L2 production of Spanish taps and trills in three linguistic contexts, (1a-c) by L1 Romanian and French speakers. Spanish has a tap and trill, which constitute a quasi-phonemic contrast – they contrast only word-medially intervocalically, (1b-c); word-initially, only the trill occurs (1a; Hualde, 2005a, 2005b).

(1)	a.	*rata*	/ˈrata/	‘rat’	(word-initial trill)
	b.	*carro*	/ˈkaro/	‘car’	(word-medial intervocalic trill)
	c.	*caro*	/ˈkaɾo/	‘expensive’	(word-medial intervocalic tap)

Romanian has been reported to mainly have the tap, with the trill occurring as a sporadic free variation allophone occurring mostly word-initially (2; Chitoran, 2002; Savu, 2012; Radu, 2016); crucially, the trill-tap substitution word-initially does not cause any change in meaning. On the other hand, French has a single rhotic, the uvular fricative /ʁ/ (3a-b) (Léon,

1992; O'Shaughnessy, 1982), which is phonetically and articulatorily highly dissimilar to both the tap and the trill.

(2)		*rege*	/ˈɾedʒe/	'king'	[ˈɾe.ɖʒe]	[ˈre.ɖʒe]
(3)	a.	*rouge*	/ʁuʒ/	'red'		
	b.	*vérité*	/veʁite/	'truth'		

While L1 Romanian speakers must recategorize their L1 tap and trill allophones as L2 phonemes in Spanish, the L1 French speakers must acquire the distribution of the tap and trill phonemes, two sounds that do not have direct L1 counterparts. Given the differences between the three languages, the L1 Romanian- and L1 French-L2 Spanish pairings can help provide insight into the aforementioned two learning scenarios in L2 phonology.

The majority of previous research has investigated the two above-mentioned scenarios separately, using different tasks, making it difficult to determine which scenario (if either) may be more challenging for learners that are acquiring the same L2. There is ample research regarding the first learning scenario (e.g., for the L2 English lateral-rhotic contrast: Riney & Flege, 1998 and Sheldon & Strange, 1982 for L1 Japanese learners; Borden, Gerber, & Milsark, 1983 and Yu & Jamieson, 1993 for Korean-speaking learners; Eckman & Iverson, 2013, for native Japanese and Korean learners acquiring English /s/-/ʃ/; Vokic, 2010 for L1 English speakers acquiring Spanish /ɾ/). Findings have shown that it is generally difficult to recategorize already existing L1 allophones as L2 phonemes, leading to both perception and production difficulties. Though some studies (e.g., Derrick & Gick, 2005 for English /ɹ/-/l/ by L1 Mandarin speakers; Flege, Munro, & Mackay, 1996 for

English /θ/ and /ð/ by native Italians speakers; Park & de Jong, 2008 for English /p b t d f v θ ð/ by native Korean speakers) have focused on the second learning problem, – acquiring L2 contrasts with no direct L1 counterparts – this is not as well-researched as the first. These limited studies have suggested that, if learners perceive enough of a difference between the L1 and L2 sounds, they may be more successful in perception, but the consequences for production are less clear. Despite this body of research, it is uncertain how learners of different L1s acquire the same L2 phonemes when each group faces a different learning challenge. The present study aims to address this gap in our understanding of L2 speech learning by incorporating both L1 Romanian and French speakers, since they each face a different learning problem when acquiring the same L2 Spanish rhotics.

A secondary objective of this dissertation is to provide a more comprehensive description of rhotics in Romanian. This was accomplished by collecting data from the L1 Romanian participants in their native language in addition to their L2 Spanish. Since previous research on Romanian is limited, it is difficult to understand how it compares with other languages with respect to the distribution and phonetic characterization of rhotics, as well as to make predictions regarding native Romanian speakers' L2 speech patterns. Claims about Romanian rhotics have been mostly observational and brief in nature (e.g., Chitoran, 2002; Savu, 2012), lacking any data collection and acoustic analyses. To my knowledge, only one preliminary acoustic study has been conducted on Romanian rhotics (Radu, 2016). This preliminary research revealed that taps are the most common (72%), followed by trills (24%), as well as approximants and fricatives (4%). Importantly, Radu (2016) reported that the trill occurred word-initially more than word-medially (39% versus

8%). Given this limited research, the present dissertation aims to provide a fine-grained analysis of the distribution and acoustic characteristics of Romanian rhotics in two linguistic contexts, and in turn, to offer empirical support for previous observational claims. This can serve as a point of departure for future studies that may either choose to expand on the topic of Romanian rhotics, – for example, investigating rhotics in phonetic contexts other than those presented here (e.g., in clusters) – or to study it as part of a new language pairing in L2 phonology, such as L1 Romanian-L2 French/Italian.

1.2 The present study: Overview of research questions, methodology and general findings

Three research questions guided this dissertation:

1: What differences are expected between the L1 Romanian and L1 French learner groups regarding the L2 production of the Spanish rhotics? If there are differences, are these related to the speakers' L1 rhotics?

2: Will learners realize trills in word-initial and word-medial intervocalic positions with equal accuracy?

3: Are there differences in the learners' rhotic accuracy in more versus less formal speech, specifically between a sentence reading and a picture description task?

In order to explore these research questions, 10 L1 Romanian, 10 L1 French speakers of intermediate Spanish proficiency, and 13 native Spanish speakers participated in a study. So as to collect data on rhotics in more naturalistic speech (a need highlighted in Colantoni & Steele, 2008), participants completed two tasks: a picture description (more

spontaneous, less formal) and a carrier sentence reading (less spontaneous, more formal). The stimuli contained initial and medial rhotics, similar to the examples in (1a-c), in order to address the second research question regarding positional effects. The stimuli included real and nonce words, the latter in an effort to reduce the effect that L1-L2 cognates may have on production. In addition, as mentioned in Section 1.1, the L1 Romanian participants completed similar tasks in their native language. These data were analysed acoustically in order to better address the first research question, as well as to expand the empirical research on Romanian rhotics.

With respect to the first research question, from an articulatory perspective, while Spanish taps are not particularly difficult to produce (Martínez Celdrán, 1997), several studies using mainly reading tasks have found that trills are difficult for learners (e.g., for L1 English speakers: Amengual, 2016; Face, 2006; Johnson, 2008; Olsen, 2012, 2016; Waltmunson, 2005; for L1 Farsi speakers: Rafat, 2008). Successful trill production requires critical positioning of the articulators, as well as meeting various aerodynamic constraints causing the tongue tip to vibrate, and sustaining adequate airflow (Ladefoged & Maddieson, 1996; Martínez Celdrán, 1997; Navarro Tomás, 1971; Solé, 2002). Due to the difficult nature of trills, previous research has found that learners often substitute the less difficult tap (Face, 2006; Johnson, 2008) or they transfer L1 sounds (Major, 1986). In this project, though from an articulatory point of view, the trill should not pose difficulties for the L1 Romanian speakers since it exists in their L1, these learners' task is to learn a new distribution of the tap and trill. Regarding the L1 French group, in addition to learning the tap-trill distribution, these speakers also have the challenge of learning new articulatory

routines involved in trill production, which are different from the articulatory routines involved in producing their L1 rhotic.

The second research question referred to whether participants would have different trill rates word-initially or medially. Positional/contextual effects for trill production exist, with intervocalic position favouring trill production more than word-initial position due to the airflow from the preceding vowel (Johnson, 2008; Lewis, 2004; Major, 1986). Similar positional effects have also been found for non-rhotic segments (e.g., Rogers & Alvord, 2014 for the Spanish spirants), with overall higher accuracy rates word-medially than initially. Thus, we might expect the L1 French group to have higher trilling rates medially than initially, due to ease of articulation in this position. Regarding the L1 Romanian group, however, recall that, in Romanian, the trill occurs predominantly word-initially (Chitoran, 2002; Radu, 2016), which may lead to this group exhibiting higher trill rates word-initially than medially as a result of transfer from their L1 (Major, 1986).

Lastly, the third research question addressed task effects. Previous research (e.g., Beebe, 1987; Colantoni & Steele, 2007; Díaz-Campos, 2006; Labov, 1972; Lindblom, 1989; Major, 1987, 2001; Rafat, 2010; Tarone, 1979, 1982, 1983; Zampini, 1994) has shown that speech style is an important variable in L2 production accuracy. Though some studies have found less target-like production in more formal tasks for certain segments (e.g., Beebe, 1987; Díaz-Campos 2006; Zampini, 1994), other research has suggested that L1-based transfer generally increases in more casual speech, but decreases in more formal styles, such as word lists (Major, 1987; Tarone, 1979, 1982, 1983). Given these differences

in speech style, we might expect both learner groups in the present study to exhibit more target-like rhotics in the reading than in the picture description task.

An additional factor to consider is orthography, since reading tasks often include visual presentation of the stimuli, while picture tasks do not. Orthography has been found to have both facilitative (e.g., Steele, 2005; Erdener & Burnham, 2005) and hindering effects (e.g., Best & Tyler, 2007; Face & Menke, 2009; Vokic, 2010; Young-Scholten & Langer, 2015; Zampini, 1994) on L2 production. The latter is especially true in cases where the L1 and the L2 use the same grapheme to represent two different sounds. As concerns the present study, word-initially, all three languages represent the rhotics using <r>. However, in word-medial position, the Spanish tap-trill contrast is represented orthographically (<r> for the tap and <rr> for the trill), which may improve rhotic accuracy word-medially (Waltmunson, 2005), as opposed to word-initially, where only <r> (corresponding to the trill) occurs, and learners may map this sound to their L1 rhotics (Young-Scholten & Langer, 2015).

For the L1 Romanian rhotics, the results from the picture description and sentence reading tasks showed that the primary rhotic in Romanian is the tap, and, that the trill occurs predominantly word-initially. Regarding the L2 Spanish rhotics, while both learner groups had relatively high tap accuracy rates, their overall trill accuracy rates were low. Though the L1 Romanian group had slightly higher accuracy rates than their L1 French counterparts, these differences were not significant when word position was taken into account. In contexts that required trills, non-target-like productions included a variety of rhotics, such as the tap and uvular fricative, for the L1 French group, versus mainly the tap for the L1 Romanian speakers. With respect to the Romanian group, the participants who produced

more trills in their L1 also produced more trills in their L2. Furthermore, both learner groups produced significantly more target-like trills medially than initially, suggesting that ease of articulation is an important factor. Lastly, the results showed that, all speakers had higher accuracy rates in the reading than in the picture description task, raising the question of the extent to which previous research on L2 Spanish rhotics may have overestimated learners' accuracy rates, since, to date, predominantly reading tasks have been used in these studies. Overall, given the results, this study suggests that trill production may be difficult for the two learner groups for different reasons. In particular, for the L1 French group, the difficulty is mainly due to articulatory limitations, but also in part due to the distributions of the two rhotics in Spanish, while for the L1 Romanian group it is primarily due to the differences in the distributions between their L1 and their L2.

1.3 Overall contributions

The present study makes several important contributions to previous research. Until now, the focus of L2 Spanish rhotic research has been overwhelmingly on L1 English populations, making it unclear whether the findings could be applied to other L1 speakers. This study expands the repertoire of L1s by investigating the L2 production of Spanish taps and trills by two groups of speakers facing different learning challenges. In doing so, it contributes to the field of L2 phonology by comparing two learning problems, namely (1) acquiring two (quasi)phonemes when both sounds exist in the L1, but as allophones instead of (quasi)phonemes; (2) acquiring the same two (quasi)phonemes when neither sound exists in the L1, within the same study, and using the same tasks in order to better compare the scenarios.

A secondary contribution is a comprehensive description of rhotics in Romanian using acoustic analyses to support previous descriptions that have been predominantly observational or only partially related to descriptions of rhotics (e.g., Savu, 2012 for a focus on the epenthetic vowels flanking the Romanian rhotic). The contribution of this empirical research is two-fold. First, it served as a basis for the predictions concerning the L1 Romanian learners' production of Spanish rhotics in the present study; second, it offers a deeper, more fine-grained understanding of rhotics in Romanian, ultimately expanding the repertoire of less well-studied languages and providing a foundation for research moving forward, such as the analysis of Romanian rhotics in other phonetic contexts not studied here (e.g., clusters, coda position).

Lastly, this study also contributes methodologically to the literature on L2 Spanish rhotics. As mentioned, the overwhelming majority of previous research has used mainly reading tasks, leading to potentially incomplete conclusions regarding production accuracy across tasks differing in the degree of formality. To address this, in addition to a reading task, the present study also employed a more spontaneous task (i.e., picture description). As such, the present study makes the contribution of more naturalistic speech in the L2 Spanish rhotic field.

1.4 Thesis structure

The following is an outline of the thesis. Chapter 2 provides a summary and critical evaluation of previous research concerning L2 phonology more generally, and L2 Spanish rhotic acquisition more specifically. Detailed articulatory and distributional descriptions of

the rhotics in the three languages (Spanish, Romanian and French) under investigation are also provided. Chapter 3 begins by reiterating the research questions guiding this study as well as the detailed corresponding hypotheses based on the previous literature presented in Chapter 2. Chapter 3 also outlines the methodology, from descriptions of the participants, to the tasks and stimuli. Next, Chapter 4 reports the results, beginning with the L1 Romanian data, followed by the Spanish rhotic data for all three speaker groups: L1 Romanian, L1 French and native Spanish controls. Lastly, Chapter 5 evaluates the hypotheses, discussing the findings in light of previous research, as well as implications for future research.

Chapter 2

Literature Review

2 Overview

In order to inform the predictions that will follow in Chapter 3, the goal of the present chapter is to review the literature regarding the two learning problems that the learners in this study face, and to discuss the effect of several factors (articulatory difficulty, position in the word, task formality, and orthography) on L2 production. Next, this chapter provides a summary of the rhotics in Romanian, French and Spanish, as well as the research that has been done on L2 Spanish rhotics specifically.

As mentioned in Chapter 1, this thesis is concerned with two learning problems: (1) acquiring the distribution of sounds that are already present in the L1 and in the target language, but are allophones in the L1, but phonemes in the L2; (2) acquiring the distribution of two contrastive L2 sounds when these do not have direct L1 counterparts. As will become clear in Sections 2.1.1 and 2.1.2, most studies have looked at one or the other learning problem individually, but not with the goal of comparing them when learners are acquiring the same L2. Generally, previous research has found that both learning scenarios are difficult for learners.

The next part of the discussion then shifts to the literature on each of the independent variables involved in this study. Specifically, this study seeks to determine the effect of position within the word, task formality, and, to a lesser extent, orthography, on the L2

production of Spanish rhotics. Section 2.2.1 summarizes general articulatory difficulties involved in producing trills, which are articulatorily difficult segments, more so than taps, and are problematic for both children acquiring Spanish as their first language (e.g., Goldstein, 2000; Jiménez, 1987), as well as for L2 learners of Spanish (Blecua, 2001; Quilis, 1981; Solé, 2002). Lindblom's (1989) H&H Theory can be related to trill production, since the theory accounts for differences in production according to different variables. As concerns rhotics, trills have been reported to be longer than normal under certain conditions (e.g., differences in task formality) due to hyperarticulation (Rafat, 2008).

In addition to trills being difficult due to aerodynamic constraints, part of their difficulty is also attributed to the adjacent segments, namely the degree of stricture associated with these, resulting in higher trill rates word-medially intervocalically than absolute word-initially for native speakers (Lewis, 2004; Solé, 2002), and L2 learners (Major, 1986). The tendency for L2 sounds to be first acquired word-medially rather than initially has also been observed for other segments, such as the French uvular rhotic (Colantoni & Steele, 2007) and the Spanish spirants (Rogers & Alvord, 2014). These positional effects are discussed in Section 2.2.2.

Task formality (Section 2.2.3) is also an important consideration in this study, since the sentence reading task is arguably more formal than the picture description task. There is some evidence in previous research to suggest, on the one hand, that learners produce more errors in less formal tasks due to L1-based transfer (Major, 1987), and on the other hand, that learners are more target-like in more spontaneous tasks, at least with some segments (Beebe, 1987; Díaz-Campos, 2006; Zampini, 1994).

Lastly, the role of orthography is discussed in Section 2.2.4, since this study involves a task in which orthography is directly involved (i.e., a sentence reading task) and one in which orthography is not present (i.e., a picture description task). As we will see, orthography has been formalized in only one (perceptual) model (Best & Tyler, 2007), and its effects have been found to be both facilitative (e.g., Erdener & Burnham, 2005; Steele, 2005) and hindering (e.g., Bassetti & Atkinson, 2015; Face & Menke, 2009; Vokic, 2010; Young-Scholten & Langer, 2015; Zampini, 1994) in L2 speech.

Section 2.3 is a detailed summary of each of the three languages involved in this study, focusing on the phonological, phonetic, and articulatory characteristics of these sounds. Lastly, Section 2.4 provides a critical discussion of the research that has been conducted on the L2 acquisition of Spanish rhotics specifically, in order to summarize the factors that have been found to affect the production of these segments. The literature review will show that the trill is usually acquired after the tap (Colantoni & Steele, 2008; Face, 2006; Olsen, 2012; Waltmunson, 2005) and the position within the word, as well as phonetic context, are important factors (Colantoni & Steele, 2008; Face, 2006; Johnson, 2008; Olsen, 2012; Waltmunson, 2005). Moreover, it has been shown that learners do not target all of the parameters of rhotics simultaneously. That is, they do not master length, percentage of voicing and manner all at the same time (Amengual, 2016; Colantoni & Steele, 2008; Johnson, 2008; Waltmunson, 2005). Additionally, as will become clear, the majority of the studies have focused on English as the L1 and have predominantly used reading tasks to elicit rhotics.

2.1 The learning problems

2.1.1 L1 allophones to L2 phonemes

Research regarding the recategorization of an L1 allophonic distribution into an L2 phonemic distribution is abundant, mainly focusing on perception, but also production, at least to some extent, which is important for the present study. A large body of this research has focused on the acquisition of the English /l/-/ɹ/ contrast by native speakers of East Asian languages (for L1 Japanese: Riney & Flege, 1998; Sheldon & Strange, 1982; for L1 Korean: Borden et al., 1983), which lack this phonemic contrast.

In their study on L1 Japanese speakers learning English, Sheldon and Strange (1982), showed accurate production of the /l/-/ɹ/ contrast, but not perception. Japanese has only one liquid /r/, phonetically described as a tap [ɾ]. Crucially, in Japanese, [l] is an allophone of /r/, albeit rarely occurring. Japanese speakers' learning problem, then, is two-fold: first, they must perceive the distinction between English /ɹ/ and /l/; second, they must learn that, except for the infrequent [l], the phonetic variants of the Japanese liquid are unlike the English /ɹ/, an alveolar approximant. Sheldon and Strange analysed the production and perception of the English /ɹ/-/l/ contrast by six L1 Japanese-L2 English speakers, whose production was rated as good (five speakers) or poor (one speaker) by four native (American) English speakers. First, all participants produced sixteen minimal pairs containing singleton liquids in all possible positions. Then, they listened to the same minimal pairs from the production experiment, and identified the minimal pairs as containing either /ɹ/ or /l/. Different versions of the perception experiment were administered to the participants, depending on which

group produced the stimuli: (i) native English speakers; (ii) other Japanese speakers, or (iii) the Japanese speaker who was being tested. Based on the production and perception results, the authors categorized the participants as follows: the 'poor' L1 Japanese-L2 English speaker made errors in both perception and production; two of the 'good' L1 Japanese-L2 English speakers had minimal errors in production, but had several errors in perception; the other three 'good' L1 Japanese-L2 English speakers were accurate in both production and perception. An important finding reported by Sheldon and Strange was that the perception of the contrast varied depending on the context in which it occurred. Specifically, /l/ was misidentified more than /ɹ/ word-initially, whereas /ɹ/ was misidentified more than /l/ in consonant clusters (i.e., *lead* was misidentified as *read*, while *breed* was misidentified as *bleed*). The authors also mentioned that the participants in their studies had reported that they were taught to produce de /l/-/ɹ/ contrast from an articulatory perspective, rather than an auditory one, possibly leading to high accuracy rates in production. These findings highlight the importance of considering position within the word in L2 speech, as well as the idea that perception does not always precede production in L2 speech.

Like Sheldon and Strange (1982), Riney and Flege (1998) also investigated the acquisition of the English /ɹ/-/l/ contrast by native Japanese speakers, but the authors only focused on production. Specifically, they examined changes in foreign accent over time by analysing the production of 11 L1 Japanese-L2 English across three experiments conducted 42 months apart, at two different times. In Experiment 1, which investigated global foreign accent, participants were recorded reading five English sentences at two different times (T1 – freshman year and T2 – senior year). Then, five native English speaker judges rated each

sentence on a scale of 1 ‘strong foreign accent’ to 9 ‘no foreign accent’. In the second experiment, the goal was to determine if the L1 Japaneses-L2 English speakers’ /ɹ/s and /l/s at T2 were more identifiable than at T1. Participants read 25 words containing singleton and cluster liquids at T1 and T2. Three trained native English listeners heard only the first part of the words through headphones and, by checking a marked box, identified what they had heard as either “R”, “L” or “neither”. The authors note that, there is a possibility that a speaker’s productions were equally identifiable at both T1 and T2, but that their production at T2 may have been more accurate than at T1. As such, in Experiment 3, the aim was to implement a more fine-grained perceptual technique that would help determine whether any improvement had occurred over time. In Experiment 3, ten untrained listeners (five were the listeners in Experiment 1) listened to the matching T1-T2 word pairs from Experiment 2. Their task was to choose the word in the pair that exhibited the better exemplar of /ɹ/ or /l/. They were told that the words were supposed to contain /ɹ/ and /l/, but that in some cases, they may be unable to make a clear decision; in such cases, they were told to guess. In Experiment 1, three of the 11 L1 Japanese-L2 English speakers’ global accents significantly improved over time, two speakers showed marginal improvement, the and six showed no changes. In Experiments 2 and 3, generally, there were not significant improvements between T1 and T2 when comparing /ɹ/ and /l/ productions[1]. The findings suggest that there is individual variability in the L2 production of the rhotic-lateral contrast, as some speakers made improvements both globally and at the liquid identifiability and accuracy levels, while others exhibited no improvements across the three experiments. Furthermore, although /l/

[1] Significant improvement between T1 and T2 occurred for /l/ in clusters, but not for either of the liquids in any of the other environments (i.e., /l/ or /ɹ/ singleton and /ɹ/ clusters).

in both singleton and clusters tended to be better than /ɹ/, against the authors' prediction, there was no significant improvement between T1 and T2 for either segment. Since the mean score for singleton /l/ identification was higher than singleton /ɹ/ at T1, Riney and Flege noted that much of the improvement had perhaps already taken place before T1. Overall, the results of this study suggest that, although there is individual variability, the English /ɹ/-/l/ contrast is difficult to acquire by speakers whose L1 lacks this contrast, and there does not seem to be significant improvement between the early stages (junior year or T1) versus the later stages of acquisition (i.e., senior year or T2).

The learning problem is similar for Korean speakers, who report having difficulty discriminating between /ɹ/ and /l/ as separate phonemes in English, since, in Korean, both sounds are grouped into one category (Jung, 1962). /r/ occurs intervocalically and it is phonetically realized as a tap [ɾ], articulated with the tip of the tongue against the alveolar ridge (Pyun, 1987); [l] occurs post-vocalically in final position, and is a light [l] (Borden et al., 1983). Neither allophone occurs in initial position (Jung, 1962), and there is no orthographic difference between the two sounds (Borden et al., 1983). This distributional difference has resulted in substantial research on the acquisition of this contrast by native Korean speakers.

Similar to Sheldon and Strange (1982), Borden et al. (1983) also investigated the perception-production link. Ten L1 Korean-L2 English speakers were trained on the rhotic-lateral contrast. Training involved verbal feedback on an identification test (nonsense syllables and words), a discrimination test (correct versus incorrect productions), practice producing the contrast, and practice in self-judgments. Then, they completed a production

and three perception experiments (identification, AX discrimination, and self-perception, which involved participants listening to their own productions and judging whether [r] and [l] were pronounced correctly). In the production experiment, participants imitated a tape recording of a native English speaker uttering 20 randomized samples of words containing /ɹ/ and /l/ (e.g., *road-load*), 20 nonsense CV syllables containing the contrast (e.g., *ra-la*), ten sentences with /ɹ/ and /l/ in similar environments (e.g., *Put the key in that lock. Don't throw that rock*), and lastly, 12 items including syllables, words and then complete sentences (e.g., *Say the sound 'li'* → *Say the word 'light'* → *Turn on that light*). Participants then completed three perception tests: an identification test, an AX discrimination test – the stimuli for both was a continuum of ten synthetic syllables from [ra] to [la] – and a self-perception test using the utterances recorded during the imitative production task. Results showed that participants were better in the discrimination than in the production task, supporting the notion that perception precedes production, a finding which contradicts Sheldon and Strange's (1982) observed tendency, whereby production (sometimes) leads perception. However, despite the extensive training, participants' production only improved short-term, with little to no carryover beyond the training session. There was some individual variation with regard to types of errors, with some speakers replacing the rhotic with the lateral, some producing a rhotic-like sound instead of the lateral, and some speakers with substitutions both ways. These findings suggest that the English liquid contrast is difficult to produce, and to some extent perceive, and learners have varying error patterns. Though Borden et al.'s (1983) findings suggest that training may not yield long-term effects, especially in production, Yu and Jamieson (1993) found that training does indeed have lasting effects, at least in what concerns perception.

In a study following procedures similar to Logan, Lively and Pisoni (1991), Yu and Jamieson (1993) trained five L1 Korean-L2 English speakers on the English /ɹ/ and /l/ contrast. For the pre- and post-tests, participants listened to minimal pairs containing /ɹ/ and /l/ in five phonetic environments, and had to indicate whether they heard /ɹ/ or /l/ by pressing "L" or "R" on a keyboard. Participants received 15-minute training sessions over 15 days; training involved a two-alternative forced-choice identification task with feedback, and participants were only trained on two of the five phonetic environments (i.e., initial singletons and initial consonant clusters, since the authors predicted these would be most difficult). Regarding position, results showed that the most difficult was final consonant cluster, while the least difficult was final singleton. Equally difficult were initial singleton, initial consonant cluster, and medial position. Overall, accuracy scores significantly increased for all participants from pre- to post-training, showing that adult Korean speakers can more accurately perceive the English /ɹ/-/l/ contrast post-training. However, it is unclear how Yu & Jamieson's findings would apply to production.

Lastly, Vokic (2010) also tested the role of L1 sound status in L2 speech production. Specifically, the main hypothesis was that the phonological status of the sound (i.e., allophone versus phoneme) and its distributional characteristics in the L1 and L2 are the main predictors of success in L2 production. Based on previous findings (e.g., Boomershine, Hall, Hume, & Johnson, 2004; Flege, 1995; Major, 1986; Zampini, 1994), it was predicted that, if an acoustically similar or identical L2 sound is a phoneme in the L1 and an allophone in the L2, and the two sounds share the same distributional properties, then the sound would be relatively easily accessed in the L2. Likewise, if an acoustically similar or identical L2

sound is an allophone in the L1 and a phoneme in the L2, and the two sounds share the same distributional properties, the sound would be relatively easy. Conversely, if the similar or identical target sound is a phoneme in the L1 and an allophone in the L2, but the two sounds do not have overlapping distributions, the sound will be problematic. The most difficult scenario would be one in which the similar or identical target sound is an allophone in the L1 and a phoneme in the L2, but the sound occurs in a position in which it never occurs in the L1. To test these predictions, Vokic (2010) analysed 12 intermediate to high-intermediate L1 English-L2 Spanish speakers' word list productions of Spanish [ð] and [ɾ]. While both sounds exist in both Spanish and English, they have different phonological statuses and distributions in the two languages; /ð/ is a phoneme in English (e.g., *this* [ðɪs]), but an allophone of the voiced stop /d/ intervocalically in Spanish (*dedo* [de.ðo] 'finger'). Conversely, intervocalic /ɾ/ is a phoneme in Spanish (e.g., *caro* [ka.ɾo] 'expensive' and *carro* [ka.ro] 'car'), but an allophone of /t/ and /d/ in North American English (e.g., *pity* [pi.ɾi]). Overall, L1 English-L2 Spanish speakers were slightly more successful in accessing L1 phonemes (i.e., /ð/) than L1 allophones (i.e., /ɾ/) when producing the L2 sounds, but this difference was not statistically significant. This tendency suggests that the phonological status of a sound (i.e., allophone versus phoneme) in the L1 may not greatly affect the production of the sound in the L2. However, some research (e.g., Eckman & Iverson, 2013) shows that the phonological status of a sound in an L1 does indeed have different consequences for the production of the L2 contrast.

Eckman and Iverson (2013) investigated the production of the English /s/-/ʃ/ (e.g., *sip* versus *ship*) contrast by 26 native Korean and 23 Japanese speakers. Japanese and

Korean both have the sounds [s] and [ʃ], and in both languages, /s/ is realized as [ʃ] before the high front vocoids [i] and [j]; however, while in Korean [s] and [ʃ] are in complementary distribution (as allophones of /s/), in Japanese the two sounds are (arguably) different phonemes (e.g., *sakai* [sakai] 'boundary' versus *shakai* [ʃakai] 'society'). However, this contrast is merged or neutralized before high front vocoids (e.g., *[siki] versus *shiki* [ʃiki] 'ceremony'). As such, the same rule (i.e., /s/ → [ʃ] before high front vocoids) is neutralizing/phonemic in Japanese, but allophonic in Korean. Broadly, the authors argued that the phonological difference between the two languages has consequences for the way in which the L2 English contrast is acquired, as well as for the type of production errors. The 49 participants completed a picture naming task, in which they produced 60 English words containing /s/ and /ʃ/ in various positions (word-initial before a high front vowel, medial position, final position and medial position at the juncture with another morpheme). Results showed opposite patterns for the two groups. The native Korean speakers were more target-like with /ʃ/ in the environment before /i/, while the native Japanese speakers were more target-like with /s/ than /ʃ/ before /i/. The authors attributed the different results to the status of the two sounds in the speakers' L1. Specifically, since the Japanese speakers produced hypercorrection errors, whereby they pronounced words containing /ʃ/ before a high font vowel as [s] instead of [ʃ], the authors argued that this group must have made a distinction between the /ʃ/-/s/ contrast. Conversely, the authors noted that the Korean speakers were successful in their performance on [ʃ] before /i/, but they were not target-like in their production of /s/ because they applied their L1 allophonic rule to English words. That is, they produced their L1 allophone [s] before the high front vowel [i]. This study suggests that it is not the mere existence of two specific sounds in the L1 that affects the L2

production of those sounds, but rather their phonemic status. That is, two sounds that exist in the L1 as allophones will have a different effect on L2 production than two sounds that exist in the L1 as phonemes.

The body of research reviewed above concerns L1 allophones to L2 phonemes, which is relevant for this thesis, since it is the learning problem faced by L1 Romanian speakers acquiring Spanish rhotics. That is, they must learn to recategorize their L1 tap and trill allophones as phonemes in L2 Spanish. To summarize, the studies reviewed in this section show that, while speakers may learn to recategorize L1 allophones as L2 phonemes over time and with training, L2 contrasts that are mapped onto the same native category are generally difficult to acquire, in both perception and production, and there may be minimal long-term improvement. A related question that arises is how the acquisition process would differ in a scenario where the L1 lacks either one or both sounds involved in the L2 contrast. This situation describes the L1 French group's acquisition problem, since, as mentioned, French has a uvular rhotic which is highly dissimilar to both Spanish rhotics. This latter learning problem is discussed in the following section.

2.1.2 The acquisition of L2 contrasts when one or both sounds are absent from the L1

The literature reviewed so far has focused on the acquisition of contrasting L2 sounds when these exist in the L1 as allophones. A different, but related, learning problem is a scenario where there are two contrasting L2 sounds that do not have L1 counterparts. The research on this learning problem is limited, and predominantly focuses on situations where one

sound exists in the L1 and the other does not. One study that focused on this learning problem is Derrick and Gick (2005), who investigated the L2 production of the English /ɹ/-/l/ contrast among two L1 Beijing Mandarin (BM) and two L1 'other' Mandarin (OM) speakers. For both groups, one speaker was experienced, while the other was new. While all Mandarin dialects have a prevocalic liquid realized as a dental-alveolar lateral approximant[2], Beijing Mandarin also has a post-vocalic variant that is similar to the English /ɹ/. The learning problems, then, are slightly different for the two groups: for the BM speakers, learning involves recategorizing an already existing L1 allophone as an L2 phoneme, while for the OM speakers it is one of learning a new sound, /ɹ/, which contrasts with an old sound (i.e., one that exists in the L1), /l/. Derrick and Gick predicted that, the BM group would produce more accurate pre- and post-vocalic /ɹ/ than the OM speakers. The stimuli, embedded in the carrier phrase *I say* _____ *again*, included minimal pairs of simple and complex onsets (e.g., *lead-read* and *glass-grass*); intervocalic position (e.g., *collect-correct*); and simple and complex codas (e.g., *deal*-dear and *cold-chord*). The stimuli produced were then presented to 13 monolingual native English speakers, who had to indicate whether the productions of the two English liquids were correct or incorrect. The results were compiled and presented according to the perception differences between: BM and OM speakers; /ɹ/ and /l/ sounds; and various syllabic contexts (i.e., onset, coda, intervocalic). The results showed that the experienced OM speaker was easier to understand than the experienced BM speaker, while the new BM speaker was easier to understand than

[2] This is similar to Cantonese, which has only /l/, but not any rhotic consonant (Chan & Li, 2000). The lack of a lateral-rhotic contrast often poses problems for native Cantonese speakers learning English, with these speakers often producing [w] and [l] instead of [ɹ] (Hung, 2000), or even deleting the segment entirely (Stibbard, 2004).

the new OM speaker for both liquids. Regarding syllabic context, all speakers produced onsets more accurately than codas, while the BM speakers produced the distinction between /ɹ/ and /l/ more accurately across all contexts. Overall, the results showed that the presence of the /ɹ/-like allophone in Beijing Mandarin significantly improves new learners' ability to produce the distinction between /ɹ/ and /l/. The findings suggest that learners may be better at producing L2 contrasts when they have to recategorize already existing L1 variants versus when they have to acquire an L2 sound that does not have an L1 counterpart. However, since only four speakers participated in the study, it is difficult to generalize these results to other populations.

The L2 acquisition of contrasts was also investigated by Flege, Munro and MacKay (1996). Their study differed from Derrick and Gick (2005) in that it analysed the production of an L2 contrast when neither sound exists in the L1. Specifically, Flege et al. analysed the L2 production of English word-initial /θ/ and /ð/[3] by L1 Italian speakers. Since these sounds do not exist in Italian, L1 Italian speakers should treat both interdentals as new sounds. 240 L1 Italian-L2 English speakers between the ages of 3 to 21 along with 24 native English controls first completed a detailed language background questionnaire, which included questions regarding their age of arrival in Canada, length of residence in Canada, % use of English, and % use of Italian. Participants then completed a production task in which they were presented with a list of words containing word-initial /θ/ and /ð/ (e.g., *thief*, *then*) and they simultaneously listened to each stimulus embedded in the phrase ____ *is the next word*.

[3] The authors also included /p/ and /t/ in their study, since these sounds have different VOTs in English and Italian.

Their task was to repeat the stimulus embedded in the carrier phrase *Now I say* ____. The data analysis involved a set of ten native English judges, who typed a number ranging from 1 to 6 on a keyboard depending on whether they heard: a "correct *th*"; a "distorted *th*"; "*s*"; "*f*"; "*t*" or "*d*". Results showed that, overall, while participants who learned English as young children had native-like productions of the two fricatives, those who had started learning English after the age of 11 had fewer target-like production of the English interdental fricatives, most of often realizing /θ/ as [t] and /ð/ as [d]. Flege et al.'s results revealed that, the L1 Italian speakers did not treat English /ð-θ/ as a contrast, and instead treated them as allophones of their L1 Italian /t-d/ contrast. The authors noted that, despite the fact that the age of L2 learning was the most important factor, there was a considerable amount of inter-speaker variability, which was affected to some degree by factors identified using a detailed language background questionnaire (factors such as motivation and language use proved to be weak predictors of target-like productions). However, since not all of the variance was accounted for by these factors, the authors offered a perceptual explanation for their results. Specifically, they noted that, if an L2 sound is to be produced accurately, it must be recognized as being phonetically distinct from any L1 sound. For example, in the case of the L1 Italian speakers in their study, Flege et al. mentioned that the L1 Italian speakers may have produced the voiced interdental fricative as /d/ because they did not perceive the phonetic distinction between English /ð/ and Italian /d/ (the closest Italian sound). These findings suggest that, regardless of the importance of the age at which a second language is learned, the inability to perceive an L2 contrast may also result in production errors. Further, the results reveal that learners assimilate L2 contrasts based on their L1 categories.

To explore the extent to which L1 categories are used in the acquisition of L2 contrasts, Park and de Jong (2008) tested 40 L1 Korean-L2 English listeners' identification of the English obstruents /p b t d f v θ ð/. The general hypotheses of this study were the following: (1) based on the Perceptual Assimilation Model (PAM) (Best & Tyler, 2007), it was predicted that if the segments of the L2 contrast are mapped onto the same L1 category, it would be difficult for learners, while L2 contrasting segments that are mapped onto different L1 categories would be easier; (2) based on the SLM, two contrastive segments that do not map onto any L1 category would not pose identification problems if the learners have acquired the L2 segmental label. As for the sounds investigated in the study, stops and sibilant fricatives are similar in Korean and English with regard to place and manner of articulation (stops in Korean, however, show a three-way tense-lax-aspirated contrast in initial position). However, the non-sibilant fricatives (/f v θ ð/) were new to the learners, since Korean only has one non-sibilant fricative (/h/). To address the hypotheses, participants listened to English nonwords and identified each target sound using both Korean (Korean labeling task) and English (Roman/IPA labeling task) consonants. For the Korean labels, participants were asked to label on a scale from 1 to 7 how good they thought their label was (i.e., goodness-of-fit judgements). The results from the Korean labeling task showed that, while the participants correctly identified /p/ and /t/ at a rate of 90% (i.e., high goodness ratings), their identification accuracy for /f v θ ð/ was under 60%. Specifically, /v/ was perceived as Korean /p/ and occasionally /t/, while /f/ was perceived as more than two Korean categories, namely /p^h/ and tense /p/, as well as /p/ in very few cases. On the other hand, regarding the interdentals, /ð/ was most often perceived as /t/ and, to a lower extent /p/, while /θ/ was perceived as tense /s/ most often, as well as a variety of other sounds such

as tense /t/, /p^h/, tense /p/, and /t/. These results show that, while some English sounds have similar Korean counterparts, the fricatives /f v θ ð/ are considered "new" segments and are therefore not confidently mapped onto L1 categories, leading the participants to be relatively good at identifying the English fricatives (in the Roman/IPA labeling task). The authors interpreted this finding as evidence that the learners were building new categories, effectively separating the fricatives from the stops. Though these results indicate that, learners can successfully identify L2 sounds that do not have L1 counterparts if they perceive the sounds as new and do not map them onto already existing L1 categories, it is unclear if these findings would translate to production accuracy as well.

The research discussed in this section focused on the acquisition of L2 contrasts when the L1 has either only one or neither of the sounds involved in the contrast. As mentioned, this is the learning problem for the L1 French group in the present study, since they must acquire the L2 Spanish tap-trill contrast and their L1 lacks both segments. To summarize, studies on the acquisition of L2 contrasts involving sounds that do not have L1 counterparts have shown that, if learners perceive enough of a difference between L1 and L2 sounds, they will develop new categories instead of mapping L2 contrasts onto already existing L1 categories, leading to successful perception (Park & de Jong, 2008), and sometimes production (Flege et al., 1996), though this remains inconclusive. Additionally, research suggests that learners who have to recategorize already existing sounds from L1 allophones to L2 phonemes have less difficulty producing the L2 contrasts than those who have to learn a new sound (Derrick & Gick, 2005); however, it is uncertain if this finding would be replicated in a study with more participants. Having discussed the literature

concerning the acquisition of L2 contrasts when the L1 lacks one or all of the segments involved, the following section turns to factors that affect segmental production.

2.2 Articulatory nature of trills, position within the word, effects of task formality and orthography on L2 phonological acquisition

2.2.1 Articulatory nature of rhotics and hyperarticulation

Across languages, from an articulatory point of view, liquids are difficult segments and are acquired by children later compared to other segments (van Lieshout, Merrick, & Goldstein, 2008). Beyond the early years, teenagers and adults may continue to display problems articulating rhotics in their L1 (e.g., Adler-Bock, Bernhardt, Gick, & Bacsfalvi, 2007 for English). Given the difficult articulatory nature of rhotics, including for native speakers, it is not surprising that L2 speakers face challenges when producing these sounds, especially if the articulatory patterns in the L1 and L2 differ.

Trills are especially difficult to articulate. The challenging nature of trills is a result of several factors, such as the critical positioning of the articulators and various aerodynamic constraints (Solé, 2002). Ladefoged and Maddieson (1996) discuss the difficulty of articulating trills generally (i.e., not Spanish trills specifically), noting that:

> One of the soft moveable parts of the vocal tract is placed close enough to another surface, so that when a current of air of the right strength passes through the aperture created by this configuration, a repeating pattern of closing and opening of the flow channel occurs […]. In its essentials this is very similar to the vibration of the vocal folds during voicing; in both cases there is no muscular action that controls each

> single vibration, but a sufficiently narrow aperture must be created and an adequate airflow through the aperture must occur. The aperture size and airflow must fall within critical limits for trilling to occur, and quite small deviations mean that it will fail (p. 217).

Evidently, trills require not only critical positioning of the articulators, but also adequate airflow. Moreover, the inability to produce trills is usually the result of the blade of the tongue being too stiff (Ladefoged, 1993, p. 169), as well as the fact that the articulation of the sound involves "considerable muscular tension and a large quantity and expiratory speed of the expelled air across and around the tongue apex" (Hammond, 1999, p. 135).

For Spanish, research has shown that, compared to trills, the production of taps is not particularly difficult, as there is more articulatory tension involved in producing trills than taps (Navarro Tomás, 1971). Martínez Celdrán (1997) notes that, to produce a tap, the tip of the tongue raises to touch the alveolar ridge but does not exert pressure (pp. 94-95). Further, for taps "the entire movement is voluntary and carried out in a very relaxed manner" (my translation) (Martínez Celdrán, 1997, pp. 94-95). Whereas taps involve active muscular movements of the tongue (i.e., the direct movement of the tongue tip to make contact with the dental or alveolar ridge), trill production relies more heavily on aerodynamic forces causing the articulators (e.g., tongue tip) to vibrate (Ladefoged & Maddieson, 1996). In addition, for the trill, the tongue body is also more highly constrained than it is for the tap (Solé, 2002). Besides being more difficult than taps, trills are also more difficult than approximants and fricatives, since they require more articulatory precision. As a result, trills are among the most challenging sounds to acquire, by native speakers (e.g., Goldstein, 2000; Jiménez, 1987) and learners (Blecua, 2001; Quilis, 1981; Solé, 2002) alike.

An articulatory theory that could be (and has been) applied to trill production, specifically the duration of these segments, is Lindblom's (1989) H&H Theory, which focuses on adaptive variability. The basic tenets of the theory are that speaking and listening are shaped by more general biological processes, and that a speaker's goal is to minimize articulatory effort as much as possible, while simultaneously ensuring that the listener understands. As such, speakers vary their production based on the communicative needs of listeners and situational demands, yielding productions that vary along a continuum between hypo- (i.e., enunciating less carefully) and hyper-speech (i.e., enunciating more carefully). Hyperarticulation involves pronouncing words more clearly and is associated with phonetic features such as longer durations and greater vowel spaces. Conversely, hypoarticulation involves pronouncing words less clearly and involves characteristics such as shorter duration, smaller vowel spaces, and sometimes even elided phonemes. For the current study, one relevant idea is related to the degree of formality, since this study involves two tasks differing in task formality. Both L1 and L2 speakers hyperarticulate in formal contexts, and hypoarticulate or undershoot in more informal ones. Regarding rhotics, hyperarticulation means that, in cases where learners produce trills, in certain contexts, they may overshoot the target, leading to a higher number of closures and longer durations than trills in either their L1 or than those produced by native Spanish speakers[4]. Blecua (1999) found that native Spanish speakers produce trills with fewer closures in fast, more casual articulation, suggesting that when trills have more closures, it is due to hyperarticulation. Solé's (2002) research corroborates this, as she found that her native Spanish population

[4] Hyperarticulation cannot affect taps, since, for these segments, a longer duration would result in a different segment altogether (e.g., /t/ or /d/).

produced trills with more contacts than previously reported (average of four contacts, but as many as five or six, versus 2-5 contacts reported by Blecua, 1999, and Recasens, 1999), noting that the trills produced in her study may have been the result of hyperarticulation due to laboratory conditions.

For learner populations, though trills are articulatorily difficult, in cases where the speakers do produce trills, these may be hyperarticulated. For example, Rafat (2008) investigated the acquisition of the Spanish trill by two L1 Farsi-L2 Spanish speakers. Based on previous research (e.g., Colantoni & Steele, 2007; Flege, 1995; Lewis, 2004), Rafat predicted that word-initially and post-consonantally the trill will be considered a 'new' sound and will therefore be acquired. Moreover, it was predicted that there would be higher trill rates word-initially than post-consonantally. Regarding intervocalic position, it was predicted that this position would favour trill production more than word-initially and post-consonantally. To test these hypotheses, the two learners read a list of Spanish words and a list of words in Farsi. Overall, Rafat found low Spanish trill production rates, with one participant producing some trills, and the other not producing any. When the trill was produced, the results revealed longer trill productions in the L1 versus the L2 (151 ms versus 175 ms), a result which was interpreted as hyperarticulation, as the L2 trills were even longer than those of native Spanish speakers. This corroborates the findings of Colantoni and Steele (2007), who reported hyperarticulation among L1 English speakers of intermediate French proficiency, noting that this group had overly long rhotics, resulting in devoicing (p. 399). Rafat noted that, although her speakers exhibited low trill rates, there is a possibility that there would be higher trill production rates in more informal tasks where

reading is not required. A similar idea was also mentioned in Colantoni and Steele (2008), who suggested future research investigate the production of Spanish rhotics in more naturalistic speech, where hyperarticulation is absent and production constraints play a greater role. In sum, Lindblom's (1989) H&H Theory can help to explain hyperarticulated productions, especially in more formal tasks, such as reading, and hypoarticulated speech in more spontaneous tasks. This research suggests that both L1 and L2 speakers both hypo and hyperarticulate depending on contextual demands (including task formality), which will be discussed in more detail in Section 2.3.3.

Taken together, the above research shows that, while taps are produced using deliberate movements and are realized in a relaxed manner (Martínez Celdrán, 1997), trills are more difficult because they require the optimal positioning of the articulators and they rely on aerodynamic forces to be produced successfully (Ladefoged & Maddieson, 1996; Martínez Celdrán, 1997; Navarro Tomás, 1971; Solé, 2002). The difficult nature of trills is also evident considering the low trilling rates among L2 speakers (Rafat, 2008), which will be further discussed in Section 2.5. In addition, in cases where speakers do produce trills, they are often overly long when compared to native speaker values (Rafat, 2008). The research also highlights the importance of investigating Spanish rhotic production in more spontaneous tasks, where hyperarticulation is less likely to occur and production constraints have more of an effect (Colantoni & Steele, 2008). The following section reviews the literature on the effect of the position of the segment in the word.

2.2.2 Effect of position in the word

The hypotheses in this study refer to the position within the word, and, as such, the effects of this variable on L2 production must be discussed. Although research has shown that the Spanish tap is generally acquired intervocalically before syllable-finally (Colantoni & Steele, 2008), the positional consideration does not apply to taps in this study, since the stimuli here involve the tap in only one position. Rather, it pertains specifically to the trill, since the stimuli involve this segment in onset position both word-initially and word-medially intervocalically.

The ease with which trills are produced may also be affected by their position within the word. Specifically, intervocalic position tends to favour trill production more than post-consonantal, since the precise requirements can be optimally met when the trill is flanked by vowels, and producing the trill with adjacent consonants could disturb the optimal conditions for trilling (Solé, 2002). Lewis (2004) established a continuum of the difficulty of Spanish trill production across word boundaries, finding an overall positive correlation between an increase in the degree of stricture of the pre-rhotic consonant and successful trill production. Specifically, he investigated trill production among native Spanish speakers in several environments: after /n/ (e.g., *Diga la frase u**n** **r**ío para mí.* 'Say the phrase a river for me'), after /l/ (e.g., *Diga la frace el rabo para mí.* 'Say the phrase a tail for me'), intervocalically (e.g., *Diga la frase mi ruta para mí.* 'Say the phrase a route for me'), post-pause absolute initial position (e.g., single words starting with /r/, such as *roto* 'broken'), and after /s/ (e.g., *Diga la frase las rejas para mí.* 'Say the phrase the bars for me'). Lewis (2004) found that the highest rates of trilling occurred after /n/, followed by after /l/,

followed by intervocalically (the rhotic was always after /i/ or /a/), then post-pause absolute initially. The lowest rates of trilling occurred after /s/. The author concluded that, as the degree of stricture of the pre-trill consonant increases, so too does the likelihood of trill production; that is, having the greatest degree of stricture, /n/ favours trill production, while /s/, having the lowest degree of stricture, disfavors trill production. The post-vocalic and absolute-initial contexts do not follow this pattern, however, since trills were produced at relatively high rates in both these contexts. Lewis (2004) concluded that, since there are no demands placed on the position of the tongue prior to articulating the trill in these contexts, speakers have more control over tongue position and the opening/size of the channel, leading to more trilling. Importantly for the present study, higher rates of trilling were found post-vocalically than word-initially post-pause.

Similar to native speaker tendencies, for learners, trilling has been reported to be easier intervocalically than it is either word-initially post-pause or post-consonantally (Johnson, 2008, p. 90). Major (1986), who tested his Ontogeny Model (OM) by analysing the production of four beginner learners of Spanish whose L1 was English, confirms these positional effects. The model holds that transfer occurs mainly at the beginning stages of L2 acquisition and decreases as the learner receives more input and as L2 components emerge. Unlike transfer errors, Major (1986) claimed that developmental errors are a result of general learning processes, which may reflect errors that are present in L1 acquisition, but also realizations that are characteristic of neither the L1 nor the L2. To test the model, participants read word and sentence lists containing the target rhotics in different positions (word-initially, medially, and finally) on seven recording sessions, spanning the duration of

an intensive eight-week course. Results showed that, three participants improved on the tap in intervocalic position only, ranging from 3% accuracy initially, to 79% percent accuracy at the end of the course. For the tap, this was attributed to a positive transfer effect; namely, the fact that the tap occurs in English as an allophone of /t/ and /d/ intervocalically. Regarding the trill, while two of the participants showed improvement (ranging from 48% initially to 100% at the end) particularly in word-medial position, the other two participants only produced one trill out of a possible total of 347 over the course of the seven recording sessions. The fact that both rhotics were more target-like word-medially intervocalically suggests that speakers may be aware that this is where the contrast is realized. In addition to the results being explained by transfer effects, it may also be the case that there is generally a greater ease of producing sounds word-medially intervocalically than in other contexts, such as word-initially. Major concluded that intervocalic position is the most advantageous, since air flow is already present from the preceding vowel, and a rapid stream of air flow is imperative in trill production; as such, intervocalically, speakers have a 'running start' (p.492). Regarding word-initial position, Major noted that the speaker must "start from scratch" (p. 492), a factor which makes this position less optimal for trilling that intervocalic position does. Evidently, word-medial intervocalic position is more likely to favour (or disfavour) mastering all of the phonetic parameters of a trill than word-initially.

Colantoni and Steele (2007) also showed positional effects, specifically for the acquisition of the French uvular rhotic /ʁ/ by native English speakers. The authors hypothesized that learners would have the least difficulty acquiring sounds involving new combinations of voicing and manner in positions that favour the phonetic implementations

of these sounds. That is, the voiced fricative was expected to be acquired first intervocalically, followed by word-initial position, and finally in coda position. This was based on the fact that voicing is favoured intervocalically compared to word-initially (Kohler, 1984; Ohala & Kawasaki, 1984). Furthermore, it was hypothesized that, since not all parameters are acquired simultaneously, learners will first target aspects of a sound's articulation that are more perceptually salient and articulatorily less difficult. 20 L1 French speakers of L2 English proficiency ranging from intermediate to advanced, and 10 native speaker controls completed a French word-reading task and a French reading passage task. In the word-reading task, stimuli included 24 words containing stop+rhotic clusters controlled for voicing and position in the word (initial versus medial), and 14 tokens containing singleton <r>; all words were presented on note cards embedded in a carrier sentence *Je dis [TARGET] encore une fois* 'I say [TARGET] again'. The reading passage task included 21 words containing singleton rhotics. Data analysis included measuring each rhotic's length (ms), calculating its percentage of voicing, and determining its manner. Results from this study showed that manner and voicing interact and should be considered together when interpreting findings. For the advanced learners, when voicing alone was considered, they followed the predicted hierarchy, voicing slightly more word-medially followed by initially (89% versus 85%, but no statistical difference), then syllable-finally (word-final: 75%; word-medial: 64%). The intermediate learners followed a different hierarchy, with word-initial onsets having the lowest rate of voicing (52%), followed by the other contexts at nearly the same rates, with no statistical differences (word-final: 67%; word-medial coda: 73%; word-medial onset: 69%). However, when considered together, manner and voicing were in a trade-off relationship. Regarding the advanced speakers, they

produced the highest proportion of approximants word-finally, suggesting that, in contexts where they were unable to control both manner and voicing, the learners targeted the voicing parameter. With regard to the intermediate learners, the highest percentage of fricatives occurred word-initially, which is also the context in which they devoiced the most, once again supporting the notion that manner was targeted at the expense of voicing in this context. Overall, learners hyperarticulated in prominent positions, such as word-initially, resulting in emphasizing some parameters, while reducing their accuracy with others. In non-prominent positions, such as word-finally, hypoarticulation occurred, which surfaced as vocalization of the French rhotic. The findings of this study suggest that, different parameters of rhotics are targeted in different positions in the word.

Positional effects have also been found for the L2 acquisition of non-rhotic segments. In their study on the L2 acquisition of the Spanish spirants [β ð ɣ] by two groups of L1 English speakers, Rogers and Alvord (2014) investigated the production of these segments in both word-initial (not absolute initial) and word-medial positions. In Spanish, spirants are allophones of the voiced stops /b d /g/. Specifically, voiced stops surface as the spirants [β ð ɣ] in all contexts, except after a pause, a nasal (am[b]os *ambos* 'both'; san[g]re *sangre* 'blood'; ven[d]e *vende* 'he/she sells'), or a lateral in the case of /d/ only (e.g.,cal[d]o *caldo* 'soup/broth') (Hualde, 2005a, p. 138). In contrast, English has the voiced stop phonemes /b d g/ as well as the interdental fricative /ð/, which occurs in words such as *there* /**ð**ɛɹ/. In Rogers and Alvord's (2014) study, the first group (UL) consisted of four learners who had completed two years of university-level Spanish, while the second group (AB) consisted of four learners who had spent time (approximately two years) in Spanish-

speaking countries. A group of three native speakers served as the baseline for the study. Participants completed two tasks (reading a fictional story and an oral interview) eliciting the production of /b d g/. In order to determine the degree of spirantization, the authors calculated the intensity differences between the target consonant and the following vowel. The greater the intensity difference between the two sounds, the less lenition. Results showed that, for the learners, but not the controls, word position had an inverse relationship with intensity differences. Specifically, when compared to word-initial stops, word-internal stops exhibited a lower intensity difference, which was indicative of a higher degree of spirantization. These positional differences were larger in the UL than the AB group. The authors noted that a possible reason for the positional differences could be due to how learners perceive word boundaries in Spanish.

In a similar vein, Shea and Curtin (2011) analysed the production of [β ð ɣ] among ten native speakers of English with low and high intermediate (LI and HI, respectively) proficiency in L2 Spanish, as well as five native Mexican Spanish controls, with the primary goal of determining whether learners use categorical or gradient phonological systems when acquiring allophones in their L2. The general hypotheses guiding this study were that, a lack of differences across the three places of articulation would indicate categorical/phonological encoding of alternations, while place asymmetries would indicate a more gradient system. To test these hypotheses, the participants completed a reading task involving the target stimuli – real and nonce words containing <b d g> in word-initial and medial position, and in stressed and unstressed syllables – embedded in the carrier sentence *Diga* _____, *por favor* 'Say _____, please'. Data analysis involved examining the presence or absence of a

release burst and calculating a ratio by dividing the target consonant intensity by the consonant + vowel (CV) intensity. The presence of a release burst and an intensity ratio closer to 0 was taken to indicate a more stop-like sound, while the absence of a release and an intensity ratio closer to 1 was interpreted as a more approximant-like sound. Overall, the more advanced learners demonstrated more nuanced realizations of the phonetic parameters than the less advanced ones. Specifically, the LI group demonstrated a main effect for context (initial versus medial), but not for consonant (i.e., <b d g>), indicating a generally categorical acquisition pattern, whereby they may be applying a rule stating that <b d g> become 'softer' word-medially. The HI group, on the other hand, showed a main effect for both context and consonant (<b> versus <d/g>), as well as an interaction, revealing a more gradient pattern that more closely resembled the controls, who also showed a main effect and interaction for context and consonant (<b> versus <g>). Similar to Rogers and Alvord (2014), both learner groups did make a distinction between the initial and medial position, leniting the consonants more medially than initially. However, the graphical trends showed that, the LI group's average CV ratios were generally lower than both the controls' and HI group's (Controls: ~0.96-0.98 medially and ~0.90-0.94 initially; HI: ~0.93-0.97 medially and ~0.89-0.90 initially; LI: 0.86-0.88 medially and ~0.83 initially).

To conclude, the research on Spanish trills specifically has shown that, when compared to word-initial position, higher rates of trilling occur word-medially (Lewis, 2004; Major, 1986). This has also been observed for L2 French rhotics, with learners targeting voicing in word-medial position before word-initial (Colantoni & Steele, 2007). Lastly, similar positional effects have also been observed for non-rhotic segments, such as

the Spanish spirants. Specifically, Rogers and Alvord (2014) showed that both of their learner groups had higher rates of spirantization in word-internal versus initial position. Further, Shea and Curtin's (2011) study demonstrated that, while both learner groups distinguished between word positions, overall, the sounds produced were more approximant-like medially than initially. Having discussed positional effects, Section 2.2.3 now turns to the effect of task formality on L2 production.

2.2.3 Effects of task formality on L2 phonological acquisition

Research (e.g., Beebe, 1987; Colantoni & Steele, 2007; Díaz-Campos, 2006; Labov, 1972; Lindblom, 1989; Major, 1987, 2001; Rafat, 2010; Tarone, 1979, 1982, 1983; Zampini, 1994) has consistently shown that task formality is a factor that influences speech production, for L1 and L2 speakers alike. Recall from Section 2.2.1 that, according to Lindblom's (1989) H&H Theory, speakers vary their speech along a continuum of hypo- and hyper-speech depending on task formality. For example, the reading of a word list is a highly controlled task, which can lead to hyperarticulated, careful speech, while higher rates of lenited forms generally occur in more casual speech (Gilbert & Madrazo, 2017). Since more formal tasks activate hyperarticulated speech, we can expect to see longer segments (Colantoni & Steele, 2007; Rafat, 2008; 2010; Thomas, 2011, p. 138) that are more heavily stressed than the ones occurring in conversational or casual speech (Thomas, 2011, p. 138). Less formal speech, such as interviews, which is more hypo-articulated, exhibits more lenition phenomena, such as vowel and cluster reduction (Diehl & Lindblom, 2000; Lindblom, 1989). For example, Rafat (2010) found that native Farsi speakers had higher trill production rates in a more formal/careful task (a word list reading task) than in a more

spontaneous one (a memorization task). These notions can also be applied to L2 speech. The idea that the degree of task formality affects production is relevant here, since the present study elicited the production of Spanish rhotics using two tasks differing in the degree of formality: a sentence reading (more formal) and a picture description (less formal) task.

One related idea is that of observed speech. As discussed by Labov (1972), the term "observer's paradox" refers to the idea that, "the aim of linguistic research in the community must be to find out how people talk when they are not being systematically observed; yet we can only obtain this data by systematic observation" (p. 32). Tarone (1979) extended Labov's (1972) idea of the "observer's paradox" to IL, claiming that IL varies along a continuum of styles modulated by the amount of attention that a speaker pays to his or her speech. She describes the nature of IL as similar to that of a chameleon, in that both are extremely sensitive to their surroundings (Tarone, 1979, p. 188). The author outlined Labov's (1972) axioms and gives both anecdotal and experimental evidence for each, ultimately linking them to IL. The first axiom states that every speaker shifts his or her pronunciation as the situation changes, which Tarone (1979) claimed is a reality L2 researchers are often faced with when analysing the speech of their participants. The second axiom is that the continuous dimension along which speakers vary their style is defined by the amount of attention they pay to their speech. Attention, Tarone explained, is not an "all-or-nothing matter, but a matter of degree" (p. 183), such that there are not only two opposing modes, but rather a range of styles. The third axiom maintains that the vernacular style (i.e., the least attention given) exhibits the most regular and systematic phonological patterns,

whereas other styles are characterized by more variability. The IL system should be the least permeable in contexts that are more informal, when little attention is paid to speech, but should be the most permeable and variable in more formal situations, since the learner is paying attention to speech. The last relevant axiom discussed holds that a formal context is created when a speaker is systematically observed, which results in the speaker paying more attention than normal. Tarone noted that researchers have been too relaxed and inconsistent in their definition of "informal", "spontaneous speech", and "formal context", giving several examples of studies that have claimed to use spontaneous speech, all while testing participants in schools, or creating testing situations where an adult asks a child specific questions that are designed to elicit a target structure. The last axiom points to the idea that, despite the care taken to maximize spontaneity and create more informal testing environments, these simply end up being "formal" situations of varying degrees, which is ultimately unavoidable. By drawing a detailed parallel between Labov's (1972) axioms and IL, Tarone (1979) highlighted the importance of task formality in L2 research, an aspect which is now more often investigated in current research.

The axioms mentioned above are indeed applicable to L2 speakers, who generally exhibit different phonological patterns depending on the experimental task conducted. Some researchers (Major, 1987, 2001, p. 96; Tarone, 1979, 1982, 1983) have suggested that L1-based transfer increases in more casual speech styles, but decreases in more formal styles, since words are mainly produced in isolation in the latter. Major (1987) also offered some evidence from a pilot study investigating several aspects of L2 English pronunciation by L1 Brazilian Portuguese speakers. One aspect the author examined was the incidence of [i]

insertion word-finally (as in [dógi] for *dog*) and word-finally in consonant clusters across four tasks differing in formality by two groups of learners with varying levels of proficiency. The difficulty for these speakers is because in Portuguese, there are no word-final consonant clusters, nor word-final singleton consonants except for /s/. As such, Portuguese learners of English often insert an [i] in loanwords that contain such sequences. Major hypothesized that the presence of [i]-insertion in his participants' L2 speech would indicate that transfer had occurred. His findings revealed an overall tendency for errors involving [i] insertion to decrease as tasks became more formal, though the differences did not prove statistically significant. Despite this, Major concluded that his results provided a trend supporting the stylistic variation aspect of his OM. Still, some studies have found results contradicting the notion that more L1-based transfer is present in less formal speech.

Beebe's (1987) study contradicts, at least in part, the prediction made by Major's (1987) OM and Tarone's (1979) "Chameleon Model". Beebe (1987) investigated L1 Thai-L2 English speakers' production of English /ɹ/ in word-initial and final position using two tasks differing in formality. In the first task, participants took part in an interview with a native English speaker (informal task, according to the author), while in the second, they read a word list (formal task). The author's results showed that, in word-final position, participants' productions of English /ɹ/ did correspond to Major's (1987) and Tarone's (1979) predictions; that is, fewer errors were attested in the reading task than in the interview. However, against Major (1987) and Tarone (1979), in word-initial position, Beebe (1987) found more errors in the word list – the more formal task – than in the interview. Specifically, in the formal tasks, many productions reflected L1 transfer.

However, this discrepancy was attributed to a prestigious rhotic variant used in Thai only in careful speech: a trilled /r/. This variant occurred in 24.4% of the rhotic productions in the formal task, leading Beebe to conclude that the participants in her study had transferred the context (i.e., careful speech) of their L1 prestigious variant into L2 speech, which resulted in more of the prestigious variant in the formal task.

Some additional evidence that more accurate pronunciation occurs in informal rather than formal tasks comes from Zampini (1994), who investigated the production of spirants [β ð ɣ] by L1 English learners of Spanish, and Díaz-Campos (2006), who looked at the same segments in addition to the voiceless stops /p t k/, syllable-final laterals, and palatal nasals. Using a reading task (formal) and an informal interview, Zampini (1994) investigated spirantization rates by second- and fourth-semester Spanish students. The author's hypothesis was that both learner groups would exhibit a more target-like pronunciation in the informal interview versus the more formal reading task. Zampini's results show that, although spirantization rates were generally low for [ð] (informal: 5.8%; formal: 4.7%), overall, the learners exhibited higher spirantization rates in the interview rather than the reading task for all three spirants. The author attributed this finding to the fact that, in conversation, learners pay more attention to communicating effectively by using adequate vocabulary and grammar, and thus focus less on pronunciation, which, in turn, increases the possibility of weakening, a process that is required for the successful production of Spanish spirants.

Zampini's (1994) findings are echoed by Díaz-Campos (2006), who also found that the L2 learners in his study had more target-like productions of the intervocalic spirants,

voiceless stops, laterals, and palatalized nasals in conversational versus read-aloud tasks. In his study, 26 study abroad and 20 Spanish language course students completed a read-aloud task and participated in an Oral Proficiency Interview (OPI). The results form this study revealed that, for all of the segments under investigation, the conversational style favoured target-like productions compared to the read-aloud task. Though revealing of task differences, these findings may only hold for the L2 production of Spanish spirants, since these sounds are characterized by a weakening process, and may therefore not be generalizable to the acquisition of other segments that require more articulatory control. Regarding both Zampini's (1994) and Díaz-Campos's (2006) results, it is unclear how the productions were determined to be target or non-target, especially in the absence of a native Spanish speaker control group to serve as a baseline. Moreover, it is unclear if the data analysis was purely autditory, or if acoustic analyses were carried out. Using a more fine-grained analysis in their study, Rogers and Alvord (2014) showed that, as task formality increased, so did intensity differences and therefore lower rates of spirantization, but only for the UL (University Learners) group and not the AB group (Abroad Learners). Similar to Zampini (1994), Rogers and Alvord (2014) also pointed out that, because successful spirant production results from articulatory reduction, it is unsurprising that learners have higher spirantization rates in more spontaneous speech (versus read speech) where they exercise more constriction and a greater degree of articulatory reduction.

In summary, task formality is a factor that influences L2 production accuracy, at least to some extent. Though some studies have found more native-like speech in casual versus more formal tasks (Díaz-Campos, 2006; Zampini, 1994), this might be particularly

true for certain segments or positions in the word (Beebe, 1987). On the other hand, other research has suggested that L2 pronunciation tends to be more target-like in more formal (read) speech than it is in conversational speech (Major, 1987; Tarone, 1979, 1982, 1983). Since formal tasks (generally) involve the presence of orthography, the effect of orthography should also be considered, especially in cases where L2 contrasts are represented through spelling, the way the Spanish tap-trill contrast is intervocalically (i.e., <r> versus <rr>). As such, the following section reviews the literature on the effects, both facilitative and hindering, of orthography on L2 production.

2.2.4 Effects of orthography on L2 phonological acquisition

Since one of the two tasks in this study involves orthography[5], and word-medially intervocalically /r/ and /ɾ/ differ orthographically, while word-initially <r> represents the trill, it is important to discuss its potential effects on the production of speech. In tasks involving the written presentation of stimuli, such as word lists or carrier phrases, speech may be more target-like than normal due to reduced processing demands (Colantoni, Steele & Escudero, 2015, p. 110), with tokens approaching their phonetic targets more than those in spontaneous speech, exhibiting less coarticulation and undershoot (Thomas, 2011, p. 138-139). Moreover, while native speakers are only affected by orthography-internal factors, L2 speakers' production is affected by the interaction between their L1 and L2 orthographies

[5] Despite orthography not being available visually in the picture task, it is possible that the speakers still had mental orthographic representations of words even if they did not have a written representation.

(Bassetti, 2008, p. 198), possibly leading them to interpret the L2 graphemes using L1 grapheme-phoneme correspondences (Young-Scholten & Langer, 2015).

There is some evidence to suggest that orthography can have positive effects on L2 production. One study that provides support for positive effects is Erdener and Burnham (2005), who investigated the role of orthographic depth in the production of a foreign language. Orthographic depth refers to "the degree to which an alphabetic system deviates from simple one-to-one grapheme-to-phoneme correspondences" (Erdener & Burnham, 2005, p. 197). If a language is characterized mainly by one-to-one grapheme-to-phoneme correspondences, it is considered transparent; if it is a language in which one grapheme corresponds to more than one sound and vice versa, then it is considered opaque (Erdener & Burnham, 2005, p. 197). Although Spanish, the target language in the present study, is considered orthographically transparent (Ellis, Natsume, Stavropoulou, Hoxhallari, Van Daal, Polyzoe, & Petalas, 2004; Rafat & Perry, 2019), some aspects of it deviate from this and can be considered opaque. For example, in Spanish, <v> corresponds to [b] or [β]. Importantly, Spanish rhotics can be considered less transparent than Spanish overall, as <r> can correspond to both a trill and a tap, depending on its position in the word (i.e., trill word-initially and tap word-medially intervocalically) (Rafat & Perry, 2019). As such, the research on orthographic depth may provide insight that could be extrapolated to the L2 acquisition of rhotics in the present study.

The participants in Erdener and Burham's (2005) study were 32 adult speakers of Australian English, an opaque language, and 32 speakers of Turkish, a transparent language. Participants completed repetition tasks in which they heard both Irish (opaque) and Spanish

(transparent). Participants did not have any prior knowledge of either of the languages tested. There were four experimental conditions: Auditory-only, AV, AV-orthography, and Auditory-orthography. Results showed that, when orthography was presented and it was transparent (Spanish), the Turkish speakers made fewer phoneme errors than Australian speakers. For example, with regard to Spanish /p/ and /b/, the percentage of confusions was 4.55% for Turkish and 15% for Australian. However, when the orthographic information was presented for Irish (opaque), the Turkish speakers' performance was worse than that of the Australian speakers. The authors explain that, since Turkish is transparent, these speakers are familiar with processing orthographic information on a grapheme-to-phoneme correspondence basis, leading them to be more accurate with transparent languages like Spanish. Conversely, the Australian speakers relied less on orthography because they are speakers of English, which is opaque, and may therefore have a more "picture-orthographic" representation of items. Overall, these findings suggest that presenting participants with orthographic information is useful in L2 pronunciation, but more so if the language is transparent, and the degree of a language's opacity determines the facilitative effects of orthography. Regarding implications for the present study, the findings may suggest that, while the orthographic presence of the Spanish rhotics word-medially may be facilitative since it is transparent, word-initially it might be hindering, since it is less transparent, with <r> corresponding to both a tap and a trill depending on the context.

Similar to Erdener and Burham (2005), facilitative results were also obtained by Steele (2005). Using a vocabulary-learning task, Steele investigated the acquisition of stop-liquid clusters by L1 Mandarin-L2 French speakers of beginner proficiency. One group of

learners (N=13) was presented the words aurally accompanied by an image, while the other group (N=10) was also provided with orthography. Stimuli consisted of French words with stop-liquid clusters, which were controlled for place of articulation and voicing of the stop, as well as the type of liquid (e.g., *drapeau* /dʁapo/ 'flag'). Results showed that, the group exposed to orthographic input was more accurate in the realization of the /ʁ/ in a stop+rhotic cluster. Conversely, the group that was not exposed to orthography produced a large portion of highly aspirated stops, which was less present in the other group's realizations. The orthographic input group also had higher rates of fricatives following the obstruents. The high proportion of aspirated stops was attributed to the similarity between Mandarin aspirated stops and French voiceless stop+rhotic clusters, whereas the presence of orthography corrected this misanalysis. The presence of orthography, then, allowed the orthography group to realize that the target sounds involved two segments, not one, which led to more target-like productions. To conclude, the presence of orthography may positively influence the L2 production of sounds, which may especially be the case if the contrast is further reflected orthographically, the way Spanish taps and trills are word-medially intervocalically.

The effects of orthography may sometimes be hindering for L2 speakers, however. Bassetti and Atkinson (2015) investigated the effect of orthography among experienced learners (14 L1 Italian-L2 English) producing real words. Italian speakers were selected, since Italian is transparent (English is more phonologically opaque than Italian) and research has suggested that native speakers of languages that are orthographically more transparent rely on orthographic input more than speakers of less transparent orthographies (Erdener &

Burnham, 2005). The authors investigated ghjthe orthographic effect at the level of segments, specifically epenthetic consonants (Study 1) and vowel length[6] (Study 2). Study 1 had stimuli containing a silent letter (e.g., <b> in the word *lamb*) and aimed to determine whether the presence of the silent letter would result in learners pronouncing it. Participants first completed a reading-aloud task in which they saw a written word on a screen and had to simply read the word aloud. Then, they completed a word repetition task with acoustic and orthographic input. Specifically, as in the first task, a written word was presented on a screen, which disappeared and simultaneously, the participants heard a recording of that word produced by a native English speaker. The participants had to produce the word. Results showed that 85% of the experienced participants pronounced the silent letters in the reading-aloud task, while only 56% of them pronounced them in the word-repetition task. The authors concluded that exposure to native spoken input reduced the effects of orthography, though this may be only temporary; however, participants still realized many of the silent letters even after being exposed to the acoustic input. Study 2 involved pairs of stimuli containing the same target vowel, but represented by either a singleton vowel or a vowel digraph (e.g., *scene* and *seen*) and sought to answer whether the vowels represented by a single letter would be shorter than those represented by two letters. Participants received a list of words and had to read each word three times in a carrier phrase. Results showed that vowels spelled with digraphs had longer durations than those spelled with single letters, showing that L2 orthographic forms influence vowel durations among

[6] The study also included two additional experiments, one investigating the effect of morphemic spelling on production, and one focusing on the effect of orthographic differences on the production of homophonous words. However, these studies will not be discussed here, since they are not directly relevant to the present topic.

learners. Overall, the findings of this study suggest that learners may sometimes over-rely on L2 grapheme-phoneme corresponded rules (e.g., pronouncing the [l] in the word *salmon*), leading to non-target-like productions.

Detrimental effects of orthography have also often been reported in situations where a speaker's L1 and L2 use the same grapheme for two different sounds. An example from Spanish are [b] and [β], which are often realized as [v][7] by L1 English speakers, since orthographically, in addition to being represented by <b>, these sounds can also be represented by <v>, as in *vaca* [baka] 'cow' and *tuvo* [tuβo] 'he/she had'. As Zampini (1994) pointed out, it is unsurprising that [b] and [β] are predominantly realized as the voiced labiodental fricative in reading tasks rather than in conversation, "for the visual reinforcement of the written word serves to influence pronunciation" (p. 478). The explanation, then, is that the presence of orthographic <v> and the transfer of L1 phonological knowledge regarding this grapheme is causing L1 English-L2 Spanish speakers to have fewer target-like realization of [b] and [β] in more formal tasks (i.e., reading) than they do in more spontaneous ones (Zampini, 1994). Another example is the case of rhotics, where the orthography connects the English alveolar approximant /ɹ/ with the Spanish /r/[8] and /ɾ/, since the languages use the same grapheme to represents different phonetic sounds, which may yield English-like realizations, especially for beginner learners (Face, 2006; Colantoni & Steele, 2006; Olsen, 2016, p. 317). This is similar to the effect orthography could be having in this study in word-initial position, where <r> is a trill in

[7] This sound does not exist in Spanish.
[8] Word-medially intervocalically, /r/ is represented by <rr> in Spanish.

Spanish, but two different rhotics in French and Romanian as we will see in Sections 2.4.1 and 2.4.2. The negative effect of a known grapheme representing a new L2 sound was reported in Young-Scholten and Langer (2015). The authors carried out a one-year longitudinal study of three American students learning German in an immersion context. Specifically, the authors investigated the production of word-initial <s>, since in English it is pronounced [s], while in German it is [z]. The results showed that, the participants' tendency to realize <s> as the English [s] rather than the German [z] did not decrease over time, showing that learners continue to apply L1-based-grapheme-phoneme rules even after having been consistently exposed to the L2 for a period of time.

This hindering effect of orthography is also highlighted in the PAM-L2 (Best & Tyler, 2007), a perceptual model. Although the present study focuses on production, the PAM-L2 may be relevant, since it recognizes the role of orthography, and also focuses on contrastive L2 segments, allowing us to explain patterns in the acquisition of the Spanish tap and trill, as well as to investigate how these contrastive sounds map onto already existing L1 sounds. This model, whose primitives are distal articulatory gestures[9], states that the ability to learn new segments depends on the perceived distance between the target L2 sounds and the most similar L1 sounds. Moreover, it considers similarity across phonetics, phonology and orthography. The PAM-L2 predicts that, at least during the initial stages of acquisition, learners will show some CLI, where their NL is the source of transfer. The pertinent postulates are as follows: (1) Two Category Assimilation: the learner is able to perceive a difference between two contrastive L2 sounds and each non-native sound is

[9] This idea is rooted in Articulatory Phonology (Browman & Goldstein, 1992).

assimilated to a different native category, which leads to excellent discrimination; (2) Single Category Assimilation: the learner is not able to perceive a difference between two contrastive L2 sounds and they are both perceived as equally good exemplars of the same native category, which leads to poor discrimination; (3) Category Goodness Assimilation: the learner assimilates both non-native sounds to the same native category, but one sound is a better fit for this category, resulting in moderate to very good discrimination. At the phonetic level, a similarity may be perceived between two sounds if these are acoustically similar, be it in manner or place of articulation, or voicing. At the phonological level, two sounds may be perceived as similar if they share distributional properties or phonotactics. The PAM-L2 also mentions the role of orthography, indicating that orthography may bias category assimilation of new L2 sounds, especially in cases where the target language (TL) sound is different from the native language (NL). Best and Tyler (2007) mentioned the acquisition of the French uvular rhotic by L1 English speakers. The authors suggest that this sound may be perceived by English speakers as an approximant /ɹ/, since in both languages, the rhotic corresponds to the grapheme <r>, and in both languages, it is similar regarding syllable structure and phonotactics. Thus, orthography can be detrimental in cases where two languages share the same graphemic representation of two different sounds, leading to category assimilation and thus, poor discrimination. As mentioned, although the present thesis focuses on production alone, this model may also have some implications for production patterns; if the learners in the present study perceive similarities between a non-native and a native sound, they will map the non-native sound onto a native sound, resulting in the non-native sound being produced using L1 articulatory gestures. As such, despite the fact that the PAM-L2 is a perception model, it may still be relevant for production studies,

as it may be able to inform the discussion on pronunciation errors. Moreover, crucially, it highlights the role of orthography in L2 speech.

Face and Menke (2009) found a negative effect of orthography even among learners at advanced stages of acquisition. The authors investigated the production of the Spanish spirants by L1 English speakers at three proficiency levels equated with their level of education in Spanish: fourth-semester students, graduating majors, and PhD Spanish students. The participants were asked to read a short story in Spanish, which was slightly modified to create more occurrences of intervocalic stops. Results showed that, while the frequency of spirant production increased with proficiency level (36% for fourth-semester students, 62% for graduating majors, and 81% for PhD students), speakers at all levels maintained a distinction between <v> and <b>. Moreover, speakers also continued to realize <b> as a stop, even in contexts where this should be a spirant. Face & Menke concluded that, "English transfer appears to prevent the production of orthographic *v* as a stop and *b* as a fricative" (pp. 50-51). Vokic (2010) provided a similar explanation regarding L1 English-L2 Spanish speakers' realization of <d> as [d] instead of target [ð], noting that, despite the fact that English has a complex etymological writing system, the sound [d] is most often represented by <d>, and speakers are relying on their L1 orthography (p. 443).

The literature reviewed in this section provides support for facilitative as well as hindering effects of orthography. Specifically, orthography may lead to more target-like productions (e.g., Steele, 2005), especially if the language is transparent regarding a particular contrast or sound (Erdener & Burnham, 2005). Conversely, orthography can be hindering in cases where the L1 and L2 use the same grapheme to represent two different

sounds (e.g., Best & Tyler, 2007; Face & Menke, 2009; Vokic, 2010; Young-Scholten & Langer, 2015; Zampini, 1994). The presence of orthography has also been found to negatively affect the production of silent letters and vowel duration if the vowels are represented orthographically by two letters instead of one (Bassetti & Atkinson, 2015).

2.2.5 Summary of articulatory nature of trills, positional effects, task formality, and orthography

Sections 2.2.1-2.2.4 discussed the literature concerning the articulatory nature of taps and trills, positional effects on the production of segments, and the effect of task formality and orthography. To summarize, previous studies have show that taps are considerably less articulatorily difficult than trills, and trills require very specific constraints to be met if they are to be produced successfully. Due to their articulatorily difficult nature, trills are challenging both for native speakers and learners. When learners do produce trills, they are often longer than those of native speakers, a result of hyperarticulation. Regarding the variables involved in this study, the effects of position within the word, task formality, and orthography were also discussed. All in all, the majority of studies have shown that L2 segments are generally acquired word-medially before word-initially. Moreover, it seems to be the case that, while some segments, such as the Spanish spirants, may be more successfully produced by learners in more spontaneous (rather than more formal) tasks, generally, it is the case that learners are overall more target-like in more formal situations, where tasks may involve reading. A related variable is orthography, which, as shown, can be both hindering or facilitative. For the present study, the orthography may be hindering word-initially, where <r> can correspond to both taps and trills (depending on the previous

segment). Additionally, <r> represents different rhotics in all three languages under investigation. Lastly, the orthography maybe facilitative word-medially, where a contrast is realized between <r> as the tap and <rr> as the trill. Having discussed the variables pertinent for this study's hypotheses, the following section turns to descriptions of rhotics in Romanian, French and Spanish in order to highlight the ways in which rhotics differ across the three languages.

2.3 Rhotics in Romanian, French, and Spanish

In this section, the properties of rhotics in Romanian, French and Spanish are discussed. It is important to note that, in all three languages (as well as cross-linguistically, more generally), phonetically, rhotics are highly variable (Colantoni, et al., 2015), so much so that there has been discussion regarding whether they could be classified as one class acoustically (e.g., Wiese, 2001). There is not one, single phonetic correlate or physical property that all rhotics share; rather, rhotics as a class are better described as having a kind of family resemblance, where each member resembles another with respect to some property (Lindau, 1985, p. 166-167). Ladefoged and Maddieson (1996) noted that one key piece of evidence that all rhotics belong to the same class from a phonological point of view is that "rhotics of one type often alternate with other rhotics" (p. 216). The authors provided an example from Farsi, where trill allophones occur word-initially, taps intervocalically and voiceless trills word-finally. Moreover, rhotic sounds can vary in manner (e.g., the Spanish trill or the French fricative), as well as place (e.g., Spanish rhotics are alveolar, while the French rhotic is uvular). However, although the manner characterizing rhotics in the three languages under investigation in this study is quite different, they all belong to the group

liquids, sharing several characteristics such as phonotactic properties and the fact that they combine with various obstruents to form consonant clusters (Colantoni & Steele, 2005).

2.3.1 Romanian rhotics

Although the literature on Romanian rhotics is very limited, there is some research that serves as a point of departure for characterizing these sounds. Generally, the Romanian rhotic is described as a "dental vibrant" (Pușcariu, 1959) or an "alveolar vibrant" (Untu, 2011). This terminology, however, lacks detail, as it does not completely shed light on whether the "vibrant" is simple or multiple (i.e., a tap or a trill), as in Spanish, for example. This confusion in terminology has been manifested in some studies. For example, in addition to Chitoran's (2002) observation that the Romanian rhotic is predominantly a flap, as well as descriptions such as Pușcariu's (1959), in their study of Romanian liquid onset and coda clusters Marin and Pouplier (2014) assumed the Romanian rhotic to be a trill, following the work of Pușcariu (1959). The confusion might be due to the various descriptions of the sound as a "vibrant", making it difficult to determine whether the rhotic is simple or multiple in nature. As such, sources such as Marin and Pouplier (2014) have perhaps erroneously translated the Romanian word "vibrant" to English "trill", when this sound is in fact acoustically (and perceptually) mainly a tap.

The rhotic is represented orthographically by a single <r> in all contexts. From a phonological point of view, regarding its distribution, /ɾ/[10] occurs in various contexts, such as word-initially after a pause, in consonant clusters, intervocalically and finally (4-8).

(4)	#rV	*ramă*	‘frame’	/ˈɾa.mə/
(5)	Cr	*crin*	‘lily’	/ˈkɾin/
(6)	rC	*învârt*	‘I/they spin’	/ɨn.ˈvɨɾt/
(7)	VrV	*mare*	‘beach’	/ˈma.ɾe/
(8)	C#	*măr*	‘apple’	/ˈməɾ/

Chitoran (2002) stated that the Romanian /ɾ/ is usually realized as the flap [ɾ], but it can occasionally surface as a trill in word-initial position (p. 10). This type of allophonic distribution, where a trill may occur word-initially more often, and a tap (or flap) surfaces intervocalically, is also prevalent in Northern Italian (Recasens, 2002), Farsi (e.g., Ladefoged & Maddieson, 1996, p. 216), and Washili Shingazidja (Patin, 2013). That the relationship between taps and trills in Romanian is allophonic is also supported by Chitoran, Vasilescu, Vieru and Lamel (2014), who noted that, in modern Romanian, the trill can occur, but the most common variant is the tap, especially in prosodically weak positions. Likewise, Schulte (2003) also noted that the rhotic in modern Romanian is clearly a tap (as it is in Bulgarian, Serbian and Macedonina), above all in unmarked, non-emphatic contexts (p. 316). This may be due to the fact that speakers tend to hyperarticulate in strong positions,

[10] Although it is not entirely clear from previous research whether the phonological representation of the Romanian rhotic is /r/ or /ɾ/, the latter notation will be used here, since the tap is the most common realization, and the trill rarely occurs.

such as word-initially. Crucially, the tap-trill substitution (in any environment) does not cause changes in meaning, and the two sounds are not in complementary distribution.

Articulatory accounts of /ɾ/ describe it as follows: "the tip of the tongue […] touches the back of the upper teeth and the alveolar ridge, or only the alveolar ridge. Then, the tip of the tongue forms into a deep groove, akin to the bowl of a spoon. […] The sides of the tongue touch the palate in a thin band, on the left and on the right, in the area where the molars are fixed in the upper maxillary" (my translation) (Pușcariu, 1959, pp. 32-33). Mafteiu (n.d.) described its articulation in similar terms: "the tip of the tongue, thin and free, lightly touches, without pressing against, the alveolar ridge behind the upper teeth, and the sides of the tongue lean on the upper molars, to block the lateral escape of the airflow. In the middle, a groove forms, through which the airflow may escape, towards the tip of the tongue" (my translation). In addition to the tongue gestures detailed by Pușcariu (1959), importantly, he further mentioned that the sound may be produced using 1-5 vibrations. Specifically, he notes that, depending on the position within the word, *r* may be produced with a different number of vibrations: word-initially, it is produced, on average, with 2.57, word-medially with 1.38, and word-finally with 1.92 tongue vibrations (Lelescu as cited in Pușcariu, 1959). Pușcariu (1959) depicted the multiple-occlusion rhotic using the symbol /*r̄*/, while the standard Romanian rhotic was depicted using *r*[11].

Even though articulatory studies of the sound do not exist, some sources offer gestural descriptions of it based on individual writers' intuitions. The Romanian rhotic is

[11] Pușcariu (1959) notes that *r̄* is a sporadic realization of *r* encountered in some regions of the country. Although these regions are not specified, it is mentioned that they are in the West and North of the country.

considered among the most difficult segments for children (Pușcariu, 1959; Mafteiu, n.d.), and the inability to correctly articulate *r* by the age of around three years is considered a pronunciation defect and speech pathology is suggested to ameliorate the issue (Mafteiu, n.d.). Beyond the early years, adults may continue to display problems articulating <r>. They usually substitute this sound with a uvular rhotic [ʁ], similar to the *r* that occurs in French. This is considered a speech impediment, namely rhoticism, and can persist into adulthood. In summary, the Romanian /ɾ/ is considered difficult by native speakers, often leading to parents overemphasizing its pronunciation to children.

Though to my knowledge, few acoustic studies have been conducted on Romanian rhotics (e.g., Radu, 2016), the research that has been done does indeed support the notion that the Romanian rhotic is in fact predominantly a tap, not a trill. In his analysis of vocoids surrounding Romanian /ɾ/, Avram (1993) notes that the primary Romanian rhotic is the tap. Similarly, Savu (2012) investigated the structure of the tap, and the flanking vowel-like elements that sometimes accompany this sound in Romanian[12]. In this study, five native Romanian speakers read isolated words and nonce words containing <r> in various contexts, such as intervocalically, and adjacent to consonants. In addition to reporting the presence of epenthetic vowels flanking the tap, Savu (2012) also briefly noted that rhotics were realized as taps 86.6% of the time, while the remaining 13.4% of realizations were trills and approximants. However, as the interest of the study was to describe the effect of taps on the acoustics of flanking vowels, and not on characterizing Romanian rhotics, there is no

[12]Epenthetic vowels flanking rhotics are not restricted to these languages. They have been reported in a number of languages, such as Greek (Baltazani & Nicolaidis, 2013), Polish (Gudurić & Petrović 2005), Hungarian (Vago & Gósy, 2007), French and Spanish (Colantoni & Steele, 2005; 2007).

mention of the contexts in which the trills were produced. For example, it is not known whether these trills occurred predominantly word-initially (a tendency which Chitoran, 2002 notes) or their specific characteristics, such as the number of closures involved. Additionally, in the case of sporadic trills, there is no mention of the number of successive closures.

Similar results were reported by Radu (2016), who focused on the phonetic realization of Romanian /ɾ/. 10 native Romanian speakers performed two tasks: a picture description and a carrier sentence reading. Stimuli included real and nonce words containing word-initial and -medial, intervocalic *r*. For example, in the picture description task, participants were presented with pictures depicting words such as *raţă* 'duck', where /ɾ/ appears word-initially, and *gară* 'train station', where /ɾ/ appears word-medially in intervocalic position. In addition to the stimuli in the picture description task, the sentence reading task also included nonce words. Some examples include *rela* 'NONCE', with the <r> word-initially, and *bera* 'NONCE', where the <r> is intervocalic.

Like Savu (2012), Radu (2016) also found that the primary realization of the Romanian rhotic is the tap, which occurred in 72% of cases overall. The second most common realization were trills, occurring 24% of the time. Other realizations included approximants (3%) and fricatives (1%). Crucially, the trill occurred word-initially more than word-medially (39% versus 8%), supporting previous observations that the trill may surface word-initially. Moreover, there were some notable differences between the tasks. Specifically, the trill was produced more often in the reading than in the picture description

task (30% versus 9%). Finally, Radu (2016) also found that the number of closures in trills ranged between 1 and 4.

Below are three waveforms and spectrograms that are representative of the Romanian rhotics. Figure 1 displays the default tap realization of word-initial /ɾ/ in Romanian, as in the word ***r****ecunoască* 'recognize (inf.)', while Figure 2 depicts a sporadic trill realization (with two successive occlusions) of word-initial [r] in Romanian, as in the word ***r****ană* 'wound (noun)'. Figure 3 depicts a word-medial tap in the word *care* 'which'. It can be observed that the tap is characterized by a short closure, while the trill is characterized by three periods of vibration, consisting of a closure and an open phase. Overall, the tap is much briefer than the trill.

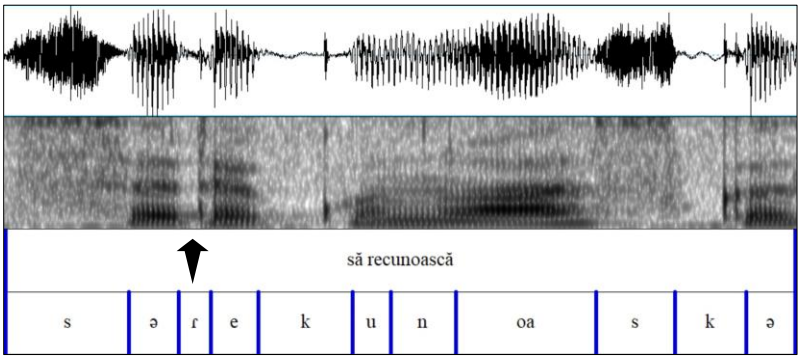

Figure 1. Tap production of word-initial /ɾ/ in a VrV sequence in *să recunoască* 'to recognize' produced by RN005, a female native speaker of Romanian (Source: http://rpd.chass.utoronto.ca/).

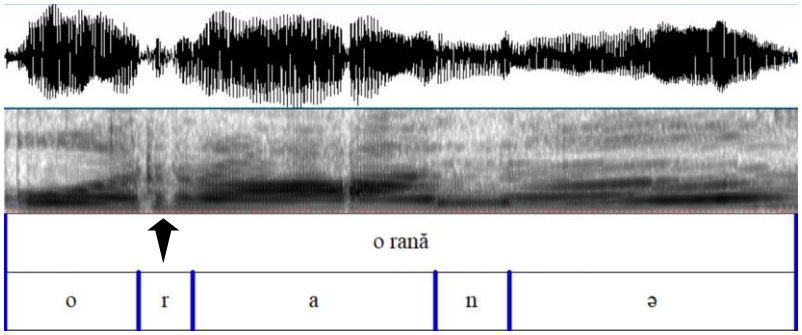

Figure 2. Two-occlusion trill production of /r/ in a VrV sequence in *o rană* 'a wound' produced by R07, a female native speaker of Romanian, in a picture description task (Source: Radu, 2016).

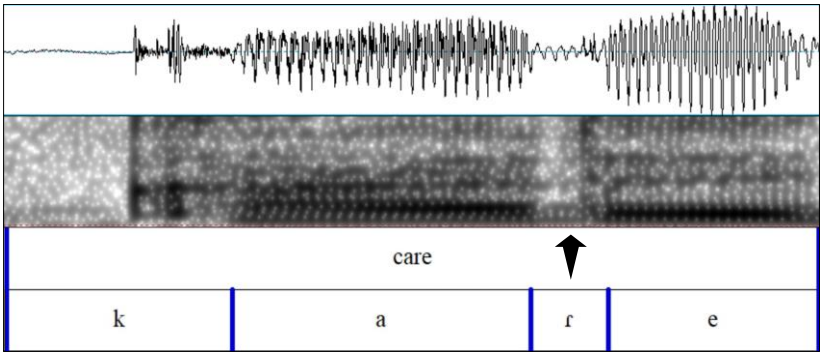

Figure 3. Tap production of word-medial /ɾ/ in the word *care* 'which' by RN005, a female native speaker of Romanian (Source: http://rpd.chass.utoronto.ca/).

In sum, based on the previous research and observations regarding Romanian rhotics, we can conclude the following. First, Romanian has mainly one rhotic, the tap /ɾ/, but the trill [r] may be produced sporadically as an allophone of the tap, predominantly word-initially (Chitoran, 2002; Pușcariu, 1959; Radu, 2016), in more careful speech (Radu, 2016). As such, it could be concluded that, in Romanian, the tap and the trill appear in free variation, where the allophones are not in complementary distribution and they do not cause

a change in meaning. Though the position in the word has been reported as a variable that affects the rhotic produced, this process is not systematic; that is, it is not always the case that trills occur word-initially in Romanian. The following section provides a detailed description of French rhotics.

2.3.2 French rhotics

Most varieties of French have been described as having only one rhotic, which is phonetically and articulatorily very different from the Spanish tap and trill. Accounts of the phonetics and pronunciation of most French dialects have described the <r> as the voiced dorsal (post-dorsal-uvular) fricative phoneme /ʁ/ (Léon, 1992; O'Shaughnessy, 1982). The <r> in most French dialects is currently described *grasseyé*, meaning "pronouncing Rs in a guttural manner" due to its frication. Its articulation is characterized by the back of the tongue against the uvula, which does not vibrate (Léon & Léon, 2009). Although the rhotic is typically described as a voiced velar fricative, realizations varying in both manner and voicing have been attested. These variants often occur in rapid speech (Sankoff & Blondeau, 2007), and are generally constrained by phonetic principles (Tousignant, 1987). They are often related to the position of <r> within the word. Namely, though most researchers have claimed that the sound is voiced, the rhotic may often be devoiced following a voiceless obstruent, as well as syllable-finally (Colantoni & Steele, 2007; Colantoni & Steele 2011), such as in *train* "train", which may be pronounced [tʁ̥ɛ]. Regarding manner, in addition to the commonly described fricative, trill, approximant realizations, as well as vocalization in codas, have also been attested (O'Shaughnessy, 1982; Colantoni & Steele, 2006b). Dialectal variation has also been attested (e.g., Sankoff & Blondeau, 2007 for Montreal French);

however, because all of the native French speakers in this study indeed had the expected uvular pronunciations of the rhotic in French (i.e., /ʁ/), as evidenced by their pronunciation in the French picture description and sentence reading tasks they completed, only this realization will be discussed here.

From an articulatory point of view, the French /ʁ/ has been described as difficult due to several of its characteristics. Specifically, the sound is both a fricative (Maddieson, 1994) and a dorsal, both characteristics which make voicing difficult (Ohala, 1997). While fricatives require abundant airflow in order to be articulated as such, this increased airflow can, in turn, inhibit voicing (Colantoni & Steele, 2007). Moreover, an additional articulatory difficulty lies in the simultaneous voicing and place of articulation of the French rhotic. Namely, in contrast with more anterior places of articulation, such as labial and coronal, voicing is typically not preferred in dorsals (Ohala, 1997).

With regard to its orthography and distribution, the French rhotic can appear in all positions and is represented orthographically by both <r> and <rr> (Léon, 1992). As examples (9-14) show, the single <r> may appear word-initially, word-medially, word-finally followed by <e>, as well as in absolute word-final position. Moreover, it can appear in obstruent + rhotic clusters, such as in (13) and (14).

(9) *rouge* ‘red’ /ʁuʒ/

(10) *vérité* ‘truth’ /veʁite/

(11) *faire* ‘to do’ /fɛʁ/

(12) *finir* ‘to finish’ /finiʁ/

(13) *croître* ‘to increase’ /kʁwatʁ/

(14) *foudre* “thunder” /fudʁ/

Unlike other consonants in word-final position (unless undergoing liaison in pre-vocalic contexts), rhotics are not deleted[13] (Proctor, 2009, p. 21).

Below are two spectrograms that are representative of the standard voiced uvular fricative /ʁ/. Figure 4 displays the rhotic in word-initial position, as in the word ***r****enonça* ‘she gave up, while Figure 5 depicts a /ʁ/ in word-medial, intervocalic position, as in the word *se**r**ait* “to be-3.sg.pres.cond.”

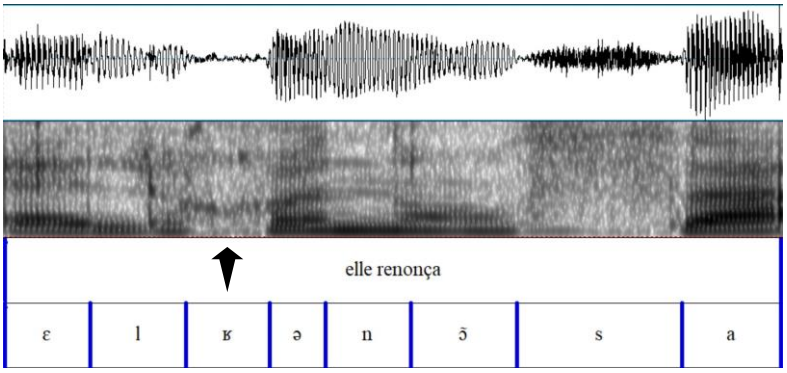

Figure 4. Production of word-initial /ʁ/ in a VrV sequence in *elle renonça* ‘she gave up’ produced by FN028, a female native speaker of French (Source: http://rpd.chass.utoronto.ca/).

[13] In fact, liquids in general, since /l/ is not deleted either.

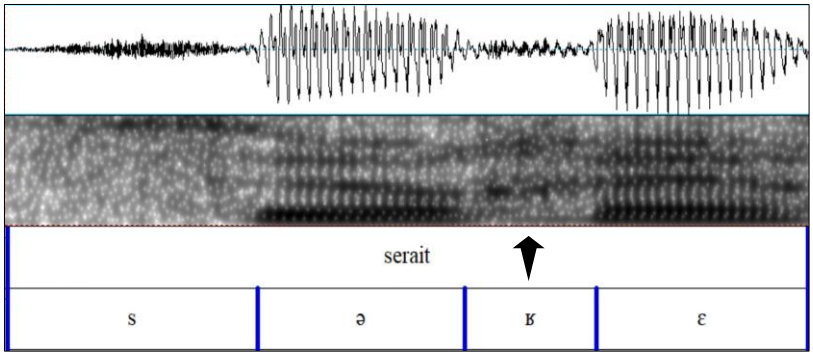

Figure 5. Production of word-medial, intervocalic /ʁ/ in the word *serait* 'to be-3.sg.pres.cond.' by FN026, a female native speaker of French (Source: http://rpd.chass.utoronto.ca/).

2.3.3 Spanish rhotics

Spanish has two rhotics: the tap /ɾ/ and the trill /r/. These two rhotics have a limited distribution and are quasi-contrastive (Hualde, 2005b). Word-medially, intervocalically, the tap and trill are contrastive, resulting in words with completely different meanings, as in examples (15) and (16). However, in all other contexts, they are in complementary distribution. Specifically, the trill is found word-initially and after /n/, /s/ or /l/ as in examples (17-20), while the tap is found in virtually all other contexts, as in (21-23).

(15)	*caro*	'expensive'	/ˈka.ɾo/
(16)	*carro*	'car'	/ˈka.ro/
(17)	*rata*	'rat'	/ˈra.ta/
(18)	*honrado*	'honest'	/on.ˈra.ðo/
(19)	*Israel*	'Israel'	/iz.ˈra.el/
(20)	*alrededor*	'around'	/al.re.ðe.ðoɾ/

(21) *broma* ‘prank’ /ˈbɾo.ma/

(22) *parte* ‘part’ /ˈpaɾ.te/

(23) *ser* ‘to be’ /ˈseɾ/

As can be seen in (15) and (16), the contrast between /ɾ/ and /r/ is also reflected in the orthography. Specifically, the tap is represented by <r> intervocalically, as well as in post-nuclear position (Quilis, 1993). The trill, /r/, is represented by <rr> intervocalically when it contrasts with the tap. In other positions, such as word-initially or after *n*, *l*, or *s*, it is represented by single <r> orthographically (Quilis, 1993).

Taps and trills are similar in that their articulation involves the tip of the tongue and the alveolar ridge, resulting in one or more occlusions; moreover, the production of both rhotics involves the tongue having a concave shape (Quilis, 1993). The tap is described as a simple alveolar vibrant (Navarro Tomás, 1971), and its articulation involves a single closure caused by the tip of the tongue striking the alveolar ridge (Navarro Tomás, 1971; Quilis, 1993). On average, its duration is approximately 25 ms (Quilis, 1993). Trills, which are described as multiple vibrants, have a duration of 85 ms on average (Quilis, 1981), but can be as long as 146 ms medially and 162 ms initially (Mendoza, Carballo, Cruz, Fresneda, Muñz and Marrero, 2003). It is important to note that the trill is not merely a sequence of taps, as trills involve more energy and muscular tension than taps do (Solé, 2002; Guitart, 2004, p. 145, my translation). Trill articulation involves two or more brief closures, which, as in the case of the tap, are caused by the tip of tongue striking the alveolar ridge (Quilis, 1993).

It is worth mentioning that, because taps and trills are only contrastive word-medially but in complementary distribution elsewhere, the status of the Spanish rhotics as separate phonemes has been the topic of debate. On one side of the debate, it has been proposed that there are two separate phonemes (Hualde, 2005b; Proctor, 2009; Solé, 2002), with a neutralization rule applying everywhere except intervocalically (Bradley, 2001), while on the other side, it has been claimed that there is only one phoneme, /ɾ/, for all contexts, except intervocalically where the underlying representation is a geminate tap (e.g., Harris, 1983, 2001). Researchers on the former side of the debate have tried to account for the limited context of the contrast (i.e., word-medially), while those on the latter side have had to account for this contrast. In his analysis, Hualde (2005b) notes that the tap and trill are quasi-phonemes. Based on the numerous minimal pairs available, as well as the syllabification of the trill as an onset (e.g., Spanish *carro* 'car' *ca.rro* versus Italian *carro* 'cart' *car.ro*), Hualde (2005b) concluded that the tap and trill are two separate phonemes. Based on Hualde (2005b) and Proctor (2009), and following other L2 studies (e.g., Johnson, 2008; Major, 1986; Olsen, 2012, 2016), the separate phoneme analysis is adopted in the present thesis. That is, it is assumed that Spanish has two separate phonemes, /r/ and / ɾ/, which contrast word-medially intervocalically, and neutralize elsewhere.

Despite normative descriptions of the Spanish rhotics, acoustic studies have generally suggested that there is a considerable amount of variation in manner, even among native speakers of Spanish, suggesting that there is individual variability (e.g., Blecua, 2001; Colantoni, 2006 for Argentine rhotics; Navarro Tomás, 1971 for Peninsular Spanish; Quilis, 1993; Solé, 2002). Both rhotics may be produced without full occlusion, yielding

approximants or fricatives (Blecua, 2001; Hammond, 1999). Regarding trills specifically, these sounds entail meeting very specific articulatory and aerodynamic requirements. To produce trills, the tip of the tongue moves forward in order to touch the alveolar ridge; this movement requires a very precise constriction of the tongue's muscles, specifically of those muscles present in the apex of the tongue (Quilis, 1993, p. 344). As such, there is a notable amount of variation in their realization (Solé, 2002). In trills with three occlusions, for example, the first two may be complete, tense closures, while the third may be realized with some frication (Quilis, 1993). Additionally, if specific aerodynamic constraints are not met, the resulting sound could be produced without tongue-tip vibrations, yielding non-trilled variants altogether (Solé, 1992). Among these variants, we also find fricatives, which may alternate with trills because they do not involve an articulation as complex as that of trills (Solé, 2002).

While there are different realizations due to articulatory constraints, there is also considerable dialectal variation. The dialectal variants could be narrowed down to four major categories, according to Hualde (2005a). The first category is assibilated rhotics, and while the tap may be produced in this manner, it is mostly the trill that exhibits these realizations (Quilis, 1993, p. 347). Assibilated rhotics are generally represented by the symbol [ř] and include both fricative or assibilated productions. According to Quilis (1993, pp. 349-350), assibilated rhotics: (a) are continuants as opposed to interrupted, as /r/ is; (b) are normally voiced (they are voiceless in only 12.5% of cases); and (c) display a turbulent frication in the higher formants. The second category includes dorsalized trills.

In addition to the dialectal variants described above, the dialects spoken in the Caribbean and some parts of Spain (i.e., Andalusia and Extremadura) are also well-known for liquid neutralization in syllable-final position (Hualde, 2005a; Proctor, 2009). For example, while in most dialects *mal* 'bad' and *mar* 'sea' are minimal pairs – the former being pronounced with /l/ while the latter being pronounced with /ɾ/ – the dialects mentioned above have neutralized the distinction between the two liquids, rendering the words homophonous. This neutralization is not necessarily in favour of one sound over the other; as such, the word pairs mentioned above could be pronounced as both [mal] and [maɾ]. In these same regions, two other phenomena have been attested: rhotacism and lambdacism. Rhoticism refers to the realization of coda laterals being realized as rhotics, and it has been attested in certain regions of Cuba, such as Havana, as well as Venezuela and Andalusia (see Proctor, 2009 for an overview). In these regions, a word such as *alguno* 'someone/somebody' may be pronounced [aɾɣuno]. Conversely, lambdacism is a phenomenon that refers to coda rhotics being realized as laterals, and has been attested in parts of Cuba, Panama, the Dominican Republic, and Puerto Rico (see Proctor, 2009, for an overview). In these dialects, *comprar* 'to buy' would be pronounced [compɾal]. Considering the wide range of phenomena that may affect rhotics in different Spanish varieties, the native speaker controls that participated in this study did not come from any of the regions mentioned above in order to avoid the possibility of different variants being produced due to dialect.

Below are two spectrograms that are representative of the standard tap and trill in contrastive environments in Spanish. Figure 6 displays the trill production of word-medial

/r/, as in the word ***arr****asa* 'flatten-3.sg.pres', while Figure 7 depicts a tap in word-medial position, as in the word *fu**r**ia* "fury".

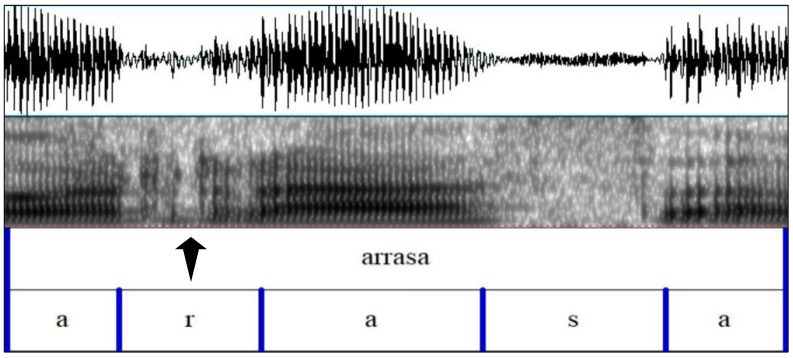

Figure 6. Trill production of word-medial /r/ in the word *arrasa* 'flatten-3.sg.pres.' produced by SN036, a female native speaker of Spanish (Source: http://rpd.chass.utoronto.ca/).

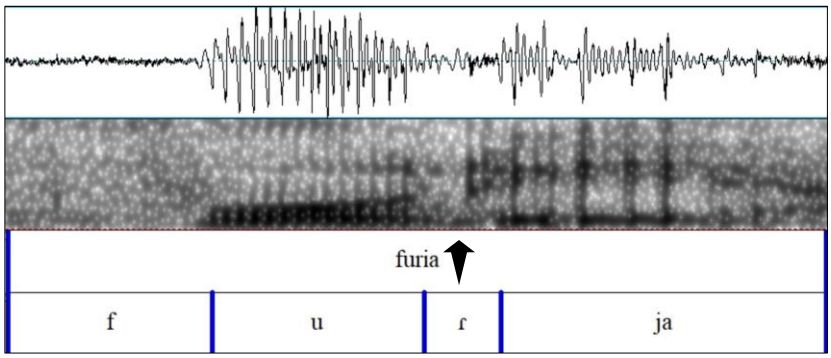

Figure 7. Tap production of word-medial /ɾ/ in the word *furia* 'fury produced by S03, a female native speaker of Spanish (Source: Radu, 2016).

The distribution of the rhotics in the three languages is summarized in Table 1.

Having described the nature of rhotics in the three languages under investigations, the next section focuses on the acquisition of the Spanish rhotics by native English speakers.

Table 1

Distribution of Rhotics in Spanish, Romanian and French

	Rhotic		
Language	**/ɾ/**	**/r/**	**/ʁ/**
Spanish	YES: intervocalically; onset cluster; word-finally	YES: intervocalically; word-initially; after /n l s/ in a different syllable; sometimes word-finally	NO
Examples	*pe**r**o* ‘but’ *b**r**oma* ‘joke’ *amo**r*** ‘love’	*pe**rr**o* ‘dog’ ***r**ápido* ‘quickly’ *al**r**ededor* ‘around’ *amor* ‘love’	
French	NO	NO	YES
Examples			***r**epas* ‘meal’ *fe**r**mer* ‘close (inf.)’ *jou**r*** ‘day’
Romanian	YES: all contexts	YES/NO: sporadically, in word-initial position	NO
Examples	***r**educe**r**e* ‘reduction’	***r**educere* ‘reduction’	

2.4 The L2 acquisition of Spanish rhotics

Studies on the L2 acquisition of Spanish rhotics are important for two reasons. One, they are directly related to the present study due to the structures under investigation, namely the acquisition of Spanish rhotics. Second, these studies provide a worthwhile insight into the acquisition of an L2 contrast when one of the sounds is present in the L1. Although the trill does not exist in English, the tap does occur as an allophone of /d/ and /t/ in intervocalic position, such as in the word *bu**tt**er* [ˈbʌ**ɾ**ər]. Overall, the objective of research on L2 Spanish rhotics has been (i) to determine which rhotic is acquired more easily (i.e., the tap or the trill), (ii) to identify if rhotics are first acquired word-initially, medially, or finally, and (iii) lastly, to investigate the order in which specific acoustic parameters of L2 rhotics

surface (i.e., duration, voicing or manner). More recently, there has also been an attempt to determine how L2 Spanish speakers realize the contrast between intervocalic taps and trills when they have not yet acquired manner (Amengual, 2016). With regard to manner, results from several studies (Colantoni & Steele, 2008; Face, 2006; Olsen, 2012, 2016; Waltmunson, 2005) have generally shown that, not only is the trill problematic for learners, the distribution and realization of the tap is also difficult, and even advanced learners do not completely master this sound.

Since Spanish has two rhotics, learners are faced with the task of acquiring both segments, as well as the context in which each of these sounds occurs. Several studies have found that learners generally do not acquire both sounds simultaneously; specifically, the consensus is that trills are acquired later than taps. Face (2006) investigated the acquisition of intervocalic Spanish rhotics by L1 English learners with different levels of proficiency (intermediate and advanced[14]). Participants were recorded reading a short story, which contained 10 tokens of each rhotic. Results showed that although the advanced learners were more successful at producing both rhotics than the intermediate group (tap accuracy: 78.7% versus 26.2%; trill accuracy: 48.5% versus 5.1%), even at the advanced level, learners had difficulty with trills. Specifically, while the intermediate learners' non-target production of the trill was generally the English voiced alveolar approximant (52%), the advanced group produced the target trill as a tap 78% of the time. Face (2006) interpreted these results as

[14] These proficiency levels were determined based on the Spanish course in which students were enrolled at the time of testing; specifically, intermediate learners corresponded to students enrolled in a fourth semester university language course, while the advanced learners corresponded to students enrolled in an upper division elective course.

evidence that learners find it easier to re-categorize their L1 tap (an allophone of intervocalic /t/ and /d/ in English), and to overgeneralize this sound to other contexts as well, than to produce the articulatory demanding trill. Similarly, Johnson (2008) investigated the production of Spanish /r/ by L1 English speakers who read a word list with the trill and tap appearing word-initially after a pause (only the trill, as in *río* 'river'), intervocalically (the trill, as in *terreno* 'terrain'; the tap, as in *pareja* 'couple'), and word-initially post-consonantally (only the trill, as in *son rojos* 'they are red'). Like Face (2006), Johnson (2008) study investigated whether the acquisition of Spanish /r/ improved among L1 English speakers across proficiency groups. Results showed that, learners' productions were characterized by stages, beginning with English /ɹ/, followed by taps, and only at later levels of proficiency being able to articulate trills. Although there was a notable "spike" in trill production between the beginner and intermediate proficiency levels (10% versus 62%, respectively), as well as the intermediate and advanced proficiencies (62% versus 82%, respectively), the advanced speakers still did not produce trills at native-like rates (90%). Generally, these studies demonstrate the difficulty associated with producing trills, and the consequences this difficulty has for learners; specifically, in the process of acquiring trills, at least until the intermediate stages of acquisition, learners appear to produce taps instead of target /r/.

In a similar vein, Olsen (2012) found that among beginner L1 English speakers with different levels of exposure to Spanish[15], rhotic accuracy increased with more exposure to

[15] Although all speakers were enrolled in a beginner Spanish course, participants used a Likert scale (1 to 5) to report varying levels of exposure to Spanish.

the TL. In this study, 51 participants completed a reading task, which was a short Spanish text containing 32 taps and four trills. Additionally, participants were also recorded reading four English words, which contained the English /ɹ/; this was done in order to determine how L1 articulatory routines (i.e., bunched or retroflex realization of English /ɹ/) affect L2 Spanish rhotic production. When controlling for amount of exposure to Spanish, learners with a more retroflex pronunciation of English /ɹ/ were able to produce the Spanish rhotics more accurately than those with a bunched-like pronunciation. Specifically, English rhotic articulation was a significant predictor of L2 Spanish rhotic accuracy only among learners who reported having little prior exposure to Spanish. Based on these results, Olsen (2012) concluded that L1 phonetic factors are less important among learners with greater L2 exposure. Thus, previous studies have shown that the tap is generally acquired before the trill (Face, 2006; Johnson, 2008), and it is possible that L1 articulatory routines can affect L2 rhotic production (Olsen, 2012). However, L1 articulatory routines do seem to diminish among more proficient speakers (Olsen, 2016), which is a consistent pattern in L2 phonology, as posited in Major's (1986) OM.

Olsen (2016) aimed to replicate the study conducted in Olsen (2012), but with more proficient learners. In the follow-up study, 35 L1 English speakers of intermediate L2 Spanish proficiency completed the same tasks as in Olsen (2012). Results showed that the more proficient group had higher tap accuracy rates than the beginners in Olsen (2012) (68.9% versus 56.3%), corroborating Face's (2006) findings of improved rhotic production over time. With regard to the trill, more intermediate participants were able to produce the trill when compared to the beginner ones (40% versus 14.6%). An interesting finding

regarding trill accuracy, however, was that the beginner participants in Olsen (2012) produced trills with higher accuracy rates than the intermediate speakers (27% versus 35.7%). The author attributed this discrepancy to a more robust sample size of trills elicited in the more recent study, noting that the accuracy rates reported in Olsen (2012) may have been due to the small sample size. In terms of L1 articulatory routines and exposure to Spanish, unlike Olsen (2012), Olsen (2016) found that these factors did not have significant effects. This follow-up study suggests that, while L1 articulatory routines may negatively affect L2 production, this effect may not persist into intermediate levels of proficiency.

As discussed in Section 2.2.2, in the development of L2 categories, learners do not target sounds in all positions equally. One of the principal factors that has been found to influence the L2 acquisition of rhotics by L1 English speakers is the position within the word and phonetic context. Specifically, several studies have found that the English phonetic context that produces allophonic taps – intervocalic /t/ and /d/ – influences the production of Spanish taps (e.g., Colantoni & Steele, 2008; Major, 1986). Using a sentence reading and a reading passage task, Colantoni and Steele (2008) investigated the L2 production of intervocalic, word-medial and word-final[16] taps among native English speakers of intermediate and advanced proficiency levels in Spanish[17], and indeed found that taps were first acquired in word-medial onsets, followed by word-medial codas and then word-finally. Their findings showed that, while learners had non-target-like realizations word-finally, overall, advanced Spanish learners were successful in producing

[16] These specific contexts were: VɾV (intervocalic, word-medial); VɾCV (word-medial); Vɾ# (word-final).
[17] This study also investigated intermediate and advanced learners of French, since the objective was to investigate interlingual difficulty.

the tap intervocalically, as evidenced by their mastery of all parameters of the tap (i.e., length, manner, voicing); intermediate learners acquired both length and manner, but not voicing. It was concluded that learners were more native-like in intervocalic position because this is where the tap/trill contrast is realized, causing learners to target this position first. Moreover, the authors attributed the mastery of the intervocalic tap to transfer from L1 English. Likewise, Olsen (2012), also found that learners were more successful at producing the taps in environments that lead to taps in English than taps in other environments (61.6% versus 45.4%), showing a clear influence of L1 phonetic context. Furthermore, Olsen (2016) showed that this effect of L1 phonology persists even among intermediate learners, as evidenced by the fact that they produced more accurate taps in environments similar to their L1. Similarly, Waltmunson (2005), who investigated the relative difficulty of tap and trill acquisition with respect to /t/ and /d/ using two reading tasks, also found an effect of word position on L2 rhotic production. The author noted that, if orthography affects accuracy, we might expect higher accuracy rates for the trill word-medially than initially, since <rr> and <r> correspond to the trill and tap, respectively. Waltmunson (2005) found that, while the trill was overall more difficult than the tap due to the need to learn new gestural skills as well as a new abstract representation, the position within the word is indeed a major factor. Namely, learners were more successful at producing /r/ word-medially than word-initially; likewise, they mastered /ɾ/ in word medial position before /ɾ/ in onset clusters. These findings are consistent with Johnson's (2008), who found that learners ceased to transfer their L1 English /ɹ/ in intervocalic position before any other position in their L2 Spanish. Moreover, he found that the trill was less difficult for the learners intervocalically than word-initially (both after a pause and after a consonant). These results can be explained by

the positional effects discussed in Section 2.2.2, namely that sounds may be more difficult in word-initial rather than medial intervocalic position. Thus, overall, these studies show that acquiring a phonemic contrast (i.e., /r/ versus /ɾ/) is not an all-or-nothing process, and learners first target contexts where a contrast is realized (i.e., intervocalically), and then move on to contexts where taps and trills are in complementary distribution (Colantoni, et al., 2015, p. 254).

In addition to difficulty in mastering target sounds in all phonetic contexts at the same time, studies have also shown that the acoustic parameters characterizing rhotics are not simultaneously acquired by speakers. Though most studies have relied on global measures (e.g., Face, 2006; Olsen, 2012), it is important to analyse the different acoustic parameters of rhotics, since not all of these are acquired simultaneously in L2 speech (Colantoni & Steele, 2008). Moreover, this kind of analysis allows for an objective identification of the differences in parameters between learners and native speakers. Colantoni and Steele (2008) found that, while learners had mastered duration, manner and voicing of taps in intervocalic position, they differed from controls in syllable-final position and word-medially, where they had only acquired length (but not manner and voicing), as well as in word-final position, where they were not native-like with regard to manner. This gradient acquisition of rhotics is also attested by Johnson (2008), who found that the number of trill occlusions was significantly lower for the learners than for native speakers. In a more recent study, Amengual (2016)[18] investigated L1 English-L2 Spanish speakers' production

[18] In addition to L1 English-L2 Spanish speakers, Amengual (2016) also investigated the production of rhotics by two groups of heritage speakers of Spanish in California. The groups were divided into (1) Spanish dominant and (2) English dominant.

of Spanish taps and trills intervocalically, a context in which the two sounds are contrastive. Using a read-aloud task, Amengual (2016) found that speakers do not master all parameters simultaneously. Namely, the high-intermediate L2 Spanish group in this study used duration to mark the distinction between intervocalic taps and trills in contrastive pairs (e.g., *carro* 'car' versus *caro* 'expensive). As such, although these learners produced trills in a non-native-like manner, using fewer occlusions than native speakers (i.e., using one or no occlusions), they were indeed still distinguishing between the two rhotics using duration; namely, trills were significantly longer in duration than taps were. Similarly, Waltmunson (2005) also measured several parameters of rhotics in his study[19]. Specifically, in addition to using auditory analysis to assess production accuracy, he compared the duration, the degree of voicing, and the presence of closures in the oral tract[20] of learners and that of native speakers. These measurements, however, were only reported for native speakers, while for learners, only global scores were reported, making it difficult to identify specific parameters in which learners and controls differed.

To recapitulate, based on previous research regarding the production of Spanish rhotics by L1 English speakers, we can conclude the following: (1) the trill is generally acquired after the tap, and even at advanced levels of proficiency, learners have difficulty with the realization and distribution of the trill (Colantoni & Steele, 2008; Face, 2006; Olsen, 2012; Waltmunson, 2005); (2) phonetic context and position within the word are important factors, since learners first master rhotics in word-medial intervocalic position,

[19] Waltmunson's (2005) study did not seek to investigate developmental stages.
[20] Waltmunson (2005) uses the term "stripes" to refer to this and explains it as "the near absence of visibility of 2 or more glottal pulses on the spectrogram" (p. 110).

and only move onto coda positions last (Colantoni & Steele, 2008; Face, 2006; Johnson, 2008; Olsen, 2012; Waltmunson, 2005); and (3) learners do not master all acoustic parameters of rhotics simultaneously, resulting in a discrepancy between learners' and controls' productions (Amengual, 2016; Colantoni & Steele, 2008; Johnson, 2008; Waltmunson, 2005), and in fact target different parameters in different contexts (Colantoni & Steele, 2008).

2.5 Chapter summary

This chapter reviewed several studies that have approached the two learning problems relevant to this study: (1) acquiring two L2 sounds that have different statuses in the L1 and the L2 (i.e., phonemic versus allophonic status); (2) acquiring L2 contrasts when one or both sounds are absent from the L1. Several factors that have been reported to affect L2 phonology were also discussed. These factors included the articulatory difficulty of trill production, as well as the effects of position within the word, task formality, and orthography. The discussion then turned to the phonetic, articulatory and phonological descriptions that have been established for the three languages under investigation (Romanian, French, and Spanish) in order to provide a solid understanding of the ways in which the rhotic in these languages are different, which may explain the difficulty of L1 Romanian and L1 French speakers when acquiring Spanish rhotics. This chapter ended with a discussion of the pertinent research regarding the acquisition of Spanish rhotics by L1 English speakers. This section considered the various factors that may affect L2 speech, exemplifying that learners show evidence of L1 transfer when producing rhotics in the L2, that they do not target rhotics in all positions simultaneously, and that orthography and the

type of task may affect their L2 rhotic production. The following chapter turns to the methodology of this study.

Chapter 3

Methodology

3 Overview

The present chapter outlines the methodology of this study and is structured as follows: the research questions and their corresponding hypotheses are presented in Section 3.1. Next, the participants that took part in the study are described in Section 3.2. In Sections 3.3 and 3.4, respectively, I provide an overview of the picture description and sentence reading tasks in the three languages. Sections 3.5 outlines the testing protocol, and, lastly, Section 3.6 provides an overview of the chapter.

To test the hypotheses presented in Section 3.1, 10 intermediate French- and 10 Romanian-speaking learners of Spanish completed four tasks. All of the learners completed two tasks in Spanish, followed by similar tasks in their respective L1s. Completing the tasks in the both the L1 and the L2 was especially relevant for the L1 Romanian group, since, as mentioned in the previous chapter, very little is known about rhotics in Romanian. For the L1 French group, the purpose of completing the tasks in French was to ensure that no dialectal variants other than the uvular fricative were produced. The first task was a picture description task, the second a sentence reading task. These tasks had the following goals. The former task sought to investigate the production of the two Spanish rhotics word-initially and word-medially, in a more spontaneous manner in order to investigate the effect of task formality, while also aiming to minimize the effect of orthography. The sentence reading task aimed to elicit rhotic productions in a more controlled manner, while also

exposing learners to orthography. The controlled nature of these tasks guaranteed that all participants produced the same target words, and the same number of stimuli. In the following section, I outline the three research questions and their corresponding hypotheses.

3.1 Research Questions and Hypotheses

The three research questions guiding this study, as well as their corresponding hypotheses are presented below. The hypotheses are based on the previous literature presented in Chapter 2.

Research Question 1: What differences are expected between the L1 Romanian and L1 French learner groups regarding the L2 production of the Spanish rhotics? If there are differences, are these related to the speakers' L1 rhotics?

Hypothesis 1: Based on the articulatory characteristics and distribution of rhotics within the word in the two L1s, some differences are expected between the two learner groups.

L1 French group: Overall, it will be less difficult for the French speakers to produce the tap than the trill, due to articulatory factors. Specifically, previous research has shown that trill production generally lags behind tap production (Face, 2006; Johnson, 2008; Waltmunson, 2005), partly due to articulatory difficulty (Ladefoged & Maddieson, 1996; Solé, 2002), resulting in the overgeneralization of the tap. In addition to the overproduction of the tap, some transfer errors are also expected (Major, 1986), where French speakers will produce their L1 uvular rhotic instead of trills.

L1 Romanian group: Overall, these learners are expected to have more target trill productions than the L1 French group, since trills occur in their L1 and would therefore not

have difficulties with the articulatory routines required to produce these sounds. Despite being able to produce trills, however, these speakers are expected nonetheless to overgeneralize taps to trill contexts, since taps are more prevalent than trills in the L1 (Avram, 1993; Chitoran, 2002; Radu, 2015; Savu, 2012). This prediction will be further nuanced below, where position in the word is considered.

Research Question 2: Will learners realize trills in word-initial (WI) and word-medial (WM) intervocalic positions with equal accuracy?

Hypothesis 2: Generally, learners are not expected to realize trills in both positions equally well.

L1 French group: This group is expected to have lower rates of trill production word-initially than word-medially, since intervocalic position favours trilling more than word-initial position (Johnson, 2008; Lewis, 2004; Major, 1986). As mentioned in Chapter 2, trills require very precise aerodynamic constraints to be met (Solé, 2002) and any deviations may result in failure to produce trills (Ladefoged & Maddieson, 1996). Therefore, any sort of disturbance in the air flow, such as having to 'start from scratch' in word-initial position could result in not producing trills successfully.

L1 Romanian group: Since trills do occur in Romanian, this group of learners will not have difficulty with trills from an articulatory point of view. Rather, the L1 distribution of trills will affect their L2 production. Recall that trills predominantly occur word-initially in Romanian. As such, this group is expected to target trills word-initially before word-medially, a result of positive transfer (Major, 1986) from their L1.

Research Question 3: Are there differences in the learners' rhotic accuracy in more versus less formal speech, specifically between a sentence reading and a picture description task?

Hypothesis 3: Both groups are expected to behave similarly across the two tasks. Overall, more target-like rhotic productions are expected in the sentence reading than in the picture description task, since transfer errors are more likely in more casual tasks (Major, 1982; 1986; Tarone 1979). In addition to increase in formality, the sentence reading task also involves written presentation of stimuli, which may lead to more target-like speech in the sentence reading versus picture description task due to reduced processing demands (Colantoni, et al., 2015, p. 110). This is expected to be especially true in word-medial position, where the orthography may additionally facilitate rhotic accuracy, since in this position, orthography makes the Spanish tap-trill contrast clear, and this is a phonemic contrast that does not exist in either learner group's L1 (Waltmunson, 2005, p. 37). In word-initial position, the orthography is opaque and may be hindering, especially if learners map the word-initial <r> in Spanish to their L1 rhotics (Young-Scholten & Langer, 2015).

3.2 Participants

3.2.1 Speaker profiles

In total, 33 participants aged 18-58 took part in the study. They were divided into the following groups, according to their L1: (i) 13 L1 Spanish speakers (the control group); (ii) 10 intermediate L1 Romanian-L2 Spanish speakers; (iii) 10 intermediate L1-French-L2 Spanish speakers. Subjects were recruited via posters, social media and email

correspondence and received $15CDN for their participation. In the case of the Romanian speakers, this was converted to RON, the national currency.

The 13 participants in the control group were required to be native Spanish speakers, but not necessarily monolingual, since they were tested in Toronto where bi/multilingualism is widespread. However, all participants reported using Spanish on a regular basis in various situations (e.g., in social situations and at home) and were not near-native speakers of any other languages, including English, in order to limit the effect that their L2 would have on their L1. Moreover, speakers of certain Caribbean Spanish varieties (namely parts of Cuba, the Dominican Republic, and Puerto Rico) were excluded from the study, since rhotics in this dialect are affected by various processes. The native Spanish speakers, who were tested on the University of Toronto campus, were from the following Spanish-speaking countries: Colombia (n=3); Chile, Mexico, and Venezuela (n=2 each); Argentina, El Salvador, Spain, and Uruguay (n=1 each). They ranged in age from 18 to 46 years old, with a mean age of 28.9 years old (SD=7.8). With regard to education, nine of the participants had completed college or university and 4 had completed high school. Nine of the participants were female, while four were male.

The L1 Romanian-L2 Spanish speakers were 10 female participants born and raised in Romania. The age range for this group of speakers was 19-35, with a mean age of 23 (SD=4.57). All were students at the University of Bucharest at the time of testing or had studied at the University in recent years. Since most of the participants were students in the Faculty of Literature and Languages, they were required to choose a minimum of two languages to study. As such, in addition to Spanish, they also had exposure to other

languages such as English (n=9), French (three participants), Hebrew (n=2), Bulgarian, German, and Russian (n=1 each). Their knowledge of the other foreign languages mentioned above ranged from beginner to near-native. However, the additional languages did not influence their production of Spanish rhotics at the time of testing. For example, none of the participants produced uvular rhotics characteristic of German (Jaskuła, 2018) or Hebrew (Cohen & Ben-David, 2016).

Lastly, the L1 French-L2 Spanish speakers were tested in Montreal or on the University of Toronto campus. The age range for this group was 21-58, with a mean age of 30 (SD=12). Nine of the 10 participants were born in France, and the other in Quebec. All participants produced uvular fricative rhotics in their L1 French. While efforts were made to find participants with only Spanish as a foreign language, this was not always possible. All the participants reported having studied English, and two participants reported having very beginner proficiency in Italian, Portuguese or Korean. Regarding their level of education, all participants had either completed or were attending university at the time of testing. Nine of the participants were female and one was male.

Regarding the learner groups, the inclusion criteria in the recruitment posters were the same for the L1 Romanian and L1 French speakers. Firstly, participants had to be at least 18 years old at the time of data collection. Secondly, they had to speak Spanish as a second language at a self-assessed intermediate level. This criterion was included to decrease the possibility of beginner or very advanced learners wanting to participate in the study. Lastly, while one of the exclusionary criteria was for participants to have little to no knowledge of other foreign languages, it was not possible to respect it. In the case of the L1

Romanian-L2 Spanish speakers, their university program required them to take courses in an additional foreign language. Moreover, most of them had at least some knowledge of English, either through formal study or exposure in media, such as television and music. As for the L1 French-L2 Spanish speakers, similar to the Romanian group, many of them had some English proficiency (ranging in self-rated overall proficiency from beginner to advanced), either due to having been exposed through media or immersion, since three of the 10 participants were living in Toronto at the time of testing, as they were on exchange for the academic year.

Before completing the tasks, participants completed a detailed language background questionnaire (adapted from Colantoni & Steele, 2007; see Appendix A) in their respective L1s. The general purpose of the questionnaire was to construct a more complete picture of the participants' linguistic background, and the entire questionnaire took approximately 10 minutes to complete. The first section included questions regarding personal information such as year of birth, birthplace, profession, and level of education. The second set of questions focused on participants' language use; specifically, the languages they use at school, at work, at home and in social situations, and the proportion of use of each. The third section elicited information about participants' second languages including the age of onset of acquisition (AoA). Finally, participants also rated their proficiency in Spanish (and in any other foreign languages) in the following areas: speaking, reading, writing, listening and overall proficiency. The overall proficiency self-ratings were transformed into a numerical score ranging from 1-4, such that 1 corresponds to beginner, 2 to intermediate, 3 to advanced, and 4 to near-native.

Table 2 provides a summary of the Romanian and French participants' biographic information. Although the age ranges for the Romanian (median = 21.6) and French (median = 22.5) participants are quite different, a Mann-Whitney U Test[21] did not reveal a significant difference between the groups, W = 60, $p = 0.468$. As can be seen, the average AoA for the L1 Romanian speakers was 13.8 (SD = 6.76), and 13.4 (SD = 1) for the L1 French speakers. An independent samples *t*-test revealed that there is no significant difference between the groups, $t(9.3672) = -0.18516$, $p = 0.857$. Participants also reported how many hours per week they dedicate to speaking, listening to, and writing in Spanish. The L1 Romanian-L2 Spanish (median = 15.5) group reported a significantly higher number of hours weekly than the L1 French-L2 Spanish group (median = 4.5), W = 13.5, $p = 0.006$. Lastly, both learner groups had an average overall self-rating of 2.3 (Romanian speakers' SD = 0.48; French speakers' SD = 0.4). A Mann-Whitney U Test did not reveal a significant difference between the L1 Romanian (median = 2) and L1 French group (median = 2), W = 53, $p = 0.893$.

Table 2

L1 Romanian-L2 Spanish and L1 French-L2 Spanish Biographical Information

	L1 Romanian-L2 Spanish			L1 French-L2 Spanish		
	Mean	SD	Range	Mean	SD	Range
Age	23	4.57	19-35	30	12	21-58
AoA (in years)	13.8	6.76	8-27	13.4	1	13-14
Spanish use (hrs/week)	18.6	17.61	2-66	5.34	5.43	0-15.25
Self-assessment of ability	2.3	0.48	1-4	2.3	0.4	1-4

[21] All of the statistics presented were run in R Core Team (2013).

3.2.2 L2 Spanish competence

The learners' L2 Spanish competence was determined using two methods: an accentedness rating and a cloze test. While the focus of this study is oral proficiency, a cloze test was also administered to determine whether the two learner groups' lexical and grammatical proficiencies were comparable. This was primarily relevant in order to ensure that both groups of speakers would be able to complete the two tasks in this study. Moreover, research has shown that there is a correlation between vocabulary size and L2 production accuracy (e.g., Bundgaard-Nielsen, Best, Kroos, & Tyler, 2012 for L2 vowel intelligibility).

Though the first step towards determining participants' proficiency was through self-ratings, the focus of the present study is phonetic and phonological proficiency, which does not always correlate with participants' self-ratings; thus, an additional measure was used. As part of the testing session, participants were asked to read the Spanish version of *The Northwind and the Sun* (*El viento del norte y el sol*, Appendix B). After testing, the learners' readings of the text were intermixed and randomized with those of the native Spanish speakers. These recordings were subsequently presented via PowerPoint presentations to a panel of five native Spanish speakers, none of whom had any formal training in linguistics nor were aware of the objectives of the study. The order of presentation of recordings was randomized for each judge. The judges' task was to listen to each recording and assign it an accentedness score ranging from 1 (*clearly nonnative; very strong foreign accent*) to 5 (*clearly native; no foreign accent*) (as described in Bongaerts, Mennen & van der Silk, 2000). Mean ratings were then calculated for each learner and are displayed in Table 3, along with the individual scores.

Table 3

Native Speaker Judges' Evaluation of L1 Romanian- and L1 French-L2 Spanish Speakers' Oral Proficiency Based on their Reading of the Northwind and the Sun *Text*

	Judges' Scores							Judges' Scores					
Participant	1	2	3	4	5	Average	Participant	1	2	3	4	5	Average
RS09	2	3	2	1	2	2	FS01	1	1	2	1	2	1.4
RS06	2	1	4	2	3	2.4	FS08	1	1	3	1	2	1.4
RS12	1	3	3	2	3	2.4	FS10	1	2	1	1	2	1.4
RS11	1	4	3	3	3	2.8	FS14	1	1	2	1	2	1.4
RS07	2	3	4	3	3	3	FS11	1	4	1	1	1	1.6
RS10	3	4	3	2	3	3	FS15	1	1	3	1	2	1.6
RS04	2	4	4	3	3	3.2	FS03	1	1	2	3	2	1.8
RS05	2	4	4	3	3	3.2	FS04	1	3	2	1	2	1.8
RS01	3	4	4	3	3	3.4	FS09	1	3	2	2	2	2
RS08	3	5	4	4	3	3.8	FS12	3	4	4	1	3	3
Group						**2.92**	Group						**1.74**

As can be observed in Table 3, the L1 Romanian speakers received a higher average accentedness score than the L1 French speakers (2.92 versus 1.74). Because a Shapiro-Wilk test revealed that the data for the L1 French group were not normally distributed ($p = 0.016$; $W = 0.722$), a Mann-Whitney U Test was conducted to compare the median oral proficiency scores of the two language groups. The test revealed a significant difference between the L1 Romanian (median=3) and L1 French group (median = 1.6) ($W = 5.5$, $p < 0.001$ ***). In order to determine whether Average Accentedness Ratings were affected by the learners' specific L1, a simple linear regression was calculated; the results of the model are summarized in Table 4. As can be seen, the L1 Romanian group received significantly

higher accentedness ratings than the L1 French group. These findings indicate that the L1 has a significant effect on Average Accentedness Rating.

Table 4

Results of a Simple Linear Regression with Average Accentedness Rating as the Dependent Variable and L1 as the Independent Variable

	Estimate	SE	t value	Pr(>|t|)
Intercept	1.7400	0.1623	10.723	$p < 0.001$ ***
LanguageRomanian	1.1800	0.2295	5.142	$p < 0.001$ ***

This is not an unexpected result, since the overall French and Spanish sound systems and phonological processes are more different than those of Romanian and Spanish, which may have implications for the accentedness scores. That is, due to greater phonological similarity between Romanian and Spanish compared to French and Spanish, it could be the case that, even beginner L1 Romanian learners of Spanish would receive higher accentedness scores than intermediate L1 French learners of Spanish. Regarding French and Spanish, in addition to the two languages having very different rhotics, Standard French contrasts up to 12 oral vowels and four nasal vowels (Féry, 2003), while Spanish has only five oral vowels (Hualde, 2005a). French also has liaison, a phonological process occurring at word boundaries involving consonant epenthesis and resyllabification (Féry, 2003), which does not exist in Spanish. Perhaps one of the most salient differences between French and Spanish is regarding their stress systems. While Spanish is a syllable-timed language with stress being realized at the level of the word, prominence in French is realized at the Accentual Phrase level (Jun & Fougeron, 2002). On the other hand, Spanish and Romanian have many aspects in common. For example, though Romanian has a slightly larger number of vowels than Spanish (seven versus five), like Spanish, the Romanian vocalic system does

not contrast nasal and oral vowels (Seiciuc, 2014), and both languages share similar diphthongs of rising sonority as a result of the "breaking" of specific Latin stressed mid vowels (Chitoran & Hualde, 2007). In addition, the stress patterns are similar to those of Spanish (Chitoran, 2002, p. 81). Overall, the differences between French and Spanish may be a source of transfer from speakers' L1 French into their L2 Spanish, resulting in lower oral proficiency scores when compared to the L1 Romanian speakers, whose L1 is more similar to Spanish.

In addition to the oral proficiency test, the two learner groups also completed a multiple-choice cloze test in Spanish. Although the concern with cloze tests is that they do not measure oral proficiency, using both a cloze test and accentedness ratings, as were used in this study, provides a more complete, global picture of L2 proficiency (Tremblay & Garrison, 2010). The cloze test used was a condensed subset of the DELE (Los Diplomas de Español como Lengua Extranjera [Diploma of Spanish as a Second Language]) previously used in Montrul & Slabakova (2003), Montrul (2005), and Montrul (2010b) (Appendix C). A complete version of the test can be found at http://international.ucla.edu/nhlrc/data/example. For this test, participants had to provide 20 missing words in a text by choosing one of three possible choices per blank. Table 5 displays each participant's score out of 20 total points.

Table 5

L1 Romanian and L1 French Speakers' Cloze Test Scores

Romanian group		French group	
Participant	Score /20	Participant	Score /20
RS04	11	FS08	10
RS09	11	FS09	10
RS05	12	FS01	12
RS10	13	FS04	12
RS06	14	FS15	12
RS12	14	FS12	13
RS11	15	FS03	14
RS07	16	FS14	14
RS01	18	FS11	15
RS08	18	FS10	16
Group average	**14.2**	Group average	**12.8**

Regarding the cloze test results, according to an independent-samples *t*-test, there was no significant difference between the L1 Romanian (M = 14.2, SD = 2.57) and the L1 French group (M = 12.8, SD = 2), $t(16.924) = 1.3612$, $p = 0.191$. As such, we can conclude that the two learner groups are of comparable proficiency levels, at least in what concerns their lexical and grammatical proficiency.

As can be seen in Tables 4 and 5 tables, oral proficiency does not always correlate with cloze test results. For example, while these two proficiency measures are comparable for participant RS08, who received the overall highest oral proficiency and cloze test scores of the Romanian group (oral proficiency score: 3.8/5; cloze test score: 18/20), they are not similar for participant RS09, who received the lowest oral proficiency score (2/5), but had the same cloze test score as RS04 (i.e., 11/20), whose oral proficiency score was higher than

RS09's (3.2/5). Similarly, while FS12 received the highest oral proficiency score (3/5), this participant got a cloze test score of 13/20, which was not the highest of the group. In fact, participant FS10 received the highest cloze test score (15/20), but one of the lowest oral proficiency scores (1.4/5). In order to explore the possibility of a correlation between the accentedness scores and cloze test results, Figure 8 depicts a scatterplot of this data. A Kendall's Tau Rank Correlation was computed (Levshina, 2015, pp. 132-133) in order to test the strength of the correlation of accentedness ratings and cloze tests results for the 20 participants. The results of the Kendall Tau test did not reveal a significant correlation between the two proficiency scores ($\tau = 0.1839202$; $n = 20$, $p = 0.289$), suggesting that oral proficiency and grammatical or lexical proficiency do not correlate.

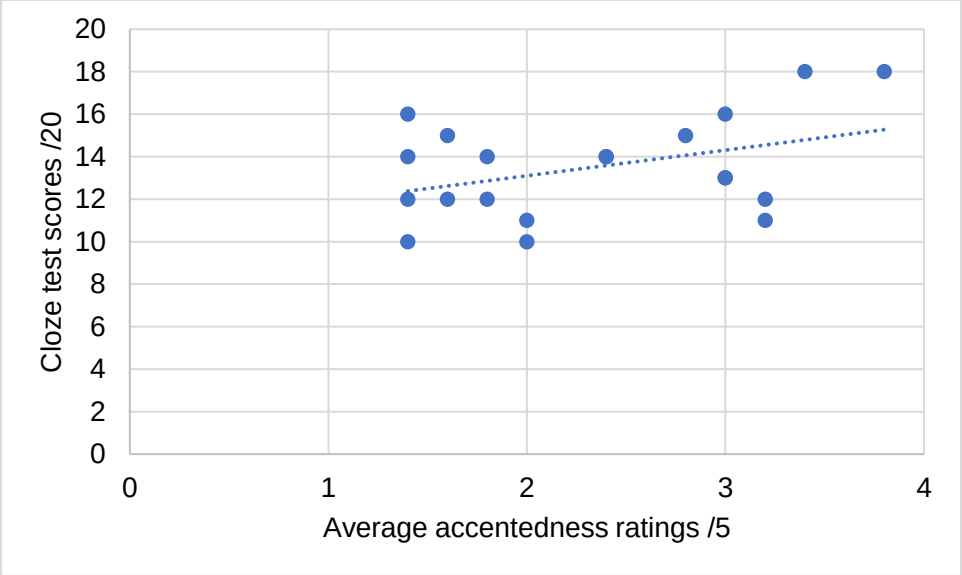

Figure 8. Scatterplot of Cloze Test Scores (y-axis) and Average Accentedness Scores (x-axis) for the 10 L1 Romanian-L2 Spanish and 10 L1 French-L2 Spanish Speakers.

Lastly, a series of simple linear effects models were run in order to investigate the if either AoA or Spanish use (hours/week) had any effect on the three proficiency measures (Self-Assessment of Ability, Cloze Test Scores, Average Accentedness Ratings), in order to get a better sense of whether those participants who spent more time using Spanish, or who had learned Spanish at an earlier age, would have higher scores for the proficiency measures. Two of these models were significant, namely the effect of Spanish use on the Cloze Test Scores, and the effect of Spanish use on Average Accentedness Scores. The results of these models are presented in Tables 6 and 7. As can be seen in the tables, the amount of Spanish use (hours/week) significantly affected two of the proficiency scores. Specifically, those participants who dedicated more time to using Spanish, received the highest Cloze Test Scores (Table 6), as well as the highest Average Accentedness Scores (Table 7).

Table 6

Results of a Simple Linear Regression with Cloze Test Scores as the Dependent Variable and Spanish use (hours/week) as the Independent Variable

	Estimate	SE	t value	Pr(>|t|)
Intercept	12.4516	0.60126	20.717	$p < 0.001$ ***
Spanish use	0.08708	0.03259	2.672	$p = 0.016$ *

Table 7

Results of a Simple Linear Regression with Average Accentedness Scores as the Dependent Variable and Spanish use (hours/week) as the Independent Variable

	Estimate	SE	t value	Pr(>|t|)
Intercept	1.94453	0.19162	10.148	$p < 0.001$ ***
Spanish use	0.03216	0.01039	3.096	$p = 0.006$ **

3.3 Picture description task

The first experimental task was a picture description task, which aimed to elicit the semi-spontaneous productions of both rhotics, since task formality may affect pronunciation. There is also the need to study rhotics in more naturalistic speech, as this is a hypoarticulated environment where phonetic production constraints play a significant role (Colantoni & Steele, 2008, p.523). Moreover, the goal was to elicit the target stimuli without the use of spelling cues, since English, French, and Spanish rhotics are all represented with the grapheme <r>, which could affect production (Colantoni & Steele, 2006; Face, 2006).

In addition to completing the picture description task in Spanish, the learners also completed a similar task in their respective L1s. This was done for two reasons: (1) since the frequency and distribution of the trill is still not clear in Romanian, it is necessary to collect as much experimental data on this as possible in order to contribute to our understanding of the sound systems of less well-studied languages; and (2) because the first research question asks whether the participants' L1 rhotic production influences their L2 Spanish production, it is important to have an understanding of each group's L1 rhotic production. However, since, as we will see, more is known about French rhotics than Romanian ones, the L1 French data was collected only to ensure that participants were indeed producing the expected uvular rhotic in their L1 French.

The task design was similar for the picture description task across all three languages. Participants were presented with one picture at a time on a computer screen via a PowerPoint presentation and were instructed to produce simple, complete sentences in order to describe each picture. Each image was a scene containing, among other images, an

image depicting a target word with either a word-initial or word-medial rhotic (e.g., *rosa* /rosa/ ‘rose’; *caro* /kaɾo/ ‘expensive’ and *carro* /karo/ ‘car’ in the Spanish task; *rege* /ɾedʒe/ ‘king’ and *mare* /maɾe/ ‘sea’ in the Romanian task; *râteau* /ʁa.to/ ‘rake’ and carré /ka.ʁe/ ‘square’ in the French task). For example, in the image depicting the target word *caro* ‘expensive’, participants saw a woman looking shocked as she checked the price of a shirt. Before beginning the actual task, participants were shown an example of an unrelated picture (not containing a rhotic) in order to practice eliciting complete sentences. In the few cases where participants did not produce the target words, they were asked questions related to the picture that would trigger the use of these words. In addition to being instructed to produce complete sentences, when completing the task in their L1, Romanian participants were also told to produce only words with indefinite articles (*un* masculine ‘a/an’; *o* feminine ‘a/an’). This is because, in Romanian, definite articles are bound morphemes, and are attached to the end of nouns. In some cases, using the definite article with some of the stimuli would result in additional syllables, changing the shape of the target bisyllabic CVCV words. For example, the word *rege* /ɾedʒe/ ‘king’ is bisyllabic and remains so when preceded by an indefinite article (*un rege* [un.ɾe.dʒe]). However, using the definite article -*le* yields the trisyllabic form *regele* ([ɾe.dʒe.le]), but the position of stress remains the same (i.e., on the first syllable of the word). As in the Spanish task, no distracters were included in either the Romanian or French tasks, since the images did not depict only target the words.

3.3.1 Spanish stimuli in the picture description task

Nineteen pictures depicting 19 Spanish words involving a word-initial or –medial rhotic were included in the picture-naming task. No distracters were included in this task, since

the nature of the picture description task helped conceal the objective of the study – as noted above, because the images used depicted not only the target word, but also other objects, participants produced several words, both with and without rhotics. In order to control several factors (word structure and shape, surrounding sounds), all of the target words were bisyllabic (CVCV), with syllable-initial stress. The vowels flanking <r> were always low or mid (i.e., /e a o/). Table 8 presents the complete set of stimuli. As can be seen, the words containing word-medial /ɾ/ and /r/ were (near)minimal pairs.

Table 8

Spanish Stimuli Set for the Picture Description Task

Word-initial (/r/ = 7)				**Word-medial (/r/ = 6; /ɾ/ = 6)**		
Sound	Orthography	IPA	Gloss	Orthography	IPA	Gloss
/r/	rata	/'ra.ta/	‘rat’	corro	/'ko.ro/	‘circle/ring’
/r/	roca	/'ro.ka/	‘rock’	cerro	/'se.ro[22]/	‘hill’
/r/	ropa	/'ro.pa/	‘clothing’	parra	/'pa.ra/	‘grapevine’
/r/	rosa	/'ro.sa/	‘rose’	carro	/'ka.ro/	‘car’
/r/	rojo	/'ro.xo/	‘red’	perra	/'pe.ra/	‘female dog’
/r/	reno	/'re.no/	‘reindeer’	zorro	/'so.ro/	‘fox’
/r/	rabo	/'ra.bo/	‘tail’			
/ɾ/	-	-	-	coro	/'ko.ɾo/	‘choir’
/ɾ/	-	-	-	cero	/'se.ɾo/	‘zero’
/ɾ/	-	-	-	cara	/'ka.ɾa/	‘face’
/ɾ/	-	-	-	caro	/'ka.ɾo/	‘expensive’
/ɾ/	-	-	-	pera	/'pe.ɾa/	‘pear’
/ɾ/	-	-	-	toro	/'to.ɾo/	‘bull’

3.3.2 French stimuli in the picture description task

The target stimuli in the French picture description task appear in Table 9 The 16 stimuli were bisyllabic words, with varying stress depending on the sentences produced by each

[22] I am adopting a Latin American pronunciation for the transcriptions involving <ce> and <ci>; that is, the transcription uses /s/ and not the interdental fricative, which is used in Spain.

participant. This is because as opposed to both Romanian and Spanish, French stress is more accurately referred to as "phrasal prominence", since it is assigned at the level of the Accentual Phrase rather than at the level of the word (Jun & Fougeron, 2002). All of the stimuli had the CVCV structure and <r> was flanked by the low and mid vowels /a, ɑ, ə, ɛ, ɔ, e, ɔ̃, o, ɑ̃, ɛ̃/. Eight of the words had <r> word-initially and the other eight word-medially, intervocalically. All of the stimuli were checked by three L1 French speakers.

Table 9

French Stimuli Set in the Picture Description Task

Word-initial (/ʁ/ = 8)				**Word-medial (/ʁ/ = 8)**		
Sound	Orthography	IPA	Gloss	Orthography	IPA	Gloss
/ʁ/	râteau	/ʁa.'to/	'rake'	carré	/ka.'ʁe/	'square'
/ʁ/	repas	/ʁə.'pa/	'meal'	marron	/ma.ʁɔ̃/	'brown'
/ʁ/	réchaud	/ʁe.'ʃo/	'camping stove'	barré	/ba.'ʁe/	'locked'
/ʁ/	radeau	/ʁa.'do/	'raft'	carreau	/ka.'ʁo/	'tile/pane'
/ʁ/	rabais	/ʁa.'bɛ/	'discount'	béret	/be.'ʁɛ/	'beret'
/ʁ/	rabbin	/ʁa.'bɛ̃/	'rabbi'	terrain	/te.'ʁɛ̃/	'field'
/ʁ/	ragoût	/ʁa.'gu/	'stew'	forêt	/fɔ.'ʁɛ/	'forest'
/ʁ/	raisin	/ʁe.'zɛ̃/	'grape'	parents	/pa.'ʁɑ̃/	'parents'

3.3.3 Romanian stimuli in the picture description task

The target stimuli in the Romanian picture description task are given in Table 10, below. The 16 stimuli were trochaic bisyllabic words in which the target segment <r> was flanked by the low and mid vowels /e, a, ə, o/. Eight of the words contained <r> word-initially, the other eight a word-medial, intervocalic <r>. All of the stimuli were checked by three L1 Romanian speakers to ensure that the words were commonly used in Romanian, and the nonce words could exist and were natural-sounding.

Table 10

Romanian Stimuli Set in the Picture Description Task

Word-initial (/ɾ/ = 8)				**Word-medial (/ɾ/ = 8)**		
Sound	Orthography	IPA	Gloss	Orthography	IPA	Gloss
/ɾ/	ramă	/'ɾa.mə/	‘frame’	sare	/'sa.ɾe/	‘salt’
/ɾ/	raţă	/'ɾa.tsə/	‘duck’	mare	/'ma.ɾe/	‘sea’
/ɾ/	rege	/'ɾe.dʒe/	‘king’	bere	/'be.ɾe/	‘beer’
/ɾ/	rază	/'ɾa.zə/	‘ray’	bară	/'ba.ɾə/	‘bar’
/ɾ/	rană	/'ɾa.nə/	‘wound’	gară	/'ga.ɾə/	‘railway station’
/ɾ/	rade	/'ɾa.de/	‘shave’ (3rd pp, fg, pres)	mere	/'me.ɾe/	‘apples’
/ɾ/	rasă	/'ɾa.sə/	‘breed’	pară	/'pa.ɾə/	‘pear’
/ɾ/	rece	/'ɾe.tʃe/	‘cold’	seră	/'se.ɾə/	‘greenhouse’

3.4 Sentence reading task

While the picture description task allowed for semi-spontaneous productions of rhotics, the purpose of the sentence reading task was to ensure that the data collected would be comparable across participants, since all of speakers would have a chance to produce the same stimuli. Unlike the picture description, the sentence reading task provided orthography. Moreover, in addition to the real words employed in the first task, the sentence reading task also included nonce words in an attempt to minimize the effect that L1-L2 cognates may have on production.

Similar to the picture description task, in addition to completing the sentence reading task in their L2, the learner groups also had to complete a similar task in their respective L1s. The purpose of this was to collect data from a more formal task, which also included orthography. This was important for two reasons: (1) there is a need for a more

comprehensive understanding of Romanian rhotics, including their production in more formal contexts, which may lead to hyperarticulated speech (Lindblom, 1990), and (2) a secondary goal was to compare data as similar as possible. As previously mentioned, the picture description task allowed for the semi-spontaneous productions of rhotics, making it more difficult to control the tokens that participants would actually produce. The sentence reading task thus aimed to elicit L1 data that would be comparable across participants, since they all had the opportunity to produce the same stimuli and number of tokens.

The sentence reading task design was similar across all three languages. On a computer screen, using PowerPoint, participants were presented with the target words embedded in a carrier phrase:

Spanish: *Digan [TARGET] otra vez* 'Say [TARGET] again'.

French: *Je dis [TARGET] de nouveau* 'I say [TARGET] again'.

Romanian: *Spun [TARGET] încă o dată* 'I/they say [TARGET] again'.

Their task was to read each sentence out loud. For example, participants read aloud sentences such as *Digan rojo otra vez* 'They say red again'(Spanish); *Spun rege încă o dată* 'I/they[23] say king again' (Romanian); *Je dis marron de nouveau* 'I say chestnut again' (French). The Spanish task had 50 sentences in total (30 target stimuli and 20 distracters), while the French and Romanian tasks each had 52 sentences in total (32 target stimuli and 20 distracters). In each of the tasks, the target stimuli and distracters were presented in

[23] The verb *spun*, from the infinitive *a spune* 'to say', has the same form for both the first person singular and third person plural; since Romanian is a pro-drop language, the verb conjugation is ambiguous when the subject is not overtly stated.

random order, which was determined using the website www.random.org. Additionally, halfway through the PowerPoint presentation, participants saw a slide which read *Sólo te hace falta la mitad* (Spanish), *Am ajuns la jumătatea experimentului* (Romanian), *On a complété la moitié de la tâche* (French) 'You are halfway through the expeiremnt', in order to give them an idea of how much of the task they had completed and to maintain their focus. At the beginning of the experiment, participants were told that some of the words were not real Spanish words. To minimize redundancy and maintain participants' focus throughout the task, each slide was a different colour.

3.4.1 Spanish stimuli in the sentence reading task

The target stimuli in the sentence reading task included all of the words presented in Table 6 in Section 3.3.1, except for the following: *rabo* 'tail', *reno* 'reindeer', *zorro* 'fox' and *toro* 'bull'. These words were only included in the picture description task as extras, in the event that the learners would not be able to produce any of the other 15 stimuli due to being unfamiliar with the vocabulary. As such, the sentence reading task included 15 real Spanish words: five containing word-initial <r>, five containing word-medial <r>, and five containing word-medial <rr>. In addition to these 15 words, 15 nonce words were also included in this task. As with the real words, five nonce words contained word-initial <r>, five had word-medial <r>, and the other five had word-medial <rr>. The main reason for using nonce words was to minimize the effect of cognates; cognates were difficult to eliminate, especially since the stimuli were restricted to CVCV with low and mid flanking vowels. The nonce words were not real words in any of the two languages participants were tested. Like the real words, they were all bisyllabic, with word-initial stress, and the vowels

flanking the rhotics were the low and mid vowels /a e o/. In addition, the task also included 20 distracters, a third of the target stimuli set. The distracters included the following: five words containing /j/ (three of these were nonce words); five monosyllabic words (three of these were nonce words); five words containing the cluster /bl/ (three were nonce words); and five words containing the cluster /kl/ (three were nonce words). All of the stimuli were checked by three native Spanish speakers. Tables 11 and 12 provide a summary of the real and nonce stimuli, respectively, while Table 13 is a summary of the distracters.

Table 11

Spanish Stimuli Set in the Sentence Reading Task (Real Words)

Word-initial (/r/ = 5)				**Word-medial (/r/ = 5; /ɾ/ = 5)**		
Sound	Orthography	IPA	Gloss	Orthography	IPA	Gloss
/r/	rata	/'ra.ta/	'rat'	corro	/'ko.ro/	'circle/ring'
/r/	roca	/'ro.ka/	'rock'	cerro	/'se.ro/	'hill'
/r/	ropa	/'ro.pa/	'clothing'	parra	/'pa.ra/	'grapevine'
/r/	rosa	/'ro.sa/	'rose'	carro	/'ka.ro/	'car'
/r/	rojo	/'ro.xo/	'red'	perra	/'pe.ra/	'female dog'
/ɾ/	-	-	-	coro	/'ko.ɾo/	'choir'
/ɾ/	-	-	-	cero	/'se.ɾo/	'zero'
/ɾ/	-	-	-	cara	/'ka.ɾa/	'face'
/ɾ/	-	-	-	caro	/'ka.ɾo/	'expensive'
/ɾ/	-	-	-	pera	/'pe.ɾa/	'pear'

Table 12

Spanish Stimuli Set in the Sentence Reading Task (Nonce Words)

Word-initial (/r/ = 5)				**Word-medial (/r/ = 5; /ɾ/ = 5)**		
Sound	Orthography	IPA	Gloss	Orthography	IPA	Gloss
/r/	rafo	/'ra.fo/	nonce	porre	/'po.re/	nonce
/r/	rebe	/'re.be/	nonce	larre	/'la.re/	nonce
/r/	reso[24]	/'re.so/	nonce	darro	/'da.ro/	nonce
/r/	reva	/'ra.ba/	nonce	garro	/'ga.ro/	nonce
/r/	rone	/'ro.ne/	nonce	terro	/'te.ro/	nonce
/ɾ/	-	-	-	pore	/'po.ɾe/	nonce
/ɾ/	-	-	-	lare	/'la.ɾe/	nonce

[24] The nonce word *reso* could be interpreted as *rezo* 'I pray', as the two words have the same pronunciation. However, orthographically, they are not the same.

/ɾ/	-	-	-	daro	/'da.ɾo/	nonce
/ɾ/	-	-	-	garo	/'ga.ɾo/	nonce
/ɾ/	-	-	-	tero[25]	/'te.ɾo/	nonce

Table 13

Spanish Distracters in the Sentence Reading Task

Type	Orthography	Gloss
/ʝ/	llanto	'sobbing'
/ʝ/	callado	'quiet'
/ʝ/	retrallar	nonce
/ʝ/	nello	nonce
/ʝ/	lledor	nonce
monosyllabic	pan	'bread'
monosyllabic	sal	'salt'
monosyllabic	len	nonce
monosyllabic	plub	nonce
monosyllabic	gar	nonce
/bl/	tabla	'board/table'
/bl/	blindado	'armoured'
/bl/	nobla	nonce
/bl/	bleta	nonce
/bl/	blecua	nonce
/kl/	claro	'clear'
/kl/	clan	'clan/group'
/kl/	clad	nonce
/kl/	clasón	nonce
/kl/	clogo	nonce

3.4.2 French stimuli in the sentence reading task

The stimuli in the French sentence reading task included all of the words presented in Table 14 – these are the same words depicted in Table 9 in Section 3.3.2. Eight of these words contained the rhotic word-initially, and the other word-medially, intervocalically. In addition to the 16 real words, 16 nonce words were included as well. Like the real words, eight of the nonce words contained <r> word-initially and the other eight contained <r>

[25] In Argentine Spanish, this word denotes a type of bird.

word-medially, intervocalically. As previously mentioned, the nonce words (see Table 15) were included in an attempt to minimize the effect of cognates, since the three languages under investigation share many lexical items. The nonce words were not real words in either French or Spanish. Like the real words, the nonce words were also bisyllabic and CVCV in structure. <r> was flanked by the low and mid vowels /a, ɑ, ə, ɛ, ɔ, e, ɔ̃, o, ɑ̃, ɛ̃/. In addition to the 32 target stimuli, 20 distracters were also included in this task. The distracters, shown in Table 16, were structured in the following way: five monosyllabic words (two real and two nonce words), five contained /pl/ word-initially (two real and two nonce words), five were vowel-initial (two real and three nonce words), and five contained /bl/ word-medially (two real and three once words).

Table 14

French Stimuli Set in the Sentence Reading Task (Real Words)

Word-initial (/ʁ/ = 8)				**Word-medial (/ʁ/ = 8)**		
Sound	Orthography	IPA	Gloss	Orthography	IPA	Gloss
/ʁ/	râteau	/ʁa.'to/	‘rake’	carré	/ka.'ʁe/	‘square’
/ʁ/	repas	/ʁə.'pa/	‘meal’	marron	/ma.ʁɔ̃/	‘brown’
/ʁ/	réchaud	/ʁe.'ʃo/	‘camping stove’	barré	/ba.'ʁe/	‘locked’
/ʁ/	radeau	/ʁa.'do/	‘raft’	carreau	/ka.'ʁo/	‘tile/pane’
/ʁ/	rabais	/ʁa.'bɛ/	‘discount’	béret	/be.'ʁɛ/	‘beret’
/ʁ/	rabbin	/ʁa.'bɛ̃/	‘rabbi’	terrain	/te.'ʁɛ̃/	‘field’
/ʁ/	ragoût	/ʁa.'gu/	‘stew’	forêt	/fɔ.'ʁɛ/	‘forest’
/ʁ/	raisin	/ʁe.'zɛ̃/	‘grape’	parents	/pa.'ʁɑ̃/	‘parents’

Table 15

French Stimuli Set in the Sentence Reading Task (Nonce Words)

Word-initial (/ʁ/ = 8)				**Word-medial (/ʁ/ = 8)**		
Sound	Orthography	IPA	Gloss	Orthography	IPA	Gloss
/ʁ/	renat	/ʁə.'na/	nonce	narreau	/na.'ʁo/	nonce
/ʁ/	ranais	/ʁa.'nɛ/	nonce	meras	/mə.'ʁa/	nonce
/ʁ/	rébas	/ʁe.'ba/	nonce	bérat	/be.'ʁa/	nonce
/ʁ/	relas	/ʁə.'la/	nonce	derais	/də.'ʁɛ/	nonce

/ʁ/	rebain	/ʁə.'bɛ̃/	nonce	derat	/də.'ʁa/	nonce
/ʁ/	refas	/ʁə.'fa/	nonce	terras	/te.'ʁa/	nonce
/ʁ/	reveau	/ʁa.'vo/	nonce	lerras	/le.'ʁa/	nonce
/ʁ/	resas	/ʁə.'sɑ/	nonce	zarant	/za.'ʁɑ̃/	nonce

Table 16

French Distracters in the Sentence Reading Task

Type	Orthography	Gloss
monosyllabic	taux	'rate'
monosyllabic	peau	'skin'
monosyllabic	paux	nonce
monosyllabic	teux	nonce
monosyllabic	caux	nonce
/pl/	plusieurs	'several'
/pl/	pleuvoir	'rain'
/pl/	plas	nonce
/pl/	ploux	nonce
/pl/	plaquis	nonce
vowel-initial	amie	'friend (fem)'
vowel-initial	ours	'bear'
vowel-initial	éblé	nonce
vowel-initial	altré	nonce
vowel-initial	ébrier	nonce
/bl/	sable	'sand'
/bl/	table	'table'
/bl/	dable	nonce
/bl/	vable	nonce
/bl/	mable	nonce

3.4.3 Romanian stimuli in the sentence reading task

The stimuli in this task consisted of the 16 words used in the picture description task (Table 10 in Section 3.3.3), outlined in Table 17. In addition, the 16 nonce words in Table 18 were also used; these nonce CVCV words were all stressed on the first syllable. None of the nonce words was real in either Romanian or Spanish. Like the real words, the target segment <r> in the nonce words was flanked by the low and mid vowels /e, a, ə, o/. Eight of the nonce words contained <r> word-initially, and the other eight word-medially,

intervocalically. Unlike the picture description task, this task contained 20 distracters, shown in Table 19: five had the diphthong /o̯a/, five had the cluster /st/, five the cluster /bl/ and five the vowel /a/. In each set of distracters, three of the five were nonce words.

Table 17

Romanian Stimuli Set in the Sentence Reading Task (Real Words)

Word-initial (/ɾ/ = 8)				**Word-medial (/ɾ/ = 8)**		
Sound	Orthography	IPA	Gloss	Orthography	IPA	Gloss
/ɾ/	ramă	/'ɾa.mə/	‘frame’	sare	/'sa.ɾe/	‘salt’
/ɾ/	raţă	/'ɾa.tsə/	‘duck’	mare	/'ma.ɾe/	‘sea’
/ɾ/	rege	/'ɾe.dʒe/	‘king’	bere	/'be.ɾe/	‘beer’
/ɾ/	rază	/'ɾa.zə/	‘ray’	bară	/'ba.ɾə/	‘bar’
/ɾ/	rană	/'ɾa.nə/	‘wound’	gară	/'ga.ɾə/	‘railway station’
/ɾ/	rade	/'ɾa.de/	‘shave (3rd pp, fg, pres)	mere	/'me.ɾe/	‘apples’
/ɾ/	rasă	/'ɾa.sə/	‘breed’	pară	/'pa.ɾə/	‘pear’
/ɾ/	rece	/'ɾe.ʧe/	‘cold’	seră	/'se.ɾə/	‘greenhouse’

Table 18

Romanian Stimuli Set in the Sentence Reading Task (Nonce Words)

Word-initial (/ɾ/ = 8)				**Word-medial (/ɾ/ = 8)**		
Sound	Orthography	IPA	Gloss	Orthography	IPA	Gloss
/ɾ/	renă	/'ɾe.nə/	nonce	fero	/'fe.ɾo/	nonce
/ɾ/	ralo	/'ɾa.lo/	nonce	meră	/'me.ɾə/	nonce
/ɾ/	reba	/'ɾe.ba/	nonce	bera	/'be.ɾa/	nonce
/ɾ/	rela	/'ɾe.la/	nonce	deră	/'de.ɾə/	nonce
/ɾ/	recă	/'ɾe.kə/	nonce	gera	/'ge.ɾa/	nonce
/ɾ/	refa	/'ɾe.fa/	nonce	hera	/'he.ɾa/	nonce
/ɾ/	revo	/'ɾe.bo/	nonce	leră	/'pe.ɾə/	nonce
/ɾ/	repă	/'ɾe.pə/	nonce	zera	/'ze.ɾa/	nonce

Table 19

Romanian Distracters in the Sentence Reading Task

Type	Orthography	Gloss
/o̯a/	broască	'frog'
/o̯a/	poartă	'gate'
/o̯a/	troabă	nonce
/o̯a/	croală	nonce
/o̯a/	floata	nonce
/st/	august	'august'
/st/	trist	'sad'
/st/	vust	nonce
/st/	birist	nonce
/st/	plust	nonce
/a/	afiș	'poster/ad'
/a/	albină	'bee'
/a/	abumă	nonce
/a/	acrună	nonce
/a/	aprotă	nonce
/bl/	bleagă	'loopy (fem)'
/bl/	blând	'gentle (masc)'
/bl/	blună	nonce
/bl/	blemă	nonce
/bl/	blonă	nonce

3.5 Testing protocol

Participants were tested individually in a quiet room. While the recording session for the native Spanish speakers took approximately 25 minutes since they completed only the Spanish tasks, the session for the two learner groups took approximately 45-50 minutes. All participants were first informed orally about the nature of the study. With regard to the objective, they were only told that the interest was to learn more about how leaners acquire Spanish as a second language and were not told their pronunciation would be tested in the study. Prior to each task, participants were given both oral and written instructions. They

were also given an overview of the order in which the tasks would be presented, as well as an estimated amount of time it would take to complete the session.

Participants began the testing session by reading and signing the consent form (Appendix D), which was written in their native language. Participants were informed that they could withdraw from the study at any time without any consequence to them. After indicating their written consent, participants completed the picture description task in Spanish, followed by the sentence reading task. Next, they read the Spanish version of *The Northwind and the Sun* (*El viento del norte y el sol*; Appendix B). After this, participants filled out the language background questionnaire. While in the case of the native Spanish speakers this came at the end of the experiment, the learner groups continued with the tasks in their L1. For the learners, the background questionnaire was completed in the middle of the experiment so as to serve as a break between the two sets of experiments in the two languages. The questionnaire was written in the learners' L1 in order to activate their L1 in preparation for the second part of the experiment. After filling out the background questionnaire, the learner groups completed the picture description, followed by the sentence reading task in their L1 (French or Romanian). Then, the learners read their L1 version of *The Northwind and the Sun – La bise et le soleil* (Appendix E) in the case of French speakers or *Crivăţul şi soarele* (Appendix F) in the case of L1 Romanian speakers. The order of the tasks reflected the reverse order of language acquisition (i.e., their L2 first, followed by their L1) for the participants. Both learner groups ended the testing session by completing the multiple-choice cloze test (Appendix C). When completing the Spanish tasks, all communication was done in Spanish, while all communication was done in French

or Romanian when completing the tasks in the respective L1s. At the end of each testing session, participants were given the option of receiving an email copy of the nonce words in each of the languages for which they completed tasks. This option was provided mainly for the learners, to ensure they knew which words constituted real Spanish words in order to prevent them from erroneously using nonce words in daily scenarios.

All of the production tasks were recorded using a Marantz professional solid-state recorder PMD661, and a unidirectional lavaliere microphone. During recording, the stimuli were digitized using a sample rate of 22,000 Hz and a 16-bit resolution.

Chapter 4

Results

4 Overview

The data for each language group was analysed acoustically and then submitted for statistical analysis. In Section 4.1.1, I describe the tokens analysed, followed by an explanation of how I carried out the acoustic analysis in Section 4.1.2. Then, for Romanian and Spanish, I present the results for manner (Romanian: Section 4.2.1; Spanish: Section 4.3.1), length (Romanian: Section 4.2.2; Spanish: Section 4.3.2) and percentage of voicing (Romanian: Section 4.2.3; Spanish: Section 4.3.3). In Section 4.2.4., I provide a brief summary of the findings for the Romanian rhotics. The selection of the acoustic parameters mentioned above is due to several reasons. While there has been a considerable amount of research on the phonetic and phonological characterization of French rhotics (e.g., Léon, 1992; Léon & Léon, 2009; O'Shaughnessy, 1982), our understanding of Romanian rhotics is limited; as such, the results presented here aim to characterize the Romanian rhotics in a more comprehensive way. They may also help to explain the patterns exhibited by the L1 Romanian speakers when producing Spanish rhotics[26]. For the Spanish rhotics, these three parameters were chosen based on previous research, which has found that learners target specific parameters first, and sometimes the parameters are in a trade-off relationship (e.g., Colantoni & Steele, 2006, 2007, 2008). Moreover, given that voicing can be linked to

[26] Recall from Chapter 3 that the L1 French data will not be acoustically analysed in this study, since there is ample previous research to characterize rhotics in French, and all of the participants in this study produced the expected uvular fricative rhotic in L1 French.

difficulties with controlling the length of segments, it is important to measure length in addition to voicing as well (Colantoni & Steele, 2006, 2007, 2008). Lastly, the L2 acquisition of sounds involves mastering different phonetic parameters at the same time, and, as such, different parameters need to be evaluated individually (Colantoni & Steele, 2008; Waltmunson, 2005).

In Section 4.4, I explore in more detail the unexpected realizations for Spanish, since the hypotheses make specific predictions about the nature of the variants produced instead of the expected rhotics. I explore the correlation of trill rates in L1 Romanian and trill rates in L2 Spanish in Section 4.5, in order to address whether learners who have higher rates of trilling their L1 Romanian also trill more in their L2 Spanish, since L1 articulatory routines may influence L2 routines (Olsen, 2012, 2016). In Section 4.6, I investigate correlations between two of the learners' individual variables, specifically AoA and weekly hours of Spanish use, and their rhotic accuracy rates in L2 Spanish. These two variables rather than other individual variables (such as the percentage of languages used in social situations, at school/work or at home) for two reasons. One is a practical reason, since all participants provided information regarding AoA and hours of use, and were consistent in the ways in which they reported this information. The second reason is due to previous research. Although these two variables are not the focus of this study and the hypotheses do not refer to them, some research has shown that AoA is a factor that significantly affects L2 speech. Specifically, an earlier AoA has been found to result in less accented vowel production among L1 Italian-L2 English speakers (Munro, Flege, & MacKay, 1996), and L1 Japanese learners of English with earlier AoAs have been shown to have less pronounced foreign

accents, as well as a higher separation of /l/ and /ɹ/ in the F2 dimensions (Ingvalson, McClelland, & Holt, 2011). Freed, Segalowitz and Dewey (2004) found that more L2 use (both spoken or written), as indicated by participants through self-reports, resulted in L2 French speech improvement, while Tremblay (2009) showed that L1 French Canadian speakers' self-reported use of L2 English was a significant predictor of the L2 perception of phonetically variable stress. As such, investigating AoA and amount of language use in more detail may provide partial explanations for the patterns found in the learners' rhotic productions in the present study. Then, in Section 4.7, I provide a more detailed description of the variability in each participant's rhotic productions, given that previous research has shown an effect of individual variability in both L1 and L2 speech (e.g., for L1 Spanish rhotics: Blecua, 2001; Colantoni, 2006; for L2: Munro et al., 1996; Riney & Flege 1998). The chapter ends with a summary of all the findings in Section 4.8.

4.1 Data analysis

4.1.1 Tokens analysed

For the Romanian rhotics, there was a maximum 480 possible tokens (10 speakers each producing 16 tokens in the picture description task and 32 tokens in the sentence reading task), however, only 435 tokens were analysed. 45 (9.4%) tokens were excluded due to poor sound quality and, in the case of RS11, 32 tokens were missing due to the sentence reading task accidentally not being recorded.

For the Spanish rhotics, there was a maximum of 1,617 possible tokens (33 speakers x 19 stimuli in the picture description task + 30 stimuli in the sentence reading task).

However, in total, 1,421 tokens were analysed: for the native Spanish controls, 590/637 tokens (92.6%); for the L1 Romanian-L2 Spanish (RO) speakers, 419/490 tokens (85.5%); lastly, for the L1 French-L2 Spanish (FR) speakers, 412/90 tokens (84.1%). Several tokens were discarded due to poor sound quality, resulting in the inability to reliably segment the rhotics. In addition, some of the stimuli in the picture description task were not produced by the learners and native Spanish speakers alike[27]. In sum, for the native Spanish controls, 47 tokens were missing (either excluded or not present at all); for the L1 Romanian-L2 Spanish speakers, 71 tokens were discarded; and for the L1 French-L2 Spanish speakers, 78 tokens were excluded from the analysis.

Regarding the picture description task for both Romanian and Spanish, while some participants followed the instructions and described the pictures in complete sentences, other speakers provided only two- or three-word descriptions, especially towards the end of the task when they may have forgotten the instructions or have become tired. These tokens were included in the analysis, as there was no reason to believe that failing to produce the target words in complete sentences affects the production of the rhotics. In both Romanian and Spanish, there were instances where some of the participants produced the target word more than once. In these instances, the most fluent realization (i.e., no false starts, no stuttering, etc.) was analysed. While this may be problematic because some speakers

[27] For example, in the picture description task, only five of the native Spanish controls produced the word *rabo* 'tail', only six of the participants in this group produced the word *cerro* 'hill', and the word *parra* 'vine' was only produced by one of the native Spanish controls. None of the speakers in either of the learner groups produced these three words in the picture description task. Among the learners, in addition to the words just mentioned, other stimuli that were not produced in the picture description task included: *roca* 'rock', *ropa* 'clothing', *corro* 'circle', *parra* 'vine', *coro* 'choir, *rata* 'rat', *reno* 'reindeer', *pera* 'pear', *zorro* 'fox', *carro* 'car'.

produced only one token and others as many as three, it was the solution that presented the lowest number of disadvantages. This solution allowed for the most precise and reliable phonetic analysis. Had the first realization of the target stimuli been analysed in all instances, more tokens may have had to be excluded if this realization was not as fluent and as clear as possible.

4.1.2 Acoustic analysis

For both Romanian and Spanish, the rhotics were segmented, labeled and subsequently analysed using Praat version 6.0.19 (Boersma & Weenink, 2016). The onset of the rhotics was determined by a decrease in F1 and intensity, while the offset was indicated by an increase F1 and intensity (Amengual, 2016; Colantoni, 2006; Johnson, 2008). Target stimuli were coded for: position in the word (WI: word-initial or WM: word-medial), task (P: picture description or R: sentence reading), and word type (real or nonce)[28]. Since a variety of rhotics were produced (e.g., tap, trill, fricative, approximant), the first step was to simply listen to the sound files and auditorily categorize each rhotic. For Romanian, the auditory analysis allowed for the identification of the following manners: taps, trills, fricatives, and some approximants. For Spanish, the following realizations were identified in the participants' productions: taps, trills, fricatives, approximants, English approximants, and uvular fricatives. Next, the realizations were analysed acoustically using Praat in order to achieve a more fine-grained categorization of place and manner. The final categorizations

[28] Though Word Type is not a variable included in the predictions of this study, some previous research (e.g., Frost, Repp, & Katz, 1988) has found that nonce words are more difficult to detect than real words. They are also associated with slower reaction times. Based on these previous findings, it was important to investigate whether there were differences in the production of rhotics in nonce versus real words.

of the rhotics relied more on the waveform and spectrogram than the auditory analysis. The following realizations were identified:

(i) Trills: These realizations were characterized by two or more closures. (Figure 9)

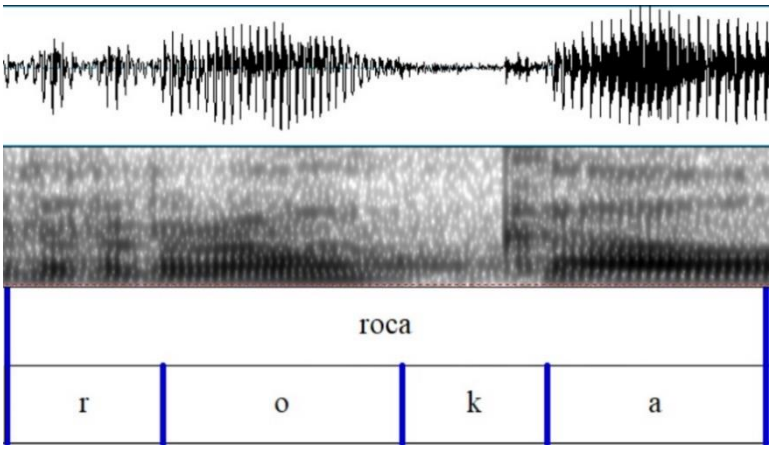

Figure 9. Word-initial trill production by a native female Romanian speaker RS01 in the word *roca* /roka/ 'rock' in the Reading Task.

(ii) Taps: These were characterized by a single, short closure. (Figure 10)

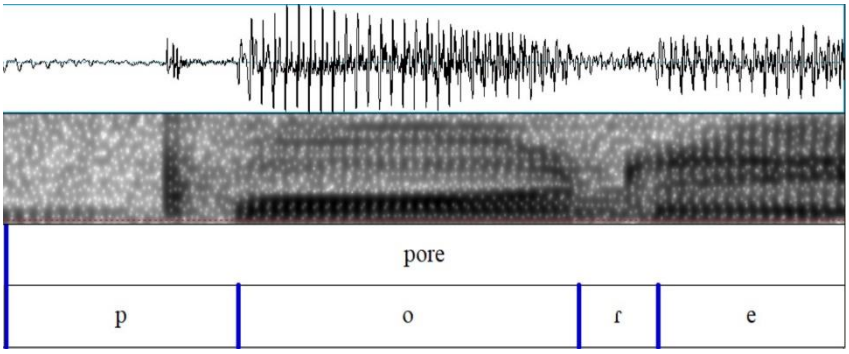

Figure 10. Word-medial tap production by a native female French speaker FS03 in the nonce word *pore* /poɾe/ in the Reading Task.

(iii) Approximants were characterized by a continuation of the formant structure throughout the rhotic and into the following vowel, and the absence of frication. (Figure 11)

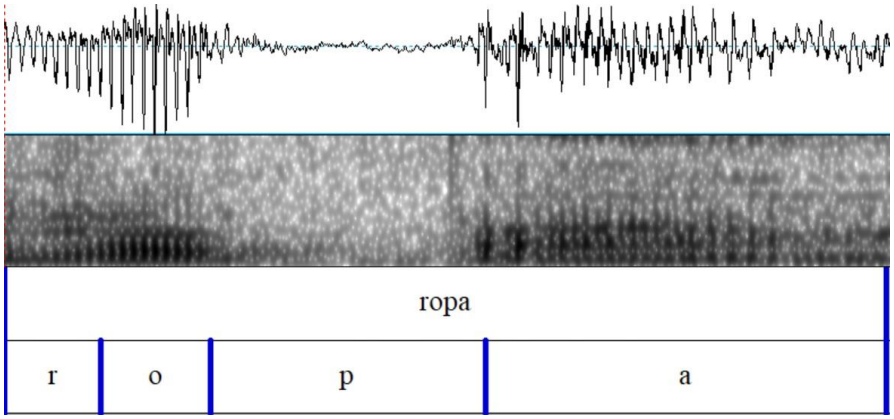

Figure 11. Word-initial approximant production by a native female Spanish speaker S08 in the word *ropa* /ropa/ 'clothes' in the Picture Task.

(iv) Taps with a fricated release were characterized by a closure followed by a fricated burst. (Figure 12)

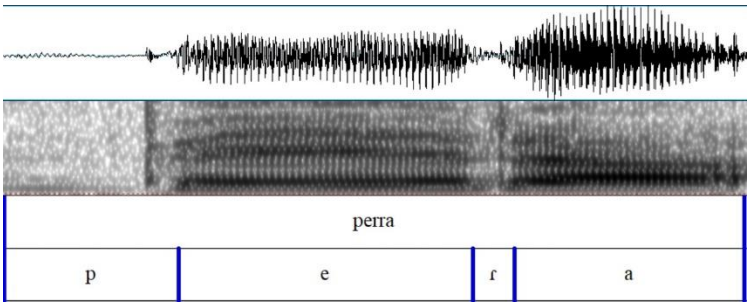

Figure 12. Word-medial production of a tap with a fricated release by a native female Romanian speaker RS09 in the word *perra* /pera/ 'dog' in the Reading Task.

(v) Fricatives were characterized by aperiodic noise in the higher frequencies. (Figure 13)

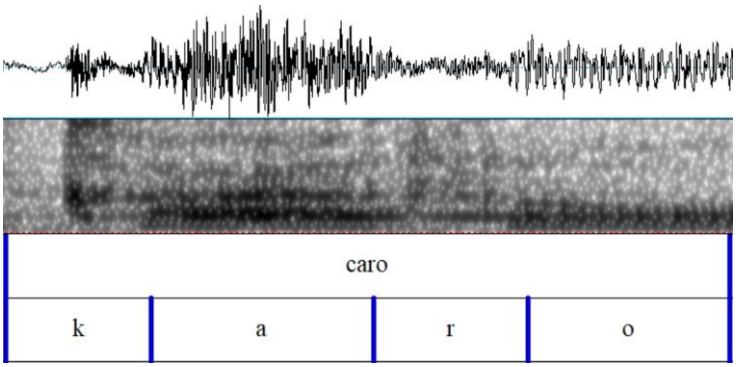

Figure 13. Word-medial fricative production by a native female Romanian speaker RS12 in the word *caro* /kaɾo/ 'car' in the Reading Task.

(vi) Uvular fricatives were categorized auditorily and by aperiodic noise in the higher frequencies (Figure 14)

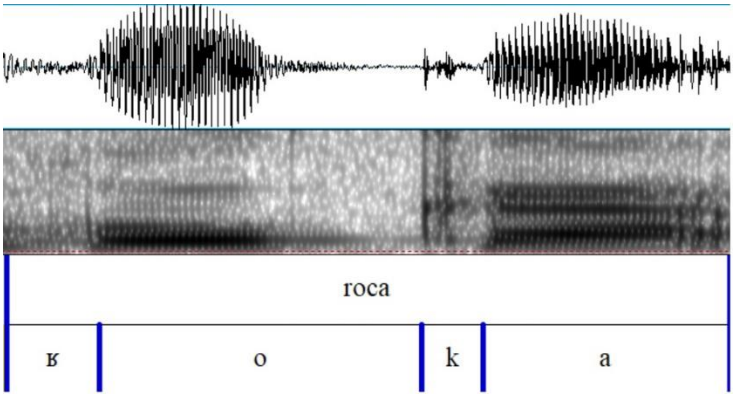

Figure 14. Word-initial production of a uvular fricative by a native female French speaker FS01 in the word *roca* /roka/ 'rock' in the Reading Task.

(vii) English approximants were identified auditorily, as well as by the absence of closures. (Figure 15)

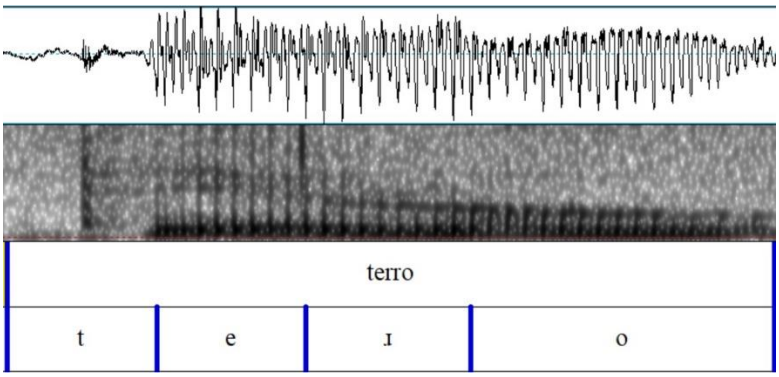

Figure 15. Word-medial production of an English approximant by a native female French speaker FS10 in the nonce word *terro* /tero/ in the Reading Task.

In addition, the length (ms) was measured and the percentage of voicing was calculated. These additional measurements were taken since, as mentioned in Chapter 2 Section 2.5, previous research has shown that learners do not target all parameters simultaneously (Colantoni & Steele, 2007; 2008). Although the length of taps was also measured, length was especially relevant in the case of trills, since these are generally longer than taps (Quilis, 1993). For trills only, the number of closures was also recorded. Percentage of voicing was calculated by dividing the duration of the presence of the F0 by the total duration of the segment. All measurements were taken at zero crossings.

In order to simplify the results for ease of presentation and to carry out the statistical analyses, for Spanish, the rhotic realizations depicted in Figures 9-15 were further grouped into the following categories: taps (including taps with a fricated release), trills, fricatives, approximants, uvular fricatives, and other (including English approximants, since there were only eight of these productions). For Romanian, there were four categories: taps, trills, fricatives, and approximants. In addition, for each rhotic, the length was measured, and the

percentage of voicing was calculated, since this portion of the analysis was exploratory in nature with the goal of providing a better characterization of Romanian rhotics, and there were no particular expectations about any of the rhotics' parameters (manner, length, % of voicing). As mentioned in Chapter 3, the French rhotics were not acoustically analysed, as participants did not produce any unexpected variants (i.e., the uvular fricative described in the literature was the expected and produced variant).

For all of the results, first, the descriptive statistics will be presented using various figures. Then, the inferential statistics will be presented for each parameter (i.e., manner, length and percentage of voicing). A detailed description of each statistical model/test will be given after the descriptive statistics are discussed. In order not to overcomplicate the graphic presentation of the results, tables with detailed totals and means (for manner), as well as means and standard deviations (for length and percentage of voicing) are presented in Appendix G and will be referred to throughout the present chapter. For the statistical analysis, the proportions of Spanish rhotic realizations (i.e., trills word-initially and word-medially intervocalically represented by <rr>; taps word-medially intervocalically represented by <r>) were compared across the three participant groups. Statistics for the length and percentage of voicing of taps and trills were only run on the taps and trills that occurred in the contexts where they are expected to occur in Spanish.

4.2 Romanian rhotics

4.2.1 Manner

Figure 16 shows the average types of rhotics produced in L1 Romanian (see Appendix G, Table G1 for totals). As previously attested in the literature, overwhelmingly, the most common realization was the tap, which occurred in 85% of cases in the present data. Though there were some trill productions as well, they occurred considerably less often than the tap, at only 8%. The number of closures in trills varied from one-three, with two being the most common (one closure: 25%; two closures: 59%; three closures: 16%). Lastly, some fricative (6.5%) and approximant realizations (0.5%) were also observed.

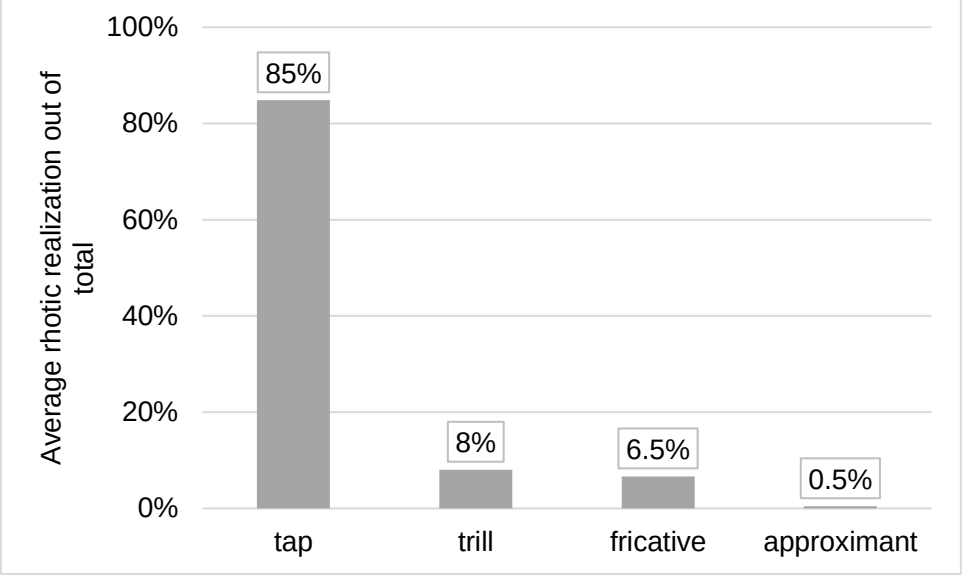

Figure 16. Average of types of rhotics produced by L1 Romanian speakers overall.

Since previous literature (Chitoran, 2002; Radu, 2016) has claimed that trill variants generally occur word-initially, it was important to investigate the speakers' realizations in

word-initial versus word-medial position. As can be seen in Figure 17 (see Appendix G, Table G2 for totals), the trill, crucially occurred almost exclusively word-initially rather than word-medially (15% versus 1%), which corroborates previous claims. Regarding the tap, it occurred more often word-medially than word-initially (98% versus 72%). Lastly, the fricative variant was observed at a low rate, and occurred mainly word-initially (12%). Approximant variants were almost non-existent (1%).

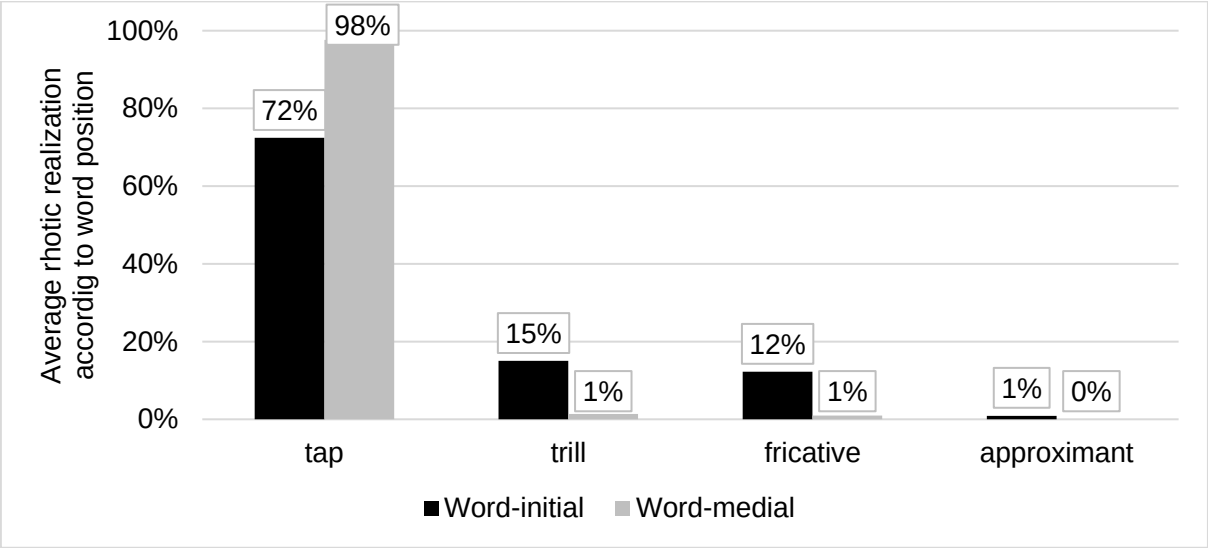

Figure 17. Average of types of Romanian rhotics in word-initial (WI) and word-medial (WM) positions.

As mentioned in Chapter 2, previous research has reported hyperarticulation in more formal tasks (e.g., Lindblom, 1989; Solé, 2002; Rafat, 2008), and, as such, we might expect more trills in the reading than in the picture description task. Figure 18 (see Appendix G, Table G3 for totals) displays the average rhotic realizations in each of the two tasks (picture description versus sentence reading). As can be seen in the figure, the tap was produced at nearly equal rates, regardless of the task (P: 87% versus R: 84%). The trill occurred slightly

more in the Picture than in the Reading task (10% versus 7%), which seems to be the opposite expectation, namely that trills would be more common in the reading than in the picture task.

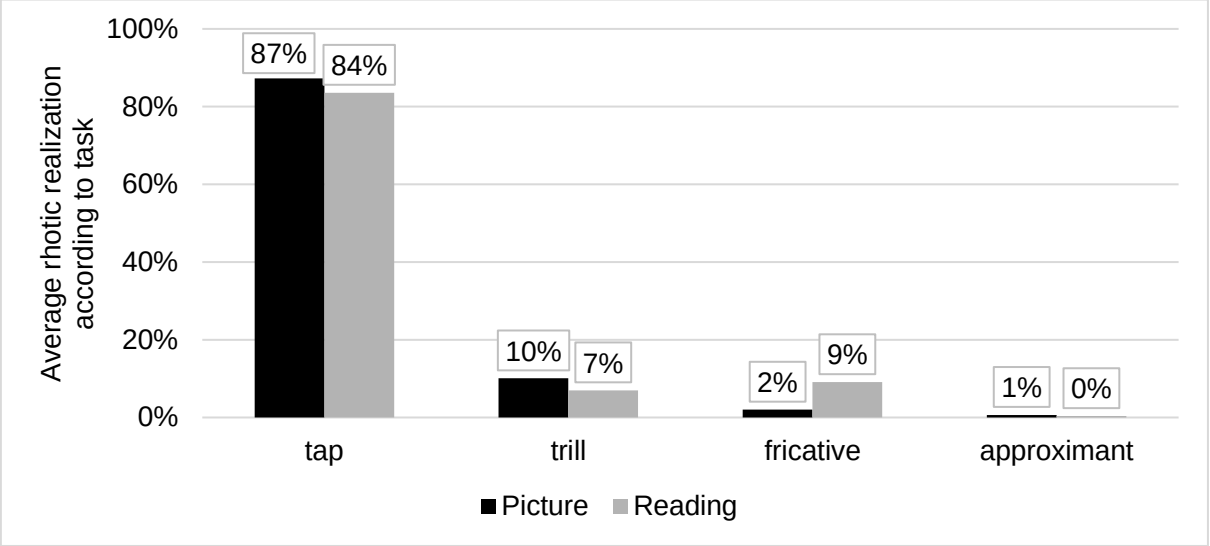

Figure 18. Average of types of Romanian rhotics in the picture description (P) and sentence reading (R) tasks.

Although word type (real versus nonce) was not a variable included in this study's research questions or predictions, it is worthwhile to discuss these results. As shown in Figure 19 (see Appendix G, Table G4 for totals), overall, there were not large differences in rhotic realizations between real and nonce words[29]. Taps were more common in real than

[29] Despite the fact that Word Type was not investigated in this study, a mixed model with Word Type (real versus nonce) as a fixed effect was also run. This factor was not significant ($p = 0.874$), suggesting the type of rhotic was not affected by whether the Word Type was real or nonce. The results of this model are presented in the following table:

Fixed effects	**Estimate**	**SE**	***z* value**	**Pr(>|z|)**
Intercept	-2.1051	0.4559	-4.618	$p < 0.001$ ***
WordTypeReal	-0.0654	0.4108	-0.159	$p = 0.874$

in nonce words (87% versus 80%; the trill realization occurred at almost the same rate in real and nonce words (7% and 8%, respectively). Lastly, the fricative realization occurred more in nonce than in real words (13% versus 4%), and the approximant hardly ever occurred in either word type.

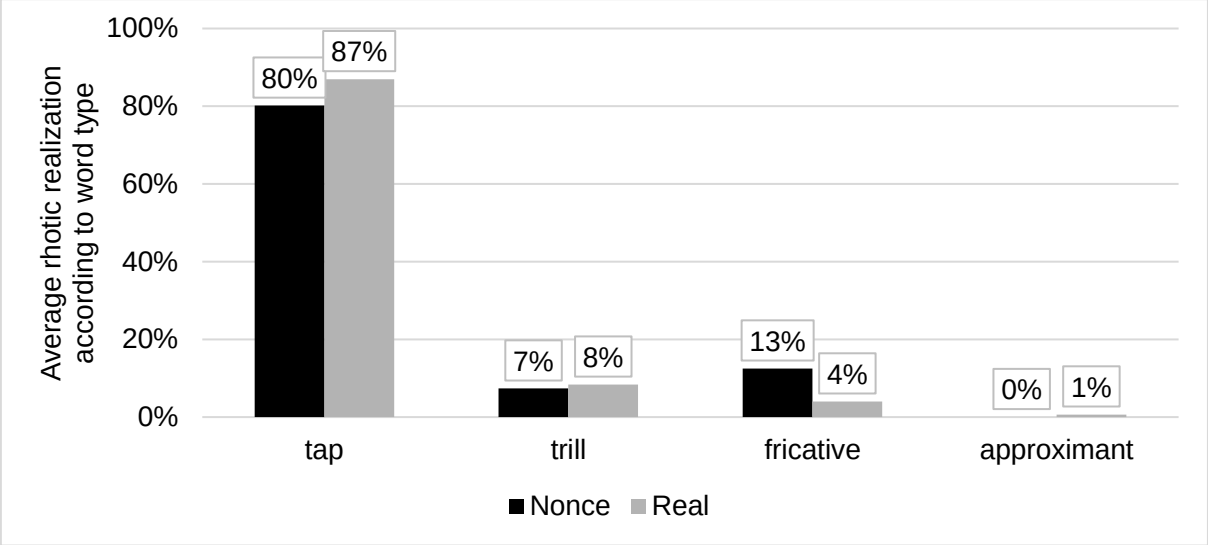

Figure 19. Average of types of Romanian rhotics in the nonce and real words.

Having presented the descriptive statistics, I now turn to the inferential statistics. Because fricative and approximant variants occurred in very few cases overall (6.5% and 0.5%, respectively; 29 tokens for the fricative and 2 for approximant out of 435 in total), and because only the tap and trill variants are the focus of the hypotheses in the current study, no statistical analyses were performed on the fricative and approximant realizations. As such, only the tap and trill realizations were analysed statistically. It is important to note that, four participants (RS04, RS05, RS07, RS11) were excluded from this particular statistical analysis. The first three participants had categorical rhotic realizations. That is, they only produced taps and, as such, there was no variance in these participants'

productions. Therefore, the effect of the independent variables on Manner could not be determined. The fourth participant was removed because, as mentioned in Section 4.1.1., the recording from this participant's reading task was lost; as such, Task could not be analysed as an independent variable for this participant. This exclusion resulted in the analysis of the productions of six participants (total of 272 tokens: 33 trills and 239 taps. With regard to position, 91% of the trills were produced word-initially (30/33) and 9% word-medially (3/33). 57% of the taps occurred word-medially (136/239) and 43% word-initially (103/239). In the reading task, the trill occurred 61% of the time (20/33), and 39% of the time in the picture description task (13/33). The tap followed a similar trend; 67% of taps (160/239) occurred in the reading task, and 33% (79/239) in the picture task. Due to the limited number of trills in this subset of the data (i.e., 33), in order to examine the effect of Position (word-initial, word-medial) and Task (reading and picture) on Manner, two non-parametric Fisher exact tests were run. The effect size was evaluated using the Crámer *V* (Levshina, 2015). The tests did not reveal a significant effect of Task (Fisher exact test = 0.760, $p = 0.556$), but did reveal a significant effect of Position (Fisher exact test = 0.076, $p < 0.001$). The effect size was moderate (see Crewson, 2005; Sheskin, 2011), with a Crámer *V* of 0.312. To summarize, though the Task did not affect whether the Romanian rhotic produced was a tap or trill, the Position in the word did, with the trill occurring significantly less word-medially than word-initially.

4.2.2 Length (ms)

Having discussed the manner of Romanian rhotics, the results for the length of the rhotics will now be presented. Figure 20 (see Appendix G, Table G5 for means and SDs; see

Appendix H, Figure H1 for box plot) displays the overall length (ms) of the rhotics in Romanian. Unsurprisingly, trills were longer than taps (74 ms versus 31 ms), and fricatives (68 ms) and approximants (60 ms) were slightly shorter than trills.

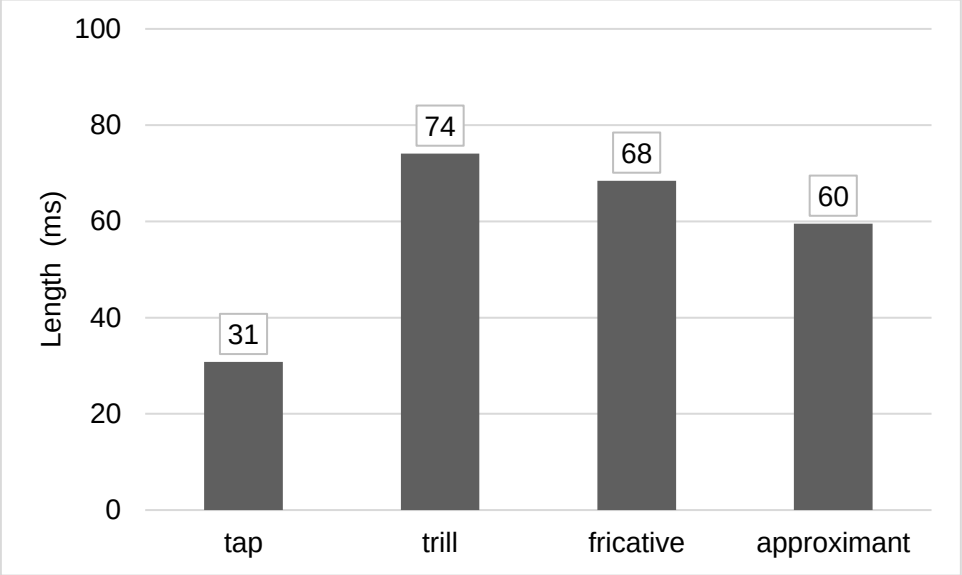

Figure 20. Overall length (ms) of Romanian rhotics produced by L1 Romanian speakers overall.

Figure 21 (see Appendix G, Table G6 for means and SDs; see Appendix H, Figure H2 for boxplots) displays the results for length (ms) of Romanian rhotics in each of the two positions (i.e., word-initial and word-medial). The tap was slightly longer word-initially than word-medially (34 ms versus 28 ms), while the trill was considerably longer word-initially than word-medially (76 ms versus 56 ms), which is a universal tendency.

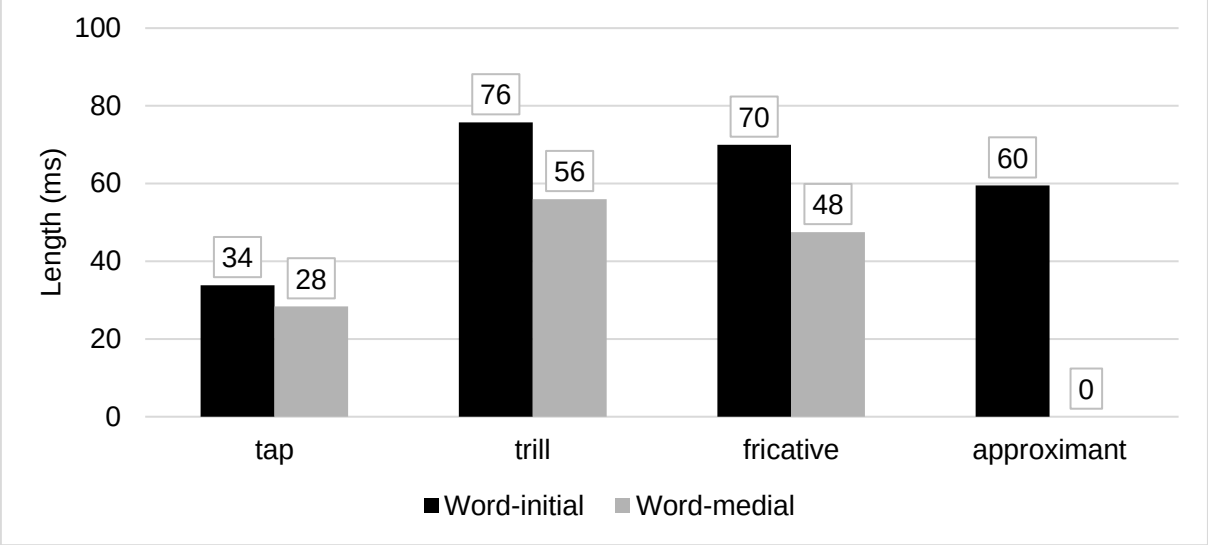

Figure 21. Length (ms) of Romanian rhotics produced by the L1 Romanian speakers in word-initial (WI) and word-medial (WM) positions.

As can be seen in Figure 22 (see Appendix G, Table G7 for means and SDs; see Appendix H, Figure H3 for boxplots), the tap had the same length in the picture description and in the sentence reading tasks (31 ms), while the trill was longer in the reading than in the picture task (81 ms versus 65 ms). That the trill was longer in the reading than in the picture task corroborates previous research that hyperarticulation occurs in read versus more spontaneous speech (Lindblom, 1989).

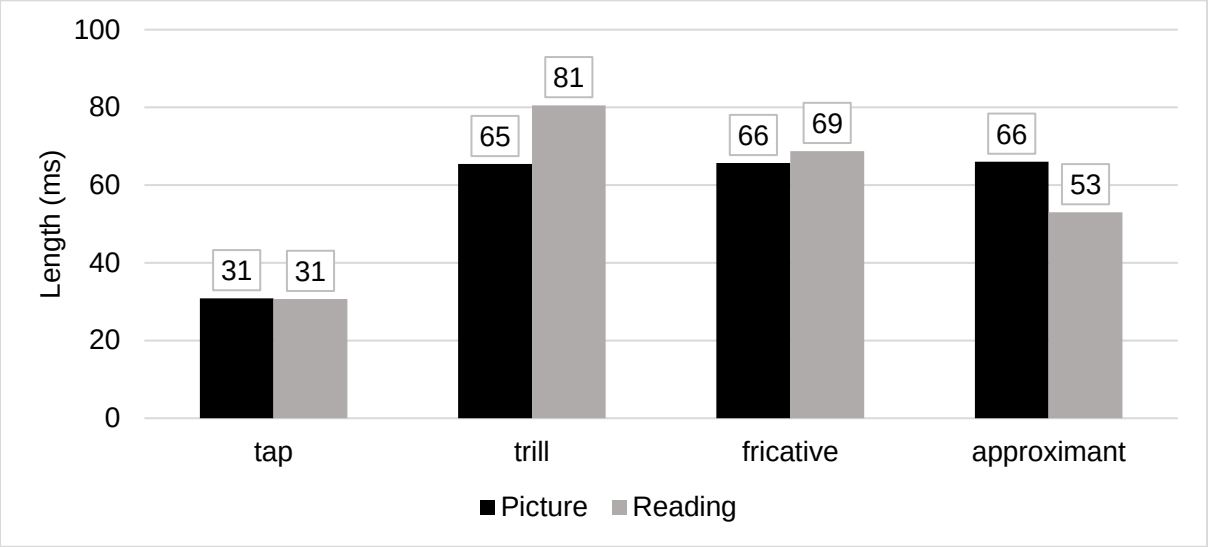

Figure 22. Length (ms) of Romanian rhotics produced by the L1 Romanian speakers in the picture description (P) and sentence reading (R) tasks.

Although no predictions were made regarding nonce versus real words, it is evident from Figure 23 (see Appendix G, Table G8 for means and SDs; see Appendix H, Figure H4 for boxplots) that for the length of the tap, the status of the word did not make a large difference; that is, the tap had equal lengths in both real and nonce words (31 ms). Both the trill and the fricative were somewhat longer in nonce than in real words (trill: 88 ms in nonce versus 68 in real words; fricative: 70 ms in nonce versus 66 in real words).

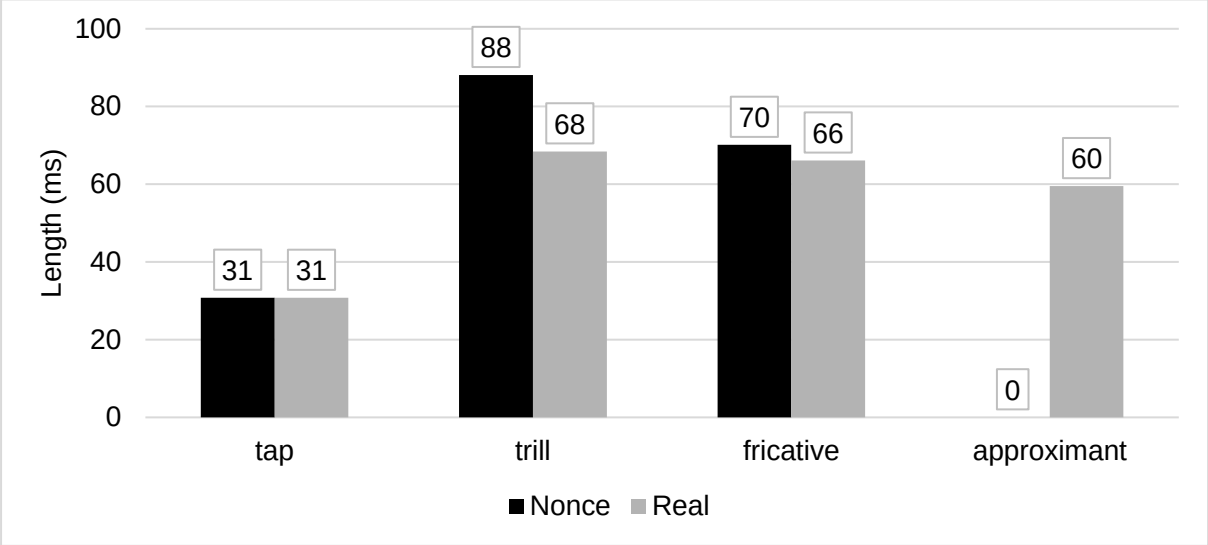

Figure 23. Length (ms) of Romanian rhotics in the nonce and real words.

In order to examine the effect of the variables of Position and Task on the Length of the rhotics, several analyses were run. For the tap realization, which, as mentioned in Section 4.2.1 was the most common, three generalized mixed effects models were fitted. All generalized mixed effects models were run using the *lme4* package in R (Bates, Mächler, Bolker, & Walker, 2015). Following an incremental approach (as in Pérez-Leroux, Peterson, Castilla-Earls, Béjar, Massam, & Roberge, 2018), the data was initially fitted with a base model (M0) including only the random effects, but no fixed effects. The next model (M1) had Length as the response variable, Task[30] as the fixed effect, and Participant and Item as random effects. The subsequent model (M2) had Length as the response variable, and Position in addition to Task as the fixed effects, and Participant and Item as random effects. The results of M2 are reported here, since a model comparison revealed that it was

[30] Word Type (i.e., nonce versus real) was included in a separate model, but was not selected as a significant factor. For the same reasons mentioned in Section 4.2.1, these results will not be reported here.

a better fit (M0: AIC = 2619.0; M1: AIC = 2620.6; M2: AIC = 2609.0)[31]. The fit of M2 was significantly better than the fit of the other two models (χ^2 = 13.06, p < 0.001). This model, displayed in Table 20, did not reveal a significant effect of Task, but did reveal a significant effect for Position, with taps being significantly shorter word-medially than -initially.

For the trill realizations, since there was a small number of tokens (35 out of a total of 435 or 8%), non-parametric tests were performed. Specifically, two Kruskal-Wallis tests were conducted to determine whether the effect of Position or Task on Length of trills was significant[32]. The tests did not reveal a significant effect of Position (χ^2(df=1) = 1.5323, p = 0.216) on trill length. However, a Kruskal-Wallis test showed that the effect of Task on trill length was significant (χ^2(df=1) = 4.5543, p = 0.033). That is, the trill was significantly longer in the sentence reading than in the picture description task.

Table 20

Results of a Generalized Linear Mixed Effects Model Examining Length (ms) of Taps with Position and Task as the Main Effects

Fixed effects	**Estimate**	**SE**	***t* value**	**Pr(>|z|)**
Intercept	3.52054	0.07059	49.875	p < 0.001 ***
PositionWM	-0.20866	0.04937	-4.226	p < 0.001 ***
TaskR	-0.02264	0.03213	-0.704	p = 0.481

4.2.3 Percentage of voicing

We now turn to the results for the percentage of voicing. Figure 24 (see Appendix G, Table G9 for means and SDs; see Appendix H, Figure H5 for boxplot) displays the overall results

[31] From this point forward, following the recommendations in Crawley (2013) and Levshina (2015), AIC values were used to determine the models that best fit the data. Smaller AIC values are indicative of better model fits (Crawley, 2013, pp. 416-417; Levshina, 2015, p. 194).

[32] As already mentioned, Word Type was not referred to in the hypotheses. However, a Kruskal-Wallis test was also conducted to determine whether the effect of Word Type on Length was significant. This test did not reveal a significant effect of Word Type on trill length (χ^2(df=1) = 3.0742, p = 0.08).

for percentage of voicing for each rhotic. Of the three types of rhotics produced, taps and trills exhibited similar percentages of voicing (88% and 89%, respectively), while fricatives and approximants were slightly more voiced (fricatives: 92%; approximants: 100%).

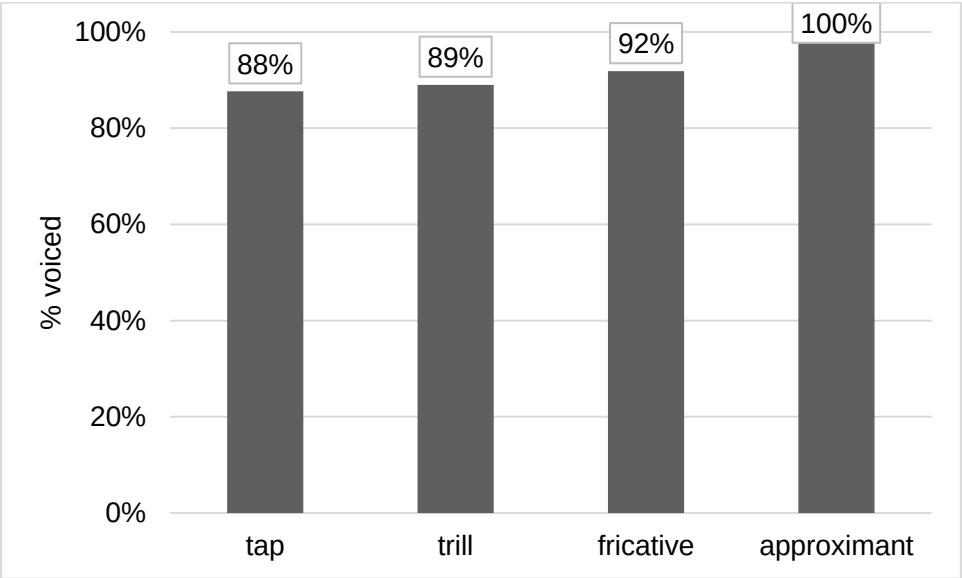

Figure 24. Overall percentage of voicing of Romanian rhotics produced by L1 Romanian speakers across the picture description and the sentence reading tasks.

Figure 25 (see Appendix G, Table G10 for means and SDs; see Appendix H, Figure H6 for boxplot) shows that percentage of voicing by rhotic type and position in the word. While for the tap, the percentage of voicing was higher word-initially than word-medially[33] (93% versus 83%, respectively), the trill was more voiced word-medially than initially (100% versus 88%). As for the fricative, it was more voiced word-initially than medially (100% versus 91%, respectively).

[33] Overall, the taps did not occur in absolute initial position; that is, they may have been word-initial, but those words may not have been the first words immediately after a pause (i.e., the first word of a sentence).

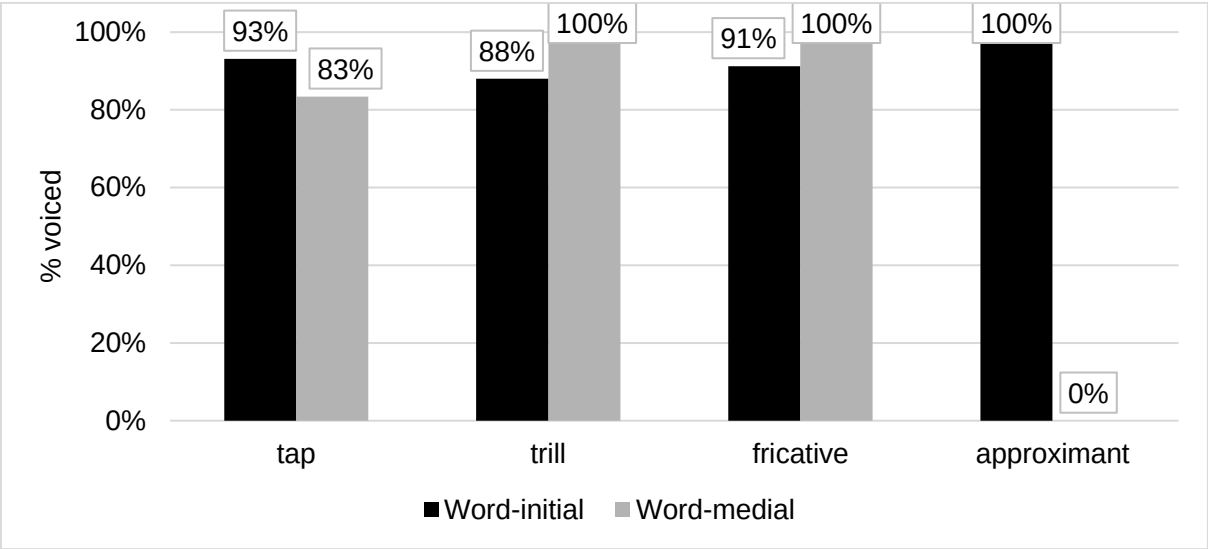

Figure 25. Percentage of voicing of Romanian rhotics in word-initial (WI) and word-medial (WM) positions.

Figure 26 (see Appendix G, Table G11 for means and SDs; see Appendix H, Figure H7 for boxplot) displays the results for the percentage of voicing for each rhotic realization by task. The tap and the trill exhibited similar trends in both tasks with both rhotics being more voiced in the reading than in the picture task (tap: 90% versus 83%; trill: 97% versus 78%). The fricative and approximant were almost voiced across the two tasks (fricative: 92%; approximant: 100%).

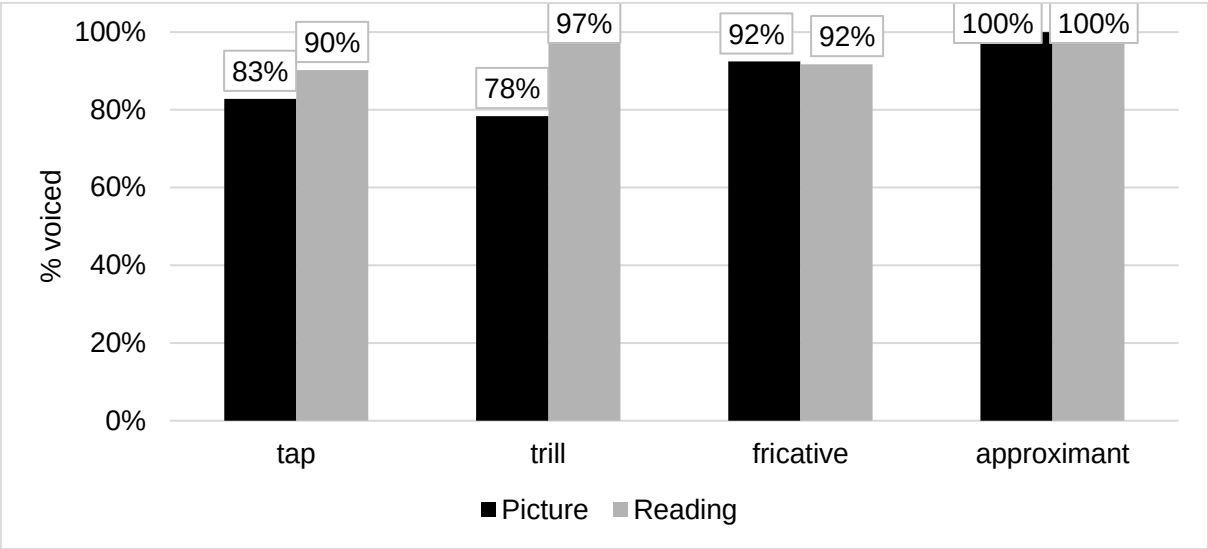

Figure 26. Percentage of voicing of Romanian rhotics in the picture description (P) and sentence reading (R) tasks.

Lastly, Figure 27 (see Appendix G, Table G12 for means and SDs; see Appendix H, Figure H8 for boxplot) represents the rhotic productions in nonce and real words. The tap and the trill followed a similar pattern, with both variants being more voiced in nonce than in real words (tap: 95% versus 85%; trill: 100% versus 85%). The fricative was slightly more voiced in real than in nonce words.

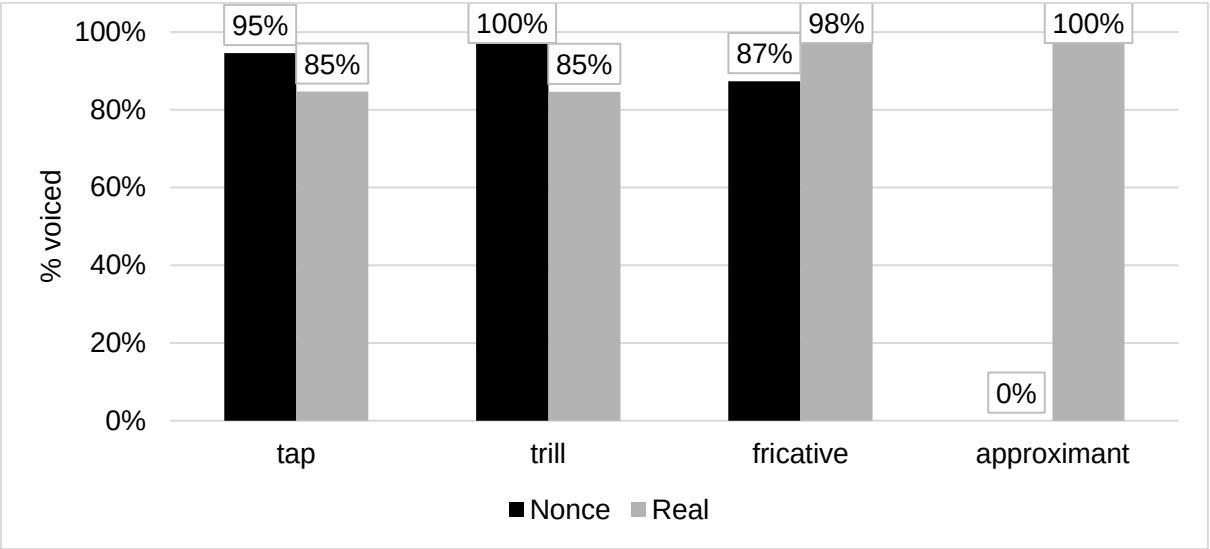

Figure 27. Percentage of voicing of Romanian rhotics in the nonce and real words.

The percentage of voicing was statistically analysed in a similar way to the length of the rhotics. That is, the effect of Position and Task on the percentage of voicing was first analysed for taps, and then for trills. Again, the fricative realizations were not statistically analysed, since only the percentage of voicing of taps and trills is relevant for the present study. For taps, two generalized mixed effects models with Percentage of Voicing as the response variable were conducted. M0 was the base model. M1 had Task as the fixed effect, and Participant[34] as the random effect. M2 had both Task and Position as fixed effects. Given that a model comparison revealed that the base model was the better fit, given that it has the lowest AIC value (M0: AIC = 4006.17; M1: AIC = 4007.593, M2: AIC = 4007.561). This difference was not statistically significant (χ^2 = 3.1538, p = 0.207). These results

[34] As mentioned in Section 4.2.1., Footnote 36, Item was not included as a random effect, since there were several stimuli that were produced exclusively with a tap, which created the problem of no variance. Eliminating Item from the random effects was the decision with the least disadvantages, since removing all of these tokens would result in a reduced data set for the analysis.

indicate that neither Position nor Task had a significant effect on the percentage of voicing in taps.

For trill realizations, two Kruskal-Wallis tests were run in order to investigate the effect of Position and Task on the percentage of voicing of the trills. The tests revealed that neither Position (χ^2(df=1) = 0.52743, p = 0.468), nor Task (χ^2(df=1) = 3.3673, p = 0.067), had a significant effect on the length of the trills[35]. As before, the fricative and approximant realizations were not analysed statistically.

To summarize, the most common rhotic in Romanian was the tap, as reported previously. Regarding trill productions, although these were not frequent, they occurred significantly less word-medially than initially. The type of task (reading versus picture) did not prove to be significant for the manner of rhotic produced. With respect to the length of rhotics, the tap was significantly shorter word-medially than initially, while the task was not significant. The trill, on the other hand, did not differ in length depending on the position within the word, but the type of task was significant, with the trill being significantly longer in the reading than in the picture task. Finally, neither the position nor the task was significant for the percentage of voicing of either taps or trills.

[35] As already mentioned, Word Type was not referenced in the hypotheses. However, a Kruskal-Wallis test was also conducted to determine whether the effect of Word Type on the percentage of voicing was significant. This test did not reveal a significant effect of Word Type on trill percentage of voicing (χ^2(df=1) = 2.2504, p = 0.137).

4.3 Spanish rhotics

4.3.1 Manner

I now turn to the manner of Spanish rhotics as produced by the three groups: L1 French-L2 Spanish, L1 Romanian-L2 Spanish and native Spanish speakers. It is important to note here that the learner groups' tap and trill (word-initial and medial) production rates were compared to the actual production rates of the native Spanish controls in this study. As mentioned in Chapter 2, Section 2.2.3, although normative descriptions of Spanish usually discuss only taps and trills, various phonetic studies suggest that there is considerable variation in the production of Spanish rhotics (e.g., Blecua, 2001; Colantoni, 2006 for Argentine rhotics; Navarro Tomás, 1971 for Peninsular Spanish; Quilis, 1993; Solé, 2002) due to aerodynamic constraints[36]. Considering this information, the learners' values were not compared to expected/anticipated tap or trill (depending on context) production rates of 100%, but rather to the actual tap and trill values observed among the Spanish controls in this study. In contexts that required taps (word-medially intervocalically corresponding to <r>), tap realizations were coded as Expected, while any other productions were coded as Other. For contexts that required trills (word-initially, and word-medially intervocalically represented by <rr>), trills were coded as Expected and any other realizations were coded as Other. These Other/Unexpected realizations will be described in more detail in Section 4.4.

[36] Though there are also several dialectal variants, as discussed in Chapter 2 Section 2.2.3, as mentioned in Section 3.2.1 of Chapter 3, none of the native Spanish speakers that participated in this study were speakers of dialects where rhotics are produced as preaspirated, dorsalized or any other variant (see the discussion of dialectal variation in Section 3.2.3 in Chapter 2).

Recall that the first prediction in this study (Hypotheses 1 and 2, Chapter 3 Section 3.1) was that the tap would be relatively easy for both learner groups. Further, although the L1 Romanian speakers were predicted to have higher trill rates than the L1 French speakers, they would still overgeneralize the tap to trill contexts. Moreover, it was predicted that, while the L1 French group would have higher trilling rates word-medially than initially, the opposite would be true for the L1 Romanian speakers, since this is where trills occur in Romanian. Table 21 presents a summary of the overall number of Expected (N=961 out of 1,421) and Other (N=460 out of 1,421) realizations for each speaker group (see Appendix G, Tables G13-G15 for means and totals for each language group). As previously mentioned, for trills, expected realizations were those that occurred either word-initially, or word-medially intervocalically in stimuli containing orthographic <rr> (where trills are expected in Spanish) while, for taps, expected realizations were those that occurred word-medially in stimuli containing orthographic <r> (where taps are expected in Spanish).

Table 21

Overall Number and Proportion of Tap and Trill (Word-initial and Medial) Productions by Language Group (L1 French, L1 Romanian, L1 Spanish) in the Expected Contexts

Rhotic expected		**Tap**		**Word-initial trill**		**Word-medial trill**	
		Expected	**Other**	**Expected**	**Other**	**Expected**	**Other**
French							
	%	70.75%	29.25%	13.01%	86.99%	42.86%	57.14%
	Total	104	43	19	127	51	68
Romanian							
	%	89.81%	10.19%	29.71%	70.29%	46.78%	53.22%
	Total	141	16	41	97	58	66
Spanish							
	%	98.05%	1.95%	86.67%	13.33%	93.71%	6.29%

Total	201	4	182	28	164	11

As can be seen in Table 21, the Romanian speakers had more expected tap realizations than the French speakers (89.81% versus 70.75%), while the controls produced the tap in most of the expected cases (98.05%). Regarding the word-initial trill context, both learner groups had low rates of the expected rhotic overall (FR: 13.01%; RO: 29.71%), while the Spanish controls once again produced the expected trills in most of the cases (86.67%). Lastly, both the Romanian and the French groups had overall low rates of the expected trill in the word-medial position (46.78% versus 42.86%, respectively), while the Spanish controls produced trills in the overwhelming majority of these cases (93.71%). The fact that the native Spanish controls did not produce the expected rhotics (i.e., taps or trills) 100% of the time is not unexpected given the variation reported for Spanish (see Section 2.2.3 in Chapter 2 for a more detailed discussion).

With respect to trill productions, the number of closures produced by all of the participants ranged from 1-6. Table 22 displays the proportion of trills that were produced with 1, 2, 3, 4, 5 and 6 closures for each speaker group. All three groups produced mainly trills with 2, followed by 3 closures. While the FR speakers did not produce any trills with 5 or 6 closures, both the RO and SPA groups produced trills with 5 closures, and one of the SPA controls produced one trill with 6 closures.

Table 22

Proportion of Target Trills Realized with 1, 2, 3, 4, 5, or 6 Closures across the Three Speaker Groups

	FR (N=70)	**RO (N=99)**	**SPA (N=346)**
# of closures			
1	20.00% (14)	25.25% (25)	14.74% (51)

2	51.43% (36)	40.40% (40)	48.84% (169)
3	18.57% (13)	22.22% (22)	28.61% (99)
4	10.00% (7)	8.08% (8)	6.65% (23)
5	0%	4.04% (4)	0.87% (3)
6	0%	0%	0.29% (1)

Note. Totals in Parentheses.

Table 23 displays the Expected and Other responses by Task for all three participant groups. As mentioned in Chapter 3 Section 3.1, Hypothesis 3 referred to the task differences between groups. It was predicted that the two learner groups would behave similarly, producing more target-like rhotics in the sentence reading than in the picture description task due to the increased task formality and the presence of orthography.

Table 23

Number and Proportion of Tap and Trill (Word-initial and Medial) Productions by Language Group (L1 French, L1 Romanian, L1 Spanish) in the Expected Contexts According to Task (Picture Description versus Sentence Reading)

Rhotic expected		**Tap**		**Word-initial trill**		**Word-medial trill**	
		Expected	**Other**	**Expected**	**Other**	**Expected**	**Other**
French							
	Picture						
	%	76.60%	23.40%	13.04%	86.96%	31.58%	68.42%
	Total	36	11	6	40	6	13
	Reading						
	%	68%	32%	13%	87%	45%	55%
	Total	68	32	13	87	45	55
Romanian							
	Picture						
	%	87.71%	12.29%	21.05%	78.95%	12.5%	87.5%
	Total	50	7	8	30	3	21
	Reading						
	%	91%	9%	33%	67%	55%	45%
	Total	91	9	33	67	55	45
Spanish							

Picture						
%	100%	0	77.78%	22.22%	84.78%	15.22%
Total	74	0	63	18	39	7
Reading						
%	96.95%	3.05%	92.25%	7.75%	96.9%	3.10%
Total	127	4	119	10	125	4

Regarding the expected tap productions, the L1 French speakers had higher tap rates in the picture than in the reading task (77% versus 68%). The L1 Romanian speakers produced a slightly higher proportion of taps in the reading than in the picture task (88% versus 91%). As for the native Spanish controls, they produced taps at nearly equal rates, at 100% of the time in the picture task and 97% of the time in the reading task. In word-initial contexts, where the trill is expected, the L1 French group produced trills at an almost identical rate across the picture and reading tasks (13.04% versus 13%, respectively), while the L1 Romanian group produced the expected trills 21% of the time in the picture task versus 33% of the time in the reading task. Lastly, the native Spanish controls produced a higher proportion of trills in the reading than in the picture task (97% versus 85%), a similar trend to the L1 Romanian group. Regarding word-medial trills, all three groups follow a similar trend, with higher rates of trill production in the reading than in the picture task. The L1 French group produced a higher proportion of trills in the reading than in the picture task (45% versus 32%), while the L1 Romanian group did so as well, but to a larger extent (55% versus 13%). Lastly, the control group produced word-medial trills in 85% of cases in the Picture task, and in 97% of cases in the Reading task.

First, in order to explore with which rhotic learners were more accurate overall, two binomial logistic regressions were fitted. All binomial logistic regressions were fitted using

the *lme4* package in R (Bates et al., 2015). The models had Response (Expected versus Other) as the dependent variable. M0 was the base model. Next, M1 had Expected Rhotic (Tap or Trill as the underlying form) and Language as the fixed effects, and Speaker and Item as the random effects. M2 had the same effects as the first, but, in addition, it also had Task as a fixed effect[37]. A model comparison revealed that M2 was a better fit (M0: AIC = 1240.396; M1: AIC = 1163.069; M2: AIC = 1152.665). This model was a significantly better fit ($\chi^2 = 12.403$, $p < 0.001$). The results of this model are presented in Table 24. These results show that, overall, speakers produced the expected trills at a lower rate than the expected taps, and Language was selected as a significant factor. In order to further explore the factor Language, which had three levels, a pairwise comparison was conducted using the *lsmeans* command in the *emmeans* with Tukey adjustments (Lenth, 2018). This comparison did not reveal a significant difference between the two learner groups ($\beta = 0.88$, $SE = 0.419$, *z-ratio* = 2.103, $p = 0.089$), but it did reveal a significant difference between the L1 French speakers and controls on one hand ($\beta = 4.04$, $SE = 0.442$, *z-ratio* = 9.141, $p < 0.001$), and the L1 Romanian speakers and the controls on the other ($\beta = 3.16$, $SE = 0.434$, *z-ratio* = 7.276, $p < 0.001$). Lastly, overall, speakers produced significantly more of the expected rhotics in the reading than the picture task[38].

[37]A third model was also conducted, exploring a Language*Expected Rhotic interaction, but this model failed to converge.

[38] Despite the fact that Word Type was not a variable investigated in this study, a mixed model with Word Type (real versus nonce) as a fixed effect was also run. This factor failed to reach significance ($p = 0.890$), suggesting that neither tap nor trill productions were affected by whether the Word Type was real or nonce. The results of this model are presented below:

Fixed effects	***OR***	**Estimate**	**SE**	***z* value**	**Pr(>|z|)**
Intercept	4.176903	1.42957	0.51308	2.786	$p = 0.005$ **
WordTypeReal	0.9359811	-0.06616	0.47946	-0.138	$p = 0.890$

Table 24

Results of a Binomial Logistic Regression Examining Response (Expected versus Unexpected) with Language, Task and Expected Rhotic as Main Effects

Fixed effects	***OR***	**Estimate**	**SE**	***z* value**	**Pr(>|z|)**
Intercept	1.903511	0.6437	0.3835	1.678	$p = 0.093$
LanguageRO	2.411864	0.8804	0.4186	2.103	$p = 0.035$ *
LanguageSPA	56.65612	4.0370	0.4416	9.141	$p < 0.001$ ***
TaskR	2.001506	0.6939	0.1955	3.549	$p < 0.001$ ***
ExpectedTrill	0.08628496	-2.4501	0.2836	-8.639	$p < 0.001$ ***

In order to explore the effect of the independent variables on the Expected versus Other rhotic productions, statistics were run on the expected tap and trill realizations separately. Beginning with taps, three binomial logistic regressions were fitted. M0 was the base model. M1 had Response (Tap versus Other) as the dependent variable, Language as the fixed effect, and Speaker and Item as the random effects; M2 had the same random effects as the first, but had Task instead of Language as the fixed effect. Lastly, M3 had the same random effects as the first two models, but had both Language and Task as the fixed effects. A model comparison revealed that M1 was the best fit (M0: AIC = 310.5774; M1: AIC = 298.7505; M2: AIC = 311.9961; M3: AIC = 300.2014), and the difference was statistically significant ($\chi^2 = 15.2456$, $p < 0.001$), suggesting that Task did not affect tap accuracy. The results of the first model are presented in Table 25. As can be observed, Language was selected as a significant factor, which was further explored with pairwise comparisons using the *lsmeans* command in the *emmeans* package with Tukey adjustments (Lenth, 2018). The comparison showed that, though the L1 Romanian speakers were more target-like than the L1 French speakers, there was no significant difference between the two learner groups ($\beta = 1.47$, $SE = 0.738$, $z\text{-}ratio = 1.987$, $p = 0.116$). Likewise, there was no significant

difference between the L1 Romanian group and Spanish controls (β = 2.00, *SE* = 0.888, *z-ratio* = 2.255, p = 0.062). However, the controls significantly outperformed the L1 French speakers (β = 3.47, *SE* = 0.874, *z-ratio* = 3.968, $p < 0.001$) in tap contexts.

Table 25

Results of a Binomial Logistic Regression Examining Response (Tap versus Other) in Tap Contexts with Language as a Main Effect

Fixed effects	***OR***	**Estimate**	**SE**	***z* value**	**Pr(>|z|)**
Intercept	3.418152	1.2291	0.52223	2.353	$p = 0.019$ *
LanguageRO	4.335773	1.4669	0.7384	1.987	$p = 0.047$ *
LanguageSPA	32.12068	3.4695	0.8743	3.968	$p < 0.001$ ***

For the trill realizations, four models were run. M0 was the base model. M1 had Response (Expected versus Unexpected) as the dependent variable, Language as the fixed effect, and Speaker and Item as random effects; M2 was the same as Model 1, except it had Task in addition to Language as fixed effects; M3 was the same as M2, but had Position added as a main effect; lastly, M4 was the same as Model 3, but additionally had the Language*Position interaction specified, since the hypotheses in this study specifically make reference to the positions in which learner groups would first target trills. A model comparison revealed that all four models were good fits (M0 = AIC =780.8166; M1 AIC = 752.4502; M2: AIC = 727.7022; M3: AIC = 707.2067; M4: AIC = 703.0771), but M4 was a significantly better fit (χ^2 = 8.1296, $p < 0.017$). The results of M4 for the trill realizations are given in Table 26. In this model, all factors were significant. Task was selected as a significant factor, with all participants producing more of the expected trills in the Reading than in the Picture task. Likewise, Position was also significant, with participants producing more of the expected trills word-medially than initially. Language was also selected as

significant, with both the Romanian and Spanish groups producing overall significantly more of the expected trills than the French group (the Romanian group to a much lesser extent than the Spanish). Lastly, the Language*Position interaction was selected as significant. In order to further explore this interaction, a pairwise comparison was carried out using the *lsmeans* command in the *emmeans* package with Tukey adjustments (Lenth, 2018). When the Language*Position interaction was further explored, there was no significant difference between the L1 French and L1 Romanian groups in WI position (β = 2.038, *SE* = 0.900, *z-ratio* = 2.265, p = 0.2089) or WM positions (β = 0.736, *SE* = 0.864, *z-ratio* = 0.853, p = 0.957), but there were several significant differences between the learners and controls. Namely, in WI position, the Spanish controls produced significantly more trills than both the L1 French (β = 6.122, *SE* = 0.901, *z-ratio* = 6.793, $p < 0.001$) and L1 Romanian speakers (β = 4.084, *SE* = 0.826, *z-ratio* = 4.944, $p < 0.001$). This was also true in WM position, where the native Spanish controls significantly outperformed the learner groups (Romanian: β = 3.994, *SE* = 0.865, *z-ratio* = 4.617, p = 0.001; French: β = 4.731, *SE* = 0.899, *z-ratio* = 5.260, $p < 0.001$). Additionally, and crucially for the predictions, both the L1 Romanian group (β = 1.016, *SE* = 0.333, *z-ratio* = 3.049, p = 0.028) and the L1 French group (β = 2.317, *SE* = 0.424, *z-ratio* = 5.471, $p < 0.001$) produced significantly more trills in WM than in WI position, and there was no significant difference between the two learner groups (β = 0.736, *SE* = 0.864, *z-ratio* = 0.853, p = 0.957).[39]

[39] Similar to the results section on the Romanian rhotics, Word Type (nonce versus real) did not have a significant effect in L2 Spanish. Moreover, none of the hypotheses make predictions about Word Type.

Table 26

*Results of a Binomial Logistic Regression Examining Response (Trill versus Other) in Trill Contexts with Task, Language, Position and Language*Position as Main Effects*

Fixed effects	***OR***	**Estimate**	**SE**	***z* value**	**Pr(>|z|)**
Intercept	0.01349414	-4.3055	0.7199	-5.980	$p < 0.001$ ***
LanguageRO	7.674476	2.0379	0.8998	2.265	$p = 0.024$ *
LanguageSPA	455.8209	6.1221	0.9012	6.793	$p < 0.001$ ***
TaskR	4.091043	1.4088	0.2519	5.593	$p < 0.001$ ***
PositionWM	10.14621	2.3171	0.4235	5.471	$p < 0.001$ ***
LanguageRO:PositionWM	0.2721505	-1.3014	0.5147	-2.528	$p = 0.012$ *
LanguageSPA:PositionWM	0.248702	-1.3915	0.5717	-2.434	$p = 0.015$ *

In summary, the results confirm the prediction that the tap was an easier segment for the learners compared to the trill, and, while the L1 Romanian group had significantly higher trilling rates than the L1 French group overall, when the Language*Position interaction was explored, there were no significant differences between the two learner groups. That is, regarding the position within the word, against the prediction made in the current study, both learner groups had significantly higher rates of trilling word-medially than initially. Overall, as well as between word positions, the controls produced significantly more of the expected trills than the L1 Romanian and French speakers.

4.3.2 Length (ms)

Having discussed the manner of the rhotics produced by the three groups, this section will discuss the length of the Spanish rhotics, since being native-like does not only involve targeting manner, but also length and percentage of voicing. As mentioned in Section 4.1.1, statistics for the length (ms) and percentage of voicing were only run on expected tap and trill realizations. That is, trills and taps that were produced in unexpected contexts were

considered unexpected and subsequently excluded from this analysis. A total of 961 (taps=446; trills=515) tokens were analysed: 174 for the L1 French-L2 Spanish speakers, 240 for the L1 Romanian-L2 Spanish speakers, and 547 for the native Spanish controls.

Regarding the overall length of taps and trills, unsurprisingly, the tap was much shorter than the trill for all speaker groups (see Appendix G, Table G25 for means and SDs). Regarding the tap realization, it was produced with a comparable length by all three speaker groups. Specifically, the average tap length for each group was 33 ms for the L1 French-L2 Spanish: 33 ms for the L1 Romanian-L2 Spanish group; and 30 ms for the native Spanish controls. The same trend was true for the trills, with mean lengths as follows: 91 ms for the L1 French-L2 Spanish group; 88 ms for the L1 Romanian-L2 Spanish group; and 86 ms for the native Spanish controls.

Figure 28 (see Appendix G, Table G25 for means and SDs; see Appendix H, Figure H9 for boxplots) displays the overall length of expected taps and trills as produced by the three speaker groups in the WI and WM positions. As can be seen, the average length of trills word-medially was comparable among the two learner groups, but was slightly longer for the SPA controls. Regarding word-medial trills, both FR and RO speakers had slightly longer trills than the SPA group. Additionally, the learners produced longer trills word-medially than initially. Lastly, tap lengths were comparable across the three groups.

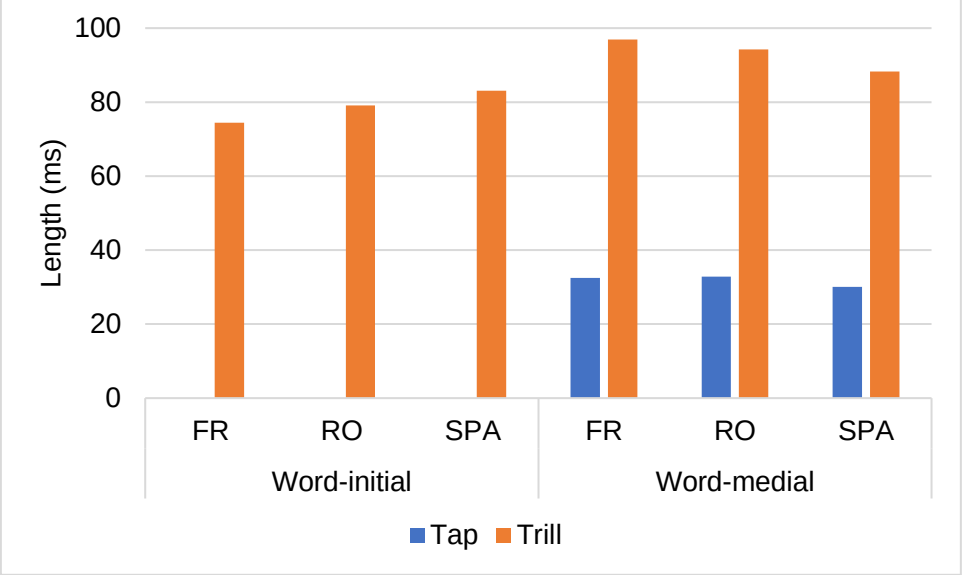

Figure 28. Length (ms) of expected Spanish taps and trills as produced by the L1 French-L2 Spanish (FR), L1 Romanian-L2 Spanish (RO) and native Spanish (SPA) groups in word-initial (WI) and word-medial (WM) positions.

Figure 29 (see Appendix G, Table G26, see Appendix H, Figure H10 for boxplots) displays the length of expected taps and trills in the Picture and Reading tasks for the three speaker groups. Regarding the tap realization, all of the groups had comparable lengths across tasks. In the Picture task, RO speakers produced the shortest trills of all the groups, while FR speakers produced the longest ones. However, in the Reading task, the three speaker groups produced trills of comparable lengths.

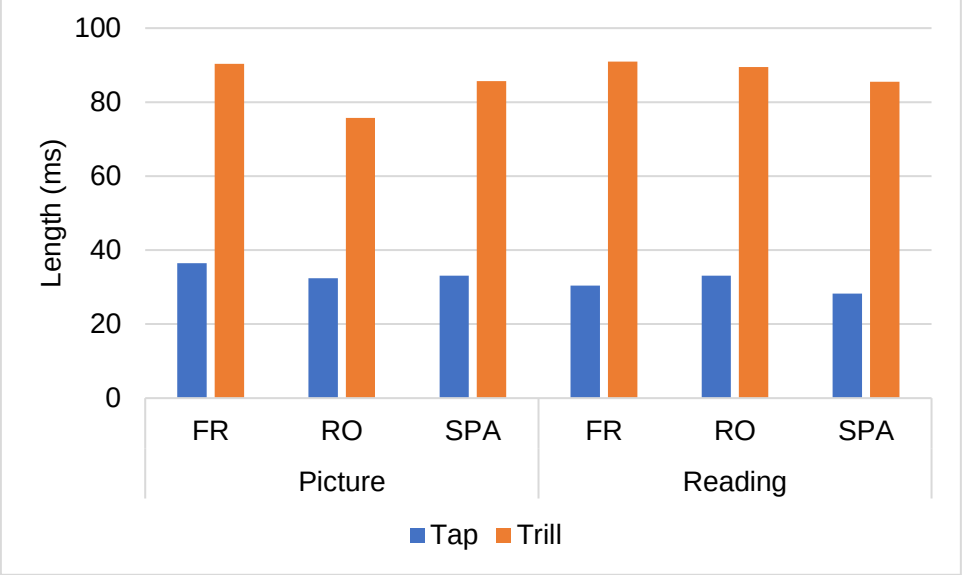

Figure 29. Average length (ms) of expected Spanish taps and trills as produced by the L1 French-L2 Spanish (FR), L1 Romanian-L2 Spanish (RO) and native Spanish (SPA) groups in the picture description and sentence reading tasks.

Lastly, Figure 30 (see Appendix G, Table G27 for means and SDs; see Appendix H, Figure H11 for boxplots) displays the length of expected taps and trills for the three speaker groups according to word type (nonce versus real). All participants' taps had similar lengths with both nonce and real words. Although the three speaker groups produced roughly equally long trills in the nonce words, the FR group produced longer trills than either the RO or SPA group in the real words.

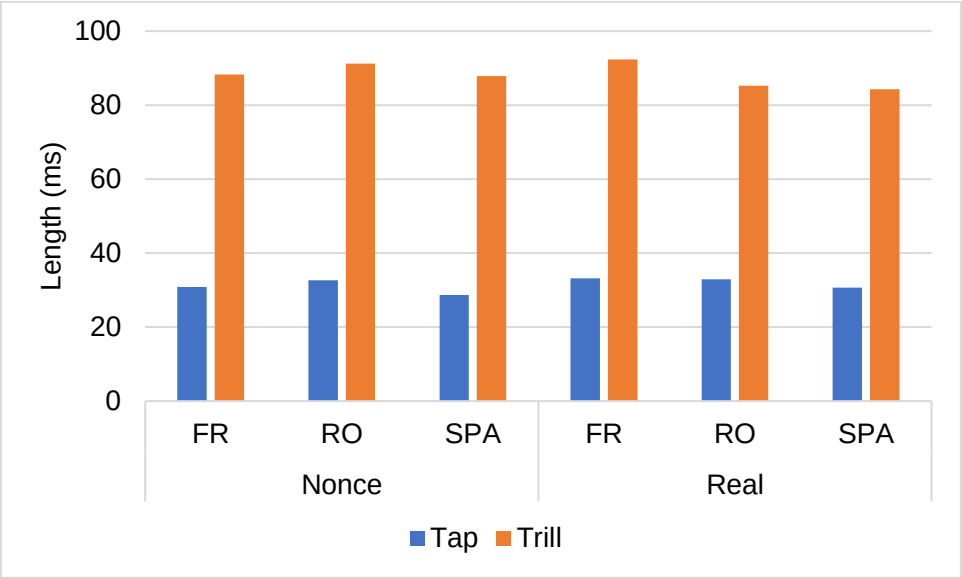

Figure 30. Average length (ms) of expected Spanish taps and trills as produced by the L1 French-L2 Spanish (FR), L1 Romanian-L2 Spanish (RO) and native Spanish (SPA) groups in nonce and real words.

In order to analyse the effect of the independent variables on expected rhotic length, separate generalized linear mixed effects models were run for taps and trills, because the length of these two rhotics is inherently different; thus, grouping all of these lengths together and running one statistical test would skew the results.

For tap length, four generalized linear mixed effects models were run: M0 was the base model; M1 had the Length of taps as the Response variable, Language as the fixed effect, and Speaker and Item as random effects; M2 had Task in addition to Language as fixed effects; M3 was the same as M2, but also tested the Language*Task interaction. A model comparison revealed that M3 was the best fit (M0: AIC = 3215.893; M1: AIC = 3218.929; M2: AIC = 3210.451; M3: AIC = 3207.476). M3 was a significantly better fit (χ^2 = 6.975, p = 0.03). The results of the third generalized mixed effects model are presented in

Table 27. However, as can be seen in the table, none of the predictors was selected as significant. That is, neither Language, nor Task, nor the Language*Task interaction had a significant effect on tap length.

Table 27

*Results of a Generalized Linear Mixed Effects Model with Length of Expected Taps as the Response Variable and Language, Task, and Language*Task as the Main Effects*

Fixed effects	**Estimate**	**SE**	***t* value**	**Pr(>|z|)**
Intercept	3.56783	0.10981	32.490	$p < 0.001$ ***
LanguageRO	-0.07768	0.15484	-0.502	$p = 0.616$
LanguageSPA	-0.07169	0.14590	-0.491	$p = 0.623$
TaskR	-07255	0.06096	-1.190	$p = 0.234$
LanguageRO:TaskR	0.06810	0.07693	0.885	$p = 0.376$
LanguageSPA:TaskR	-0.09502	0.07193	-1.321	$p = 0.187$

The next set of models was five generalized linear mixed effects models that had the Length of trills as the Response variable. M0 was the base model. M1 had Language as the fixed effect, and Speaker and Item as the random effects; M2 had Task in addition to Language as fixed effects; M3 had Position in addition to Task and Language as fixed effects. Additionally, because the hypotheses refer to the positions in the word in which each language group would be more native-like, M4 additionally explored the Language*Position interaction in addition to the fixed effects in M3. A model comparison revealed that M4 was the best fit (M0: AIC = 4634.910; M1: AIC = 4637.846; M2: AIC = 4639.212; M3: AIC = 4635.165; M4: AIC = 4631.627); that is, the model specifying the Language*Position interaction was a significantly better fit ($\chi^2 = 15.283$, $p = 0.018$). The results of M4 are presented in Table 28. In this model, Position selected as significant, with word-medial trills being significantly longer than word-initial ones for all speakers. The Language*Position interaction was also selected as significant. This interaction was

explored with a pairwise comparison using the *lsmeans* command in the *emmeans* package with Tukey adjustments (Lenth, 2018). It revealed that the significant interactions were for the L1 French and L1 Romanian speakers, who had significantly longer trills word-medially than they did word-initially (French: $\beta = 2.099$ $SE = 0.0736$, $z\text{-}ratio = 2.851$, $p = 0.05$; Romanian: $\beta = 0.1869$ $SE = 0.0591$, $z\text{-}ratio = 3.163$, $p = 0.02$).

Table 28

*Results of a Generalized Linear Mixed Effects Model Examining Length of Expected Trills with Language, Task, Position and Language*Position as the Main Effects*

Fixed effects	**Estimate**	**SE**	***t* value**	**Pr(>|z|)**
Intercept	4.32377	0.10555	40.962	$p < 0.001$ ***
LanguageRO	-0.05268	0.12863	-0.410	$p = 0.682$
LanguageSPA	0.10753	0.12394	0.868	$p = 0.386$
TaskR	-0.02435	0.02864	-0.850	$p = 0.395$
PositionWM	0.20995	0.07363	2.851	$p = 0.004$ *
LanguageRO:PositionWM	-0.02301	0.08995	-0.256	$p = 0.798$
LanguageSPA:PositionWM	-0.15502	0.07551	-2.053	$p = 0.04$ *

4.3.3 Percentage of voicing

Having discussed both the manner and length results, this last section focuses on the percentage of voicing of Spanish rhotics. Figure 31 (see Appendix G, Table G28 for means and SDs) shows the overall percentage of voicing for expected taps and trills for each of the three speaker groups. Overall, all of the groups voiced the taps more than the trills. However, the French and Spanish groups did so to a larger extent than the Romanian group, whose taps and trills were voiced at comparable rates.

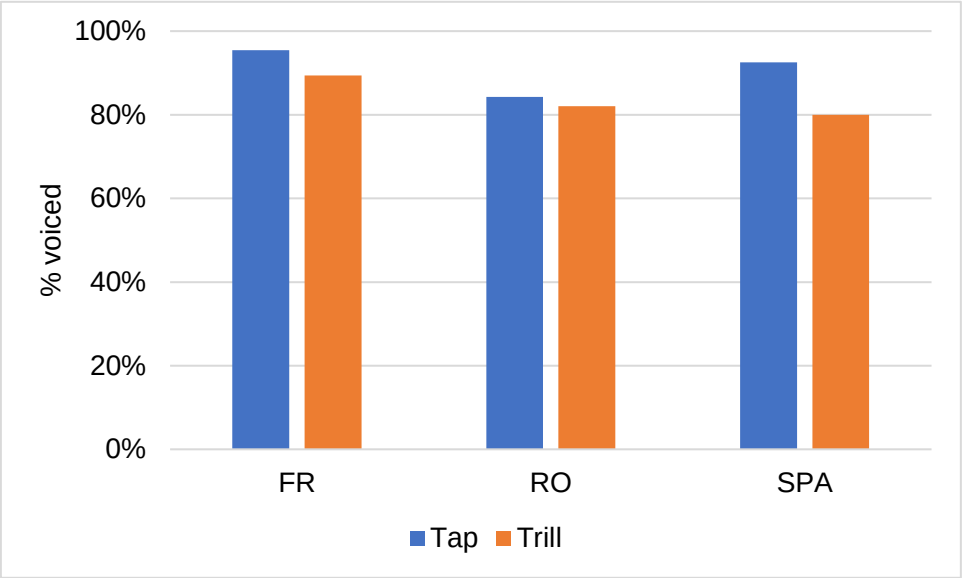

Figure 31. Average percentage of voicing of Spanish rhotics as produced by the L1 French-L2 Spanish (FR), L1 Romanian-L2 Spanish (RO) and native Spanish (SPA) groups.

Figure 32 (see Appendix G, Table G28 for means and SDs; see Appendix H, Figure H12 for boxplots) displays the percentage of voicing results for the three groups in word-initial (WI) and word-medial (WM) positions. Regarding the trill, which occurred in both positions, the two learner groups produced trills that were slightly more voiced in word-initial than -medial position. However, the SPA control group produced trills that were more voiced word-medially than word-initially. In terms of taps, the RO group produced taps with the least voicing of the three groups, while the FR speakers had taps that were slightly more voiced than the SPA group's.

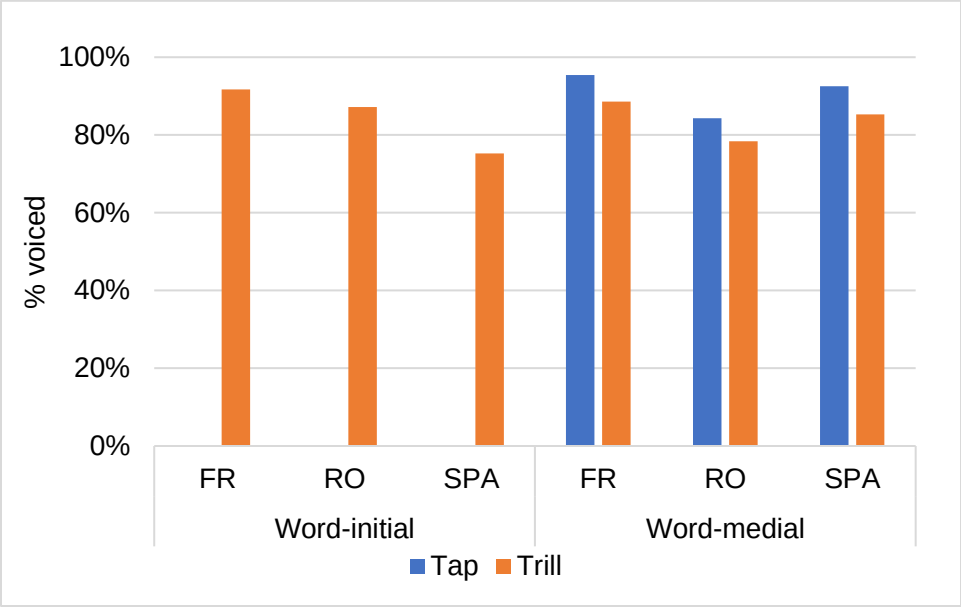

Figure 32. Average percentage of voicing of Spanish rhotics as produced by the L1 French-L2 Spanish (FR), L1 Romanian-L2 Spanish (RO) and native Spanish (SPA) groups in word-initial and word-medial positions.

Figure 33 (see Appendix G, Table G29 for means and SDs; see Appendix H, Figure H13 for boxplots) shows the percentage of voicing of taps and trills as produced by the three groups in the picture description (P) and sentence reading (R) tasks. Regarding taps, both the FR and SPA groups produced taps that were more voiced in the R than in the P task. The RO speakers, however, had almost equally voiced taps in both tasks. FR speakers produced trills with slightly more voicing in the P than in the R task, while RO speakers' trills were almost equally voiced in the two tasks. The SPA controls' trills were only slightly more voiced in the R than in the P task.

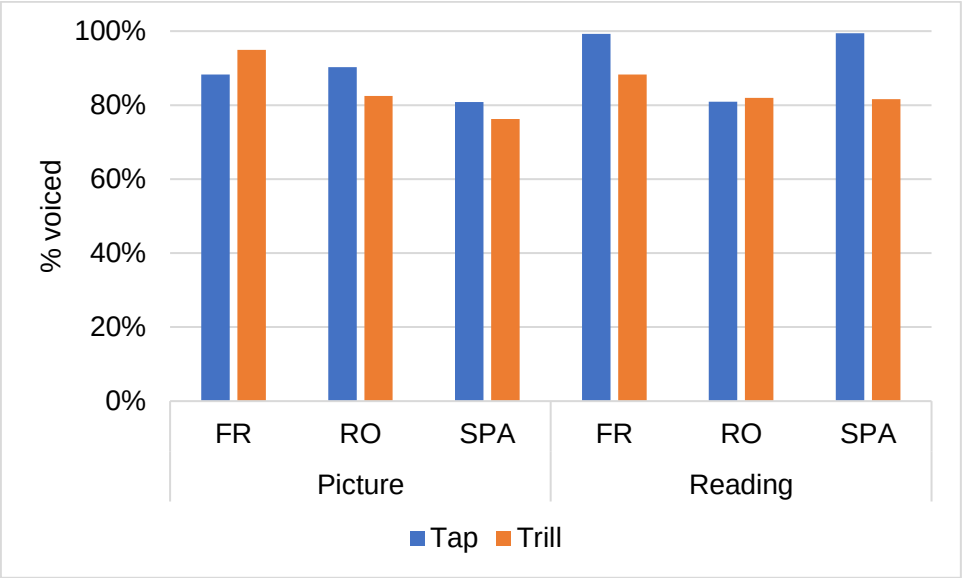

Figure 33. Average percentage of voicing of Spanish rhotics as produced by the L1 French-L2 Spanish (FR), L1 Romanian-L2 Spanish (RO) and native Spanish (SPA) groups in the picture description and sentence reading tasks.

Finally, Figure 34 (see Appendix G, Table 30 for means and SDs; see Appendix H, Figure H14 for boxplots) displays the percentage of voicing results for taps and trills in real and nonce words for all three speaker groups. For the FR and SPA groups, the tap was more voiced in nonce than in real words, while the RO group exhibited the opposite trend. The FR speakers voiced trills only slightly more in real than in nonce words, but both the RO and SPA groups voiced trills at almost equal rates in the two types of words.

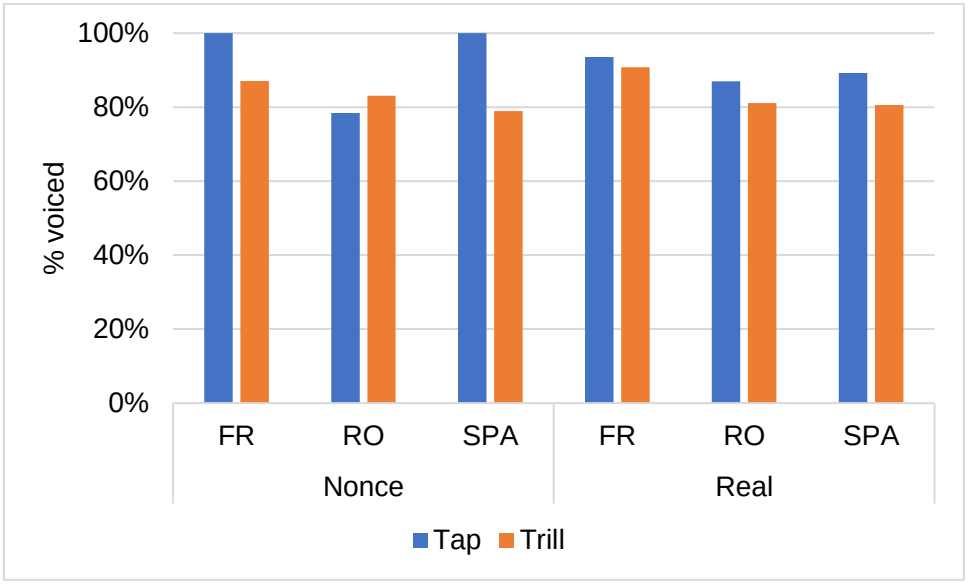

Figure 34. Average percentage of voicing of Spanish rhotics as produced by the L1 French-L2 Spanish (FR), L1 Romanian-L2 Spanish (RO) and native Spanish (SPA) groups in nonce and real words.

As with the Length variable discussed in the preceding section, separate generalized mixed effects models were fitted for Percentage of voicing for taps and trills. The first set of models explored taps and had Percentage of Voicing as the Response variable. M0 was the base model. M1 in this set had Language as the fixed effect, and Speaker[40] as the random effect. M2 had the same variables as the model just described, but, in addition to Language as the main effect, it also had Task. A model comparison revealed that the base model was the best fit (M0: AIC = 4754.300; M1: AIC = 4755.794; M2: AIC = 4755.965). However, it was not a significantly better fit (χ^2 = 2.5054, p = 0.286). This indicates that neither

[40] For the same reason mentioned for the Length analysis for the native Romanian speakers (in Romanian), Item was not included as a random effect.

Language nor Task were selected as significant; that is, there were no percentages of tap voicing differences among the three groups or between the two tasks.

The next set of four models investigated the Percentage of Voicing in trills as the Response variable. M0 was the base model. M1 in this set had Language as the main effect and Speaker as the random effect; M2 had Task in addition to Language, while M3 included Position. The fourth model included all of the fixed effects as in M3, but also explored the Language*Position interaction. A model comparison revealed that the base model was the best fit (M0: AIC = 5398.481; M1: AIC = 5401.453; M2: AIC = 5403.385; M3: AIC = 5404.212; M4: AIC = 5404.552). However, the base model was not significantly better (χ^2 = 1.0273, p = 0.598). These results suggest that none of the independent variables had a significant effect on the percentage of voicing of trills.

4.4 Unexpected realizations

Although the unexpected realizations were excluded from the statistical analyses (length and percentage of voicing), they are now going to be explored in more detail, since the hypotheses in this study do refer to the types of non-native-like productions. That is, in cases where participants did not produce the expected rhotic, what types of rhotics did they produce and are these related to their L1s? Recall that, in Hypothesis 1 (Chapter 3 Section 3.1), it was predicted that both learner groups would overproduce taps, overgeneralizing them to trill contexts, and, in addition, the L1 French group was also expected to produce some of their L1 French uvulars. In total, there were 460 unexpected realizations (out of 1,421 realizations in total; i.e., 32% of total realizations): 238 for the French group, 179 for

the Romanian group, and 43 for the native Spanish controls. Figure 35 (See Appendix G, Tables G16-24 for means and totals for each language group's non-native-like realizations by position, task, and word type) displays the overall non-native-like realizations of Spanish rhotics by the L1 French-L2 Spanish, L1 Romanian-L2 Spanish and native Spanish speakers. The figure displays participants' unexpected productions in tap contexts (on the left, first three bars) and in trill contexts (on the right, last three bars). As can be seen in the figure, in contexts where a tap was expected (left side of the figure), the RO group produced mainly trills and some fricatives (non-native-like realizations occurred in 10.19% of cases, as shown in Table 21, and Table G14 in Appendix G), while the FR group produced a variety of rhotics (non-native like 29.25% of the time, as in Table 21, and Table G13 in Appendix G): mainly trills, followed by fricatives, approximants, the French uvular, and very few Other realizations. As for the SPA group, in contexts where the tap was expected but not produced (only 1.95% of the time, as shown in Table 21, and Table G15, Appendix G), they produced trills 75% of the time and approximants in 25%[41] of cases. In contexts where the trill was expected but not produced, the RO group produced mainly the tap (62.22% of cases were non-native-like, as in Table G14, Appendix G), while the FR group produced mainly taps, followed by uvulars, fricatives, and some approximants (72.1% of cases were non-native-like, as displayed in Table G13, Appendix G). The SPA group, however, produced mainly fricatives, followed by taps, and some approximants (10.13% of cases were non-native-like, as can be seen in Table G15, Appendix G).

[41] Recall that, the native Spanish controls only had four instances of non-tap realizations in contexts that require taps. That is, in 3/4 tap-requiring contexts. Of these 4, two trills were produced in nonce words, one trill was produced a real word, and the approximant was also produced in a real word.

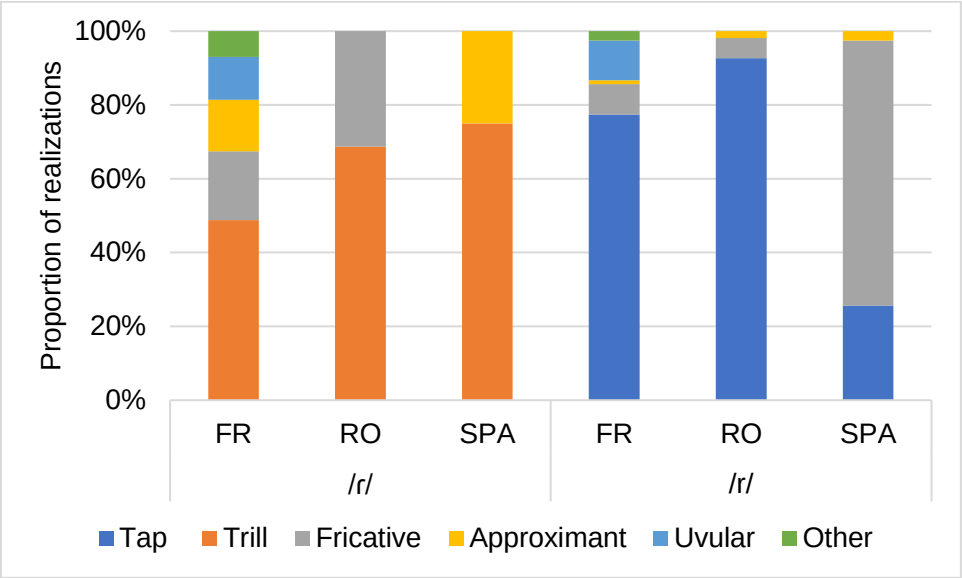

Figure 35. Overall manner of unexpected rhotics as produced by the L1 French-L2 Spanish (FR), L1 Romanian-L2 Spanish (RO) and native Spanish (SPA) groups in tap (left, first three bars) and in trill contexts (right, last three bars).

Recall from Chapter 3 that one of the RQs and its corresponding prediction referred to the types of non-native-like productions learners would have according to the position of the rhotic within the word. Figure 36 displays the results for WI on the left (first three bars) and for WM position on the right (last three bars). The expected rhotic is displayed along the bottom; although this is not necessary for WI position, as only the trill is expected, it is important for WM position, where both the tap and trill were expected depending on the word. As the figure shows, in WI position, where the trill is expected, the FR speakers' non-native-like productions included mainly taps, followed by uvulars, fricatives, and very few approximants, while the RO speakers produced overwhelmingly the tap (88.66%) instead of the expected trill. The SPA controls produced mainly fricatives, followed by taps, and lastly, very few approximants. In WM position when the tap was expected, FR speakers

produced various rhotics: mostly trills, followed by fricatives, approximants, uvulars, and some other realizations. The RO speakers produced trills overwhelmingly, and some fricatives. The SPA control group produced mainly trills, and some approximants. With regard to WM position when trills were expected, FR speakers' non-native-like realizations were taps, fricatives, uvulars, and other, while the RO speakers produced predominantly taps (98.48%). Lastly, the native SPA group produced somewhat more fricatives than taps. That several different variants were produced is unsurprising, given the amount of rhotic variation previously reported for both native speakers (e.g., Blecua, 2001; Colantoni, 2006 for Argentine rhotics; Navarro Tomás, 1971 for Peninsular Spanish; Quilis, 1993; Solé, 2002), and, partly due to L1 transfer effects, for learners (e.g., Amengual, 2016; Face, 2006).

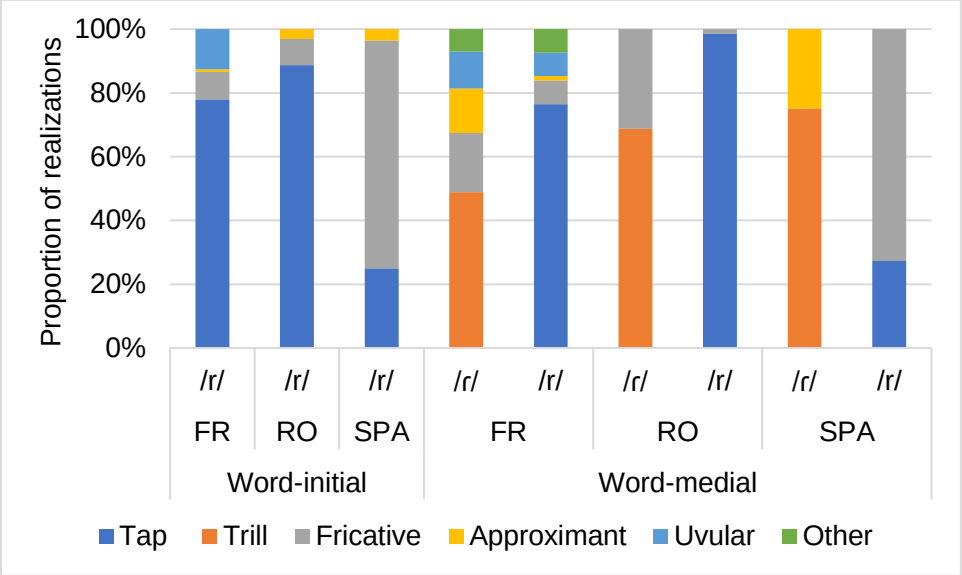

Figure 36. Manner of unexpected rhotics as produced by the L1 French-L2 Spanish (FR), L1 Romanian-L2 Spanish (RO) and native Spanish (SPA) groups in word-initial (the first three bars) and word-medial positions.

The third research question and its respective hypothesis referred to differences between the two tasks in this study. Figure 37 displays the overall types of unexpected realizations in each of the two tasks: picture description on the left (first five bars) and sentence reading on the right (last six bars). In the picture task when the expected rhotic was a tap, FR speakers' non-native productions included trills almost half the time, followed by fricatives, uvulars, and other, while the RO group produced mainly trills, and some fricatives. The absence of a bar for the SPA controls is due to the fact that this group did not have any unexpected realizations in this context. With regard to the expected trill, the FR group produced mainly taps, followed by uvulars, fricatives, and finally, approximants. In contrast, the RO group produced mainly taps (98.04%). The SPA control group produced mostly fricatives, followed by taps, and some approximants. In the reading task in situations where the tap was expected, the FR group produced unexpected trills, approximants, fricatives, uvulars, and other. The RO group produced predominantly trills and fricatives to a lesser extent. The SPA control group produced only trills and some approximants. Finally, when trills were expected, the FR group's production included mostly taps, and some uvulars, fricatives and other, while the RO group produced mainly taps, and very few fricatives and approximants. The SPA group produced taps and fricatives at comparable rates (42.86% versus 57.14%).

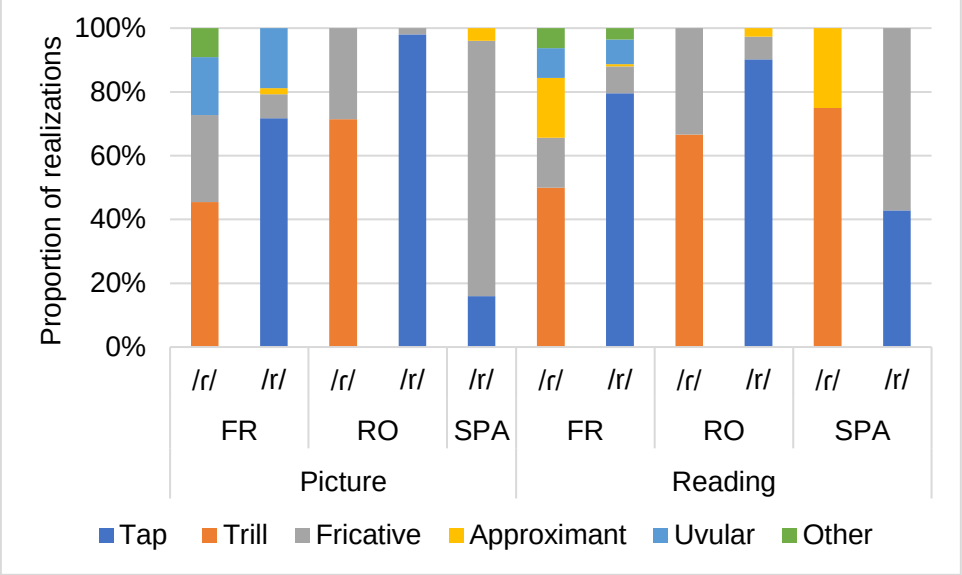

Figure 37. Manner of unexpected rhotics as produced by the L1 French-L2 Spanish (FR), L1 Romanian-L2 Spanish (RO) and native Spanish (SPA) groups in the picture description (the first five bars) and sentence reading (the last six bars) tasks.

Figure 38 displays the overall unexpected realizations in the nonce (first six bars) and real words (last six bars). For the nonce words when the context required taps, the FR speakers produced various non-target-like realizations, including trills, and to a lesser degree, fricatives, approximants, uvulars, and other realizations. The RO group produced trills and fricatives, while the SPA group produced trills and approximants. When trills were expected, the FR group produced mostly taps, followed by uvulars, approximants, and fricatives, while the RO group produced mainly taps. The native SPA group produced fricatives, and some taps. As for real words, when the tap was expected, FR speakers produced trills predominantly, followed by some approximants, uvulars, fricatives and some other realizations (4.35%). The RO speakers produced mostly trills and some fricatives, while the SPA speakers produced only trills. Finally, in contexts where trills were expected

in real words, FR speakers produced mainly taps, followed by uvulars, fricatives, and very few approximants. Again, the RO group produced largely taps, while the SPA speakers produced mainly fricatives, followed by taps, and minimal approximants.

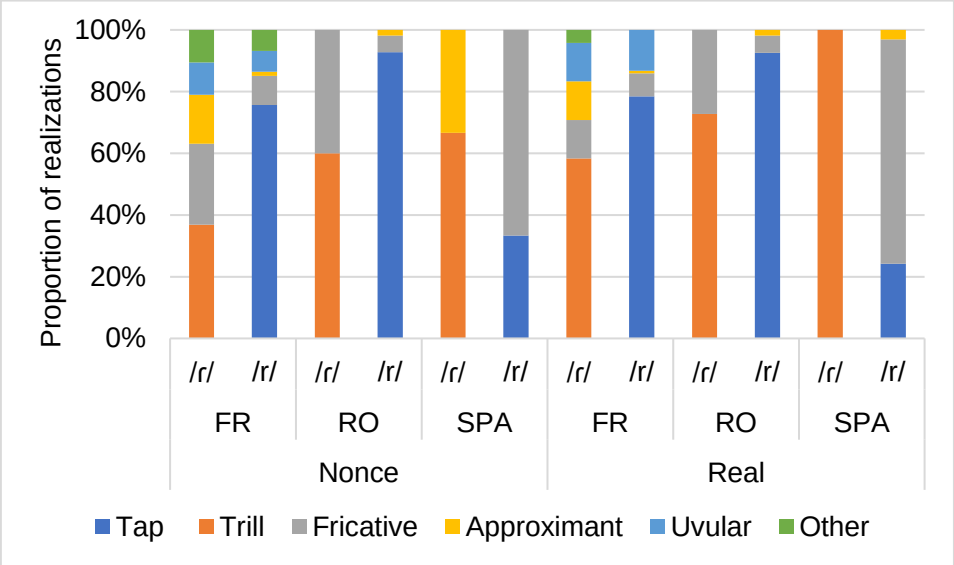

Figure 38. Manner of unexpected rhotics as produced by the L1 French-L2 Spanish (FR), L1 Romanian-L2 Spanish (RO) and native Spanish (SPA) groups in nonce words.

4.5 Trilling rates in L1 Romanian and L2 Spanish

Although the hypotheses do not make specific predictions about the correlation of trilling rates in Romanian and trilling rates in Spanish, this relationship, if it exists, can be revealing of L1 transfer, since L1 articulatory routines have been shown to affect L2 routines (Olsen, 2012, 2016). That is, it was of interest to explore whether the L1 Romanian speakers who produced more trills in their L1 Romanian also did so in their L2 Spanish. The plot in Figure 39 depicts the relationship between each participant's rate of trilling in L1 Romanian (x-axis) and their rate of trilling in L2 Spanish (y-axis) in contexts requiring trills in Spanish.

As the scatterplot shows, there was a positive correlation; that is, generally, the more participants produced trills in L1 Romanian, the more trills they produced in the appropriate contexts in L2 Spanish[42]. For example, RS01 produced trills the most in Romanian out of all ten participants (i.e., 27.08%), and also had the highest rate of trill production in L2 Spanish (81.48%). Participants RS04 and RS05, however, produced 0 trills in L1 Romanian and also had the lowest rates of trill production in L2 Spanish (11.11% and 7.69%, respectively). In order to test the strength of this correlation, a Pearson Correlation Coefficient was computed. The results of this Pearson Correlation test revealed a moderate positive correlation (Levshina, 2015, p. 119) between trilling rates in L1 Romanian and trilling rates in L2 Spanish ($R=0.706$; $n=10$, $p = 0.022$). That is, overall, those participants who produced more trills in their L1 Romanian also produced more trills in their L2 Spanish. This is consistent with Olsen's (2012, 2016) findings that L1 routines affect L2 articulation.

[42] This is not true for every single participant, however. For example, RS08 had low trilling rates in Romanian (4.35%), but produced trills in L2 Spanish at a higher rate (46.15%).

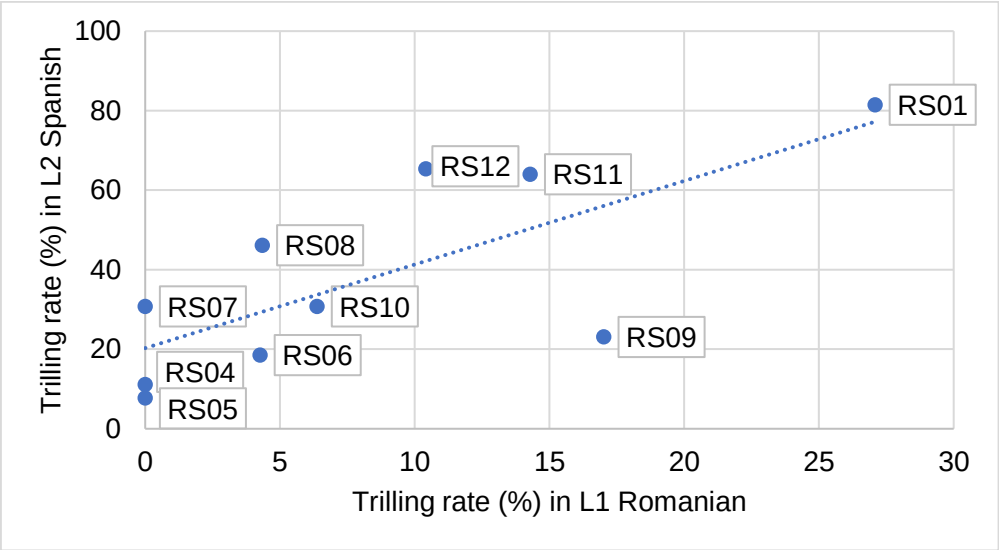

Figure 39. Scatterplot of trilling rates in L1 Romanian and L2 Spanish by L1 Romanian-L2 Spanish speakers.

4.6 Learner variables

Although the hypotheses here did not refer to individual learner variables, two of these individual variables are explored in this section in order to gain some insight into whether these factors influenced any of the participants' accuracy rates with either of the rhotics. As previously mentioned, both AoA (Munro et al., 1996; Ingvalson, McClelland, & Holt, 2011) and L2 use (Freed et al., 2004; Tremblay, 2009) have been found to affect L2 speech, and thus may provide a more complete picture of the results of this study. The plots in Figures 40 and 41 depict the relationship between each participant's AoA and their rate of tap and trill production in L2 Spanish, respectively. The plots in Figures 42 and 43 depict the relationship between each participant's use of Spanish/week and their rate of tap and trill production in L2 Spanish, respectively.

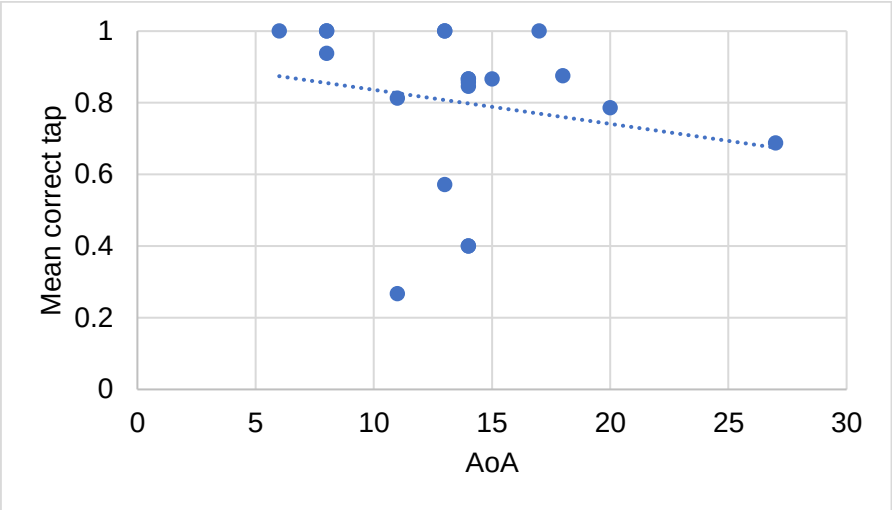

Figure 40. Scatterplot of AoA and mean tap rates for the L1 French and L1 Romanian participants.

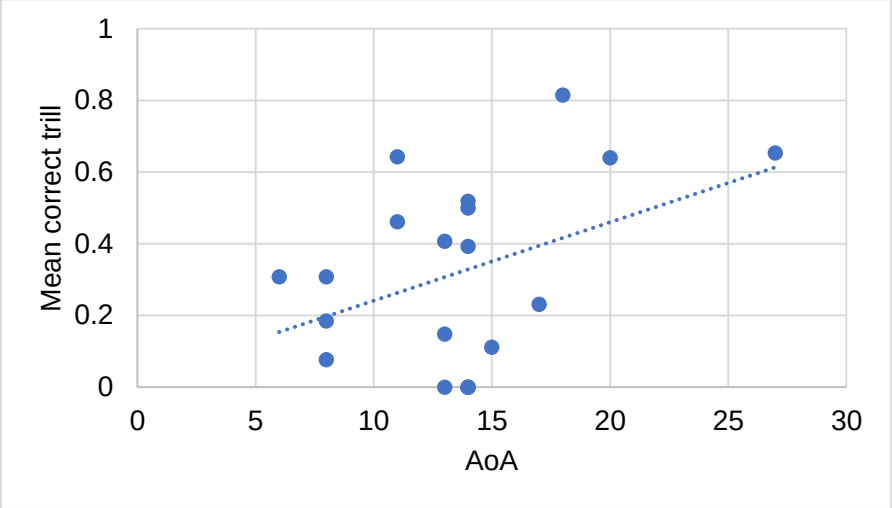

Figure 41. Scatterplot of AoA and mean trill rates for the L1 French and L1 Romanian participants.

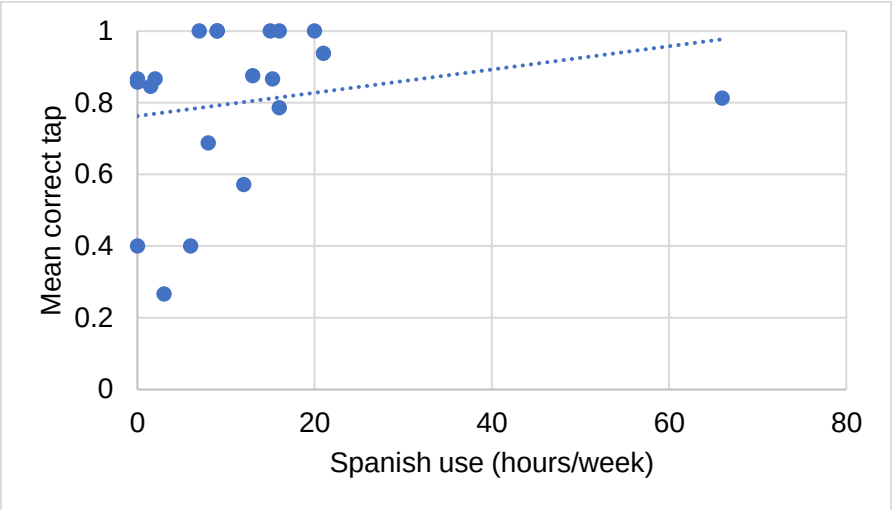

Figure 42. Scatterplot of Spanish use/week and mean tap rates for the L1 French and L1 Romanian participants.

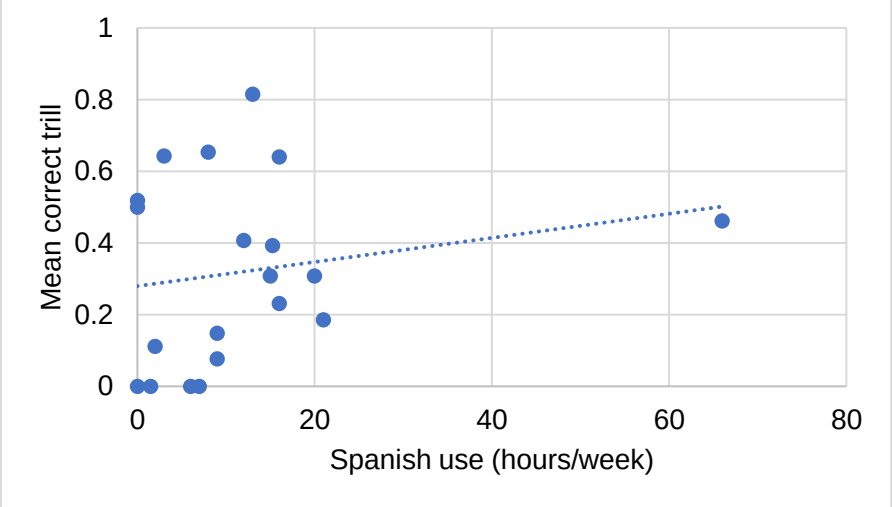

Figure 43. Scatterplot of Spanish use/week and mean trill rates for the L1 French and L1 Romanian participants.

The following correlations were run: mean accuracy and age of onset of acquisition (AoA) for taps and trills separately[43], and mean accuracy and hours/week speaking, writing

[43] The correlations for tap and trill accuracy rates were computed separately rather than combining accuracy rates, since, as was previously mentioned, trill accuracy rates were overall much lower than tap accuracy rates.

and listening to Spanish (taps and trills separately). In order to test the strength of these correlations, four Kendall's Tau Rank Correlations were computed (Levshina, 2015, pp. 132-133). The results of the first Kendall Tau test did not reveal a significant correlation between mean tap accuracy rates and the AoA of the participants (τ= -0.2300975; n=20, p = 0.189), suggesting that, it is not the case that as the rank of AoA increases, the rank of tap accuracy decreases. Likewise, the results of the second Kendall Tau did not show a significant correlation between the mean tap accuracy rates and the weekly hours of Spanish use (τ = 0.2530047; n=20, p = 0.136). This suggests that the mean tap accuracy rate does not depend on the number of hours per week spent speaking, writing or listening to Spanish. Regarding trill accuracy rates, similar results were obtained. That is, there was no significant correlation between the learners' mean trill accuracy rates and AoA (τ = 0.193093; n=20, p = 0.258) or between their trill accuracy rates and hours per week of Spanish use (τ = 0.1739156; n=20, p = 0.295). Overall, these results suggest that, while the learners varied regarding factors such as AoA and hours/week of Spanish use, these variables did not play a predictive role on either tap or trill accuracy for the specific participants in this study.

4.7 Individual Variability

In light of previous research noting that individual variation that occurs in speech (e.g., for L1 Spanish rhotics: Blecua, 2001; Colantoni, 2006; for L2: Borden et al., 1983; Munro et al., 1996; Riney & Flege, 1998), this section investigates more closely each learner's performance in order to better understand the results obtained. Figures 44 and 45 depict the mean accuracy rates for expected taps and trills, respectively, for each participant, with L1 Romanian-L2 Spanish speakers in yellow and the L1 French-L2 Spanish speakers in blue.

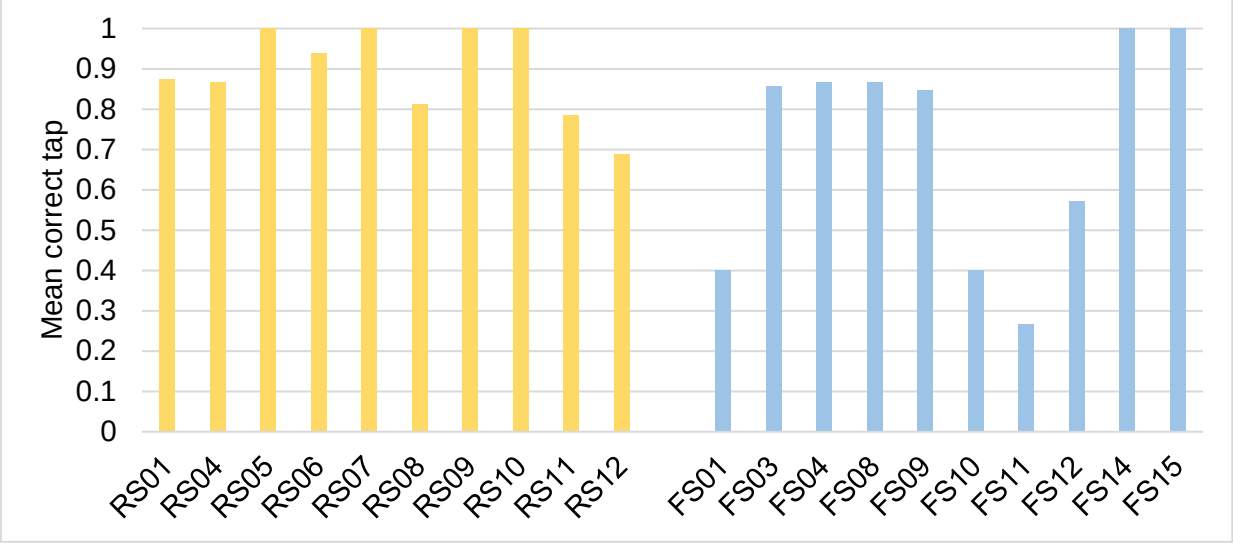

Figure 44. Individual learner results for tap realizations.

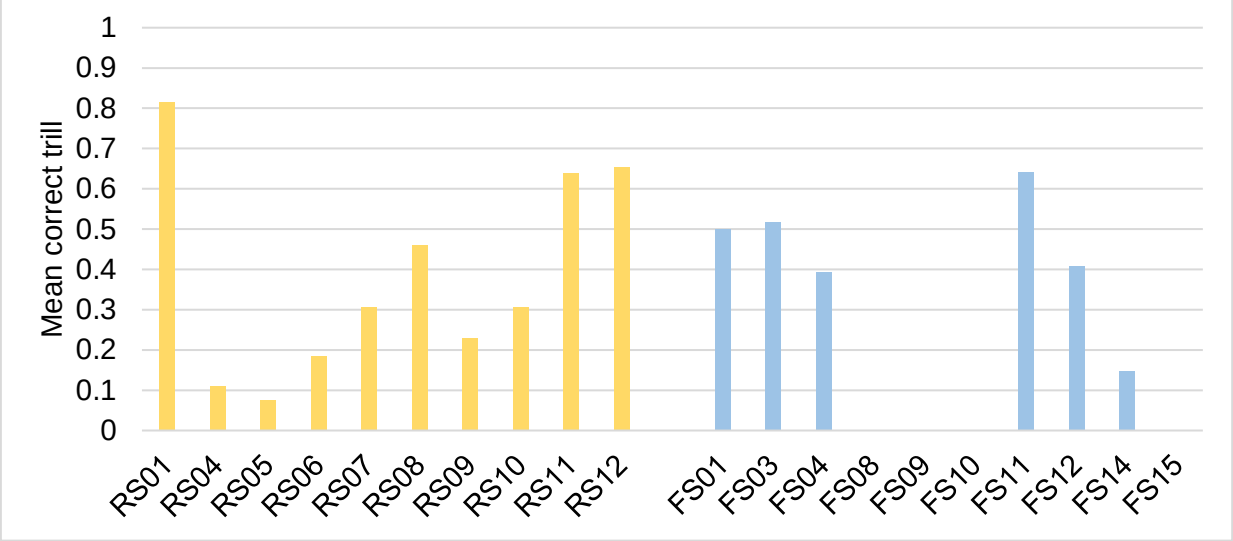

Figure 45. Individual learner results for trill realizations.

As can be seen in Figure 44, the L1 Romanian-L2 Spanish group did not exhibit a large degree of variability, ranging from 69% to 100%; however, the L1 French-L2 Spanish group showed a larger degree of variability, ranging from rates as low as 27% to as high as 100%. With regard to trill accuracy, Figure 45 shows more variability with trill than tap accuracy for both learner groups. Trilling was less successful than tap production for the participants

overall. Specifically, for the L1 Romanian-L2 Spanish group, the trill rates varied from 7.7% to 81%, while for the L1 French-L2 Spanish group, accuracy ranged from 0% to 64% production rates. Crucially, it is evident, that, while all of the L1 Romanian participants produced trills to some degree, there were several L1 French participants who did not produce any of the expected trills. Specifically, participants FS08, FS09, FS10 and FS15 did not produce any of the expected target trills. These four participants' personal profiles (AoA and hours/week using Spanish) do not seem to explain their 0% trilling rate. Regarding their individual variables, all four participants learned Spanish between 13 and 14 years old, and, and, regarding their weekly Spanish usage, while FS08 and FS09 spent 0-1.5 hours (respectively) using the language, FS10 and FS15 reported spending 6 and 7 hours, respectively. FS11, the participant who produced the highest rates of trilling (64%) in the L1 French group, learned Spanish at 11 years old and spent 3 hours per week using Spanish, which is less than FS10 and FS15. Regarding the L1 Romanian speakers, RS05 had the lowest trilling rates of this group (7.7%), followed by RS04 (11%). Their AoAs were 8 and 15 years old, respectively, and they spent 9 and 2 hours/week using Spanish, respectively. Finally, RS01, who had the highest trilling rates among the L1 Romanian group (81%) learned Spanish at 15 years old, and reported spending 13 hours per week using Spanish. It could be the case that, for RS01, the high trilling rates are due to spending a more substantial amount of time using Spanish, but this is not a clear trend for the other participants. To conclude, there does not seem to be anything specific in the personal information provided by the participants that may account for their rates of trilling.

4.8 Chapter summary

This chapter provided a detailed summary of the results of this study, beginning with the results for Romanian rhotics, and then turning to the results for the Spanish rhotics. For the Romanian rhotics, the results showed that, the most common realization was the tap, as has been attested in the somewhat limited literature (Chitoran, 2002; Radu, 2016; Savu, 2002). The other realizations were trills and to a much lesser degree, fricatives and approximants. Although trills were uncommon, they occurred significantly more word-initially than word-medially. The other independent variable, Task, did not have a significant effect on the type of rhotic produced by Romanian speakers. All in all, corroborating previous findings, it can be concluded that, in Romanian, the tap is the most common rhotic realization, at least in the contexts investigated in the present study; moreover, though trills do not occur often, when they do, they are more likely to occur word-initially than word-medially. In terms of the length of Romanian rhotics, trills were (unsurprisingly) longer than taps. Moreover, taps were significantly shorter word-medially than word-initially, and trills were significantly longer in the sentence reading than in the picture description task. Lastly, the results show that neither Position, nor Task had a significant effect on the percentage of voicing for taps or trills. With regard to the number of closures, most trills were produced with two, followed by one and finally, very few tokens with three closures.

Regarding Spanish rhotics, overall, across all conditions, the native Spanish controls produced significantly more of the expected rhotics than both the learner groups. Moreover, all participants produced the expected trills at lower rates than the expected taps, and all in all, all speaker groups produced significantly more of the expected rhotics in the Reading

than in the Picture task. In terms of the tap realization specifically, Language had a significant effect. Further analysis of this factor did not reveal differences between the two learner groups, or between the L1 Romanian speakers and the Spanish controls. However, it did reveal a significant difference between the L1 French group and the Spanish controls, with the former producing fewer of the expected taps than the latter. With respect to trills, the following factors and interactions were significant: Language, Task, Position and Language*Position. Overall, the Spanish controls produced significantly more of the expected trills than the L1 French speakers, as did the L1 Romanian speakers, although to a much lesser degree. All participants produced more of the expected trills in the Reading than in the Picture task, and, overall, they also produced more of the expected trills word-medially than initially. The Language*Position interaction was further explored, revealing that, unsurprisingly, the controls produced significantly more of the expected trills than both learner groups both initially and medially. With regard to the learners, crucially, against the predictions made, both the L1 Romanian and French speakers produced significantly more of the expected trills word-medially than initially. Furthermore, there was no significant difference between the learner groups in either position. Regarding the length of the target rhotics produced, none of the predictors was significant. For trill length, Task was not significant, but Position was, with significantly shorter segments word-medially than initially for all participants. Moreover, the Language*Position interaction was significant, with the L1 French speakers producing significantly longer trills word-medially than initially. However, the difference was only marginally significant. For the last correlate of rhotics, none of the independent variables had a significant effect on the percentage of voicing for either of the rhotics. Regarding the trill variant, two closures were the most

common for the Spanish controls and L1 French speakers, but, for the Romanian speakers, one closure was the most prevalent. Moreover, while the Spanish group produced trills with as many as six closures (not very frequent), the learner groups did not produce trills with more than four closures.

The closer analysis of the unexpected variants revealed that, whereas the L1 Romanian group produced overwhelmingly taps when trills were expected, the L1 French group produced a variety of rhotics, including predominantly taps, but also variants such as their L1 uvular rhotic, fricatives, and approximants. Additionally, there was a moderate positive correlation between rates of trilling in L1 Romanian and L2 Spanish. That is, generally, there was a tendency for speakers who produced more trills in their L1 Romanian to also produce more trills in their L2 Spanish. Further, the individual variables investigated, AoA and hours per week of Spanish use, did not correlate with either tap or trill accuracy in this study. In other words, it was not the case that those participants who had started learning Spanish at an earlier age were more target-like with either rhotic; likewise, it was not the case that those participants who spent more hours per week using Spanish were more native-like with taps and trills. Finally, the individual analysis showed a considerable amount of variability for trill accuracy rates specifically, but not for tap accuracy.

Chapter 5

Discussion and Conclusion

5 Overview

In this chapter, I discuss the results presented in Chapter 4 in light of the research questions and hypotheses in Chapter 3. The implications of this study's findings considering the previous research presented in Chapter 2 are also discussed. Then, I suggest some directions for further studies, and discuss the overall contributions and conclusions of the present study.

5.1 Romanian rhotics

One of the aims of the present study was to better characterize Romanian rhotics. Though this was not the primary goal, the results pertaining to the L1 Romanian rhotics are presented first, since they are useful for interpreting the L2 results, which are presented in Section 5.2. The results presented in Chapter 4 showed that, as reported in previous studies (Chitoran, 2002; Radu, 2016; Savu, 2012), the Romanian speakers produced more taps than trills in their L1 (85% versus 8% of the total data set). Moreover, while the task (picture description versus sentence reading) did not have a significant effect on the type of rhotic produced, the position of the rhotic in the word was significant. Specifically, participants produced the trill significantly more word-initially than word-medially (15% versus 1%). These tendencies confirm previous research, specifically Chitoran (2002) and Radu (2016), who found that the trill occurs predominantly word-initially rather than medially in Romanian. Though these tendencies were present in this study, it is important to remember that there

were very few trill productions overall; as such, these findings should be interpreted with caution.

The length of the rhotics was also analysed, in order to better characterize these segments in Romanian. Taps were significantly shorter word-medially than initially, but there were no significant positional length differences for the trill. However, the effect of task proved to be significant, with trills being significantly longer in the sentence reading than the picture description task. This finding is consistent with the fact that speakers tend to hyperarticulate in more formal speech leading to longer segment durations (Lindblom, 1989; Rafat, 2008). The last parameter analysed was the percentage of voicing. The statistical analyses performed did not reveal any significant effects of the independent variables on the percentage of voicing. Additionally, it was also discovered that most trills were produced with two as opposed to a single or three closures. Taken together, these results confirm the observations in previous work. Importantly, they provide experimental evidence that the primary Romanian rhotic is the tap, but that the trill can occur in specific positions, such as word-initially; that is, although it sometimes occurred word-medially in the present data, it was produced significantly more at the beginning of a word, a result which also provides support for observational claims made by Chitoran (2002), and preliminary claims by Radu (2016). The findings concerning length show that there are some positional effects for taps, but not for trills. Moreover, in cases where trills occur, they are longer in reading tasks than in picture description tasks.

Beyond the contribution to the literature on rhotics cross-linguistically more generally, these results, specifically those concerning manner, are relevant when interpreting the L1 Romanian group's L2 productions of Spanish rhotics, since some

predictions made specific reference to the Romanian speakers' L1 rhotics. In Section 5.2, I evaluate the hypotheses in light of the findings, as well as discuss how these relate to previous research.

5.2 Hypothesis evaluation

5.2.1 Overall production differences in the production of Spanish rhotics between learner groups

In this study, the first research aim was to determine what differences exist between the L1 Romanian and L1 French learner groups' production of Spanish rhotics. Further, it was of interest to explore whether the differences in L2 rhotic production, were they to exist, are related to the learners' L1s. It was predicted that there would be differences between the two learner groups, based on the articulatory characteristics and distributions of rhotics within the word in the two L1s. Specifically, the L1 French group was expected to have less difficulty producing the tap than the trill. This prediction was based on research showing that trill production generally lags behind tap production (e.g., Face, 2006; Johnson, 2008; Waltmunson, 2005) due, in part, to the articulatory difficulty of trills (Ladefoged & Maddieson, 1996; Solé, 2002). This articulatory difficulty often results in the overproduction of taps. In addition to overgeneralizing taps to trill contexts, it was predicted that the L1 French group's non-target-like realizations would also include some transfer errors from the L1 (Major, 1986). Specifically, this group was also expected to produce their L1 uvular rhotic instead of trills. Regarding the L1 Romanian speakers, I predicted that, in comparison with the L1 French group, they would have more target-like trill productions,

since trills do occur in Romanian, albeit infrequently. As such, these speakers were not expected to have difficulties with the articulatory routines involved in trill production. However, they were still expected to overgeneralize taps to trill contexts (Major, 1986). This was based on previous research, which has shown that taps are more prevalent than trills in Romanian (Avram, 1993; Chitoran, 2002; Radu, 2016; Savu, 2012).

Overall, the results presented in Chapter 4 provide some support for the differences predicted for the two learner groups. In the productions of Spanish rhotics, both learner groups were significantly more target-like in tap (i.e., word-medially intervocalically represented by <r> orthographically) than trill contexts (French: 71% for taps; Romanian: 90% for taps). Both learner groups had relatively high tap accuracy rates, and the differences between the two groups were not significant. Moreover, while there was no significant difference between the L1 Romanian group and controls, there was a significant difference between the L1 French speakers and the controls, who produced taps in almost all expected cases (98%). While the L1 Romanian group produced mostly trills as their non-target realization (69%), the L1 French group produced a variety of rhotics including trills (49%), fricatives (19%), approximants (14%), and the uvular rhotic (12%). With regard to trill production, the L1 Romanians were overall more accurate than their French-speaking peers, but both learner groups had relatively low target trill production rates (French: 13% for word-initial trills, 43% for word-medial trill; Romanian: 30% for word-initial trills, 47% for word-medial trills), especially word-initially (this will be further discussed in the following section, when position in the word is considered). It is worth noting that, not even the native Spanish controls produced trills in all of the anticipated contexts. This is not unexpected, as previous studies (e.g., Blecua, 2001; Colantoni, 2006; Navarro Tomás, 1971; Quilis, 1993;

Solé, 2002) have also found similar trends among native Spanish speakers. In cases where the expected trills were not produced, the native Spanish controls produced mainly fricatives. As for the learners, the L1 French group produced predominantly taps (77%), but also some of their L1 uvular rhotics (11%), fricatives (8%) and very few approximants. The Romanian group produced taps in almost all cases (93%). Taken together, these findings support previous research claiming that trill production lags behind tap production (Face, 2006; Johnson, 2008). This could be due to the articulatory demands of trills (Ladefoged & Maddieson, 1996; Solé, 2002). The fact that the L1 Romanian group overwhelmingly produced trills instead of target taps, while the L1 French group produced a variety of rhotics, provides some support for the notion that trills may not be articulatorily difficult for Romanian speakers in the same way as they are for French speakers. It is worthwhile to note that, during testing, some of the French participants explicitly mentioned aspects of Spanish pronunciation that they still found challenging, including the difficulty of 'rolling' their <r>s[44]. In these cases, the L1 French speakers in this study resolved this issue by producing either the less articulatorily difficult taps or their L1 uvular rhotic. That the L1 French group produced very few uvulars compared to taps (11% versus 77%) shows that some learning has taken place, and instead of exclusively transferring their L1 rhotics (as perhaps beginner learners would do), they are using a sound that exists in Spanish. The fact that the L1 Romanian group produced overwhelmingly taps in trill contexts could be due to the fact

[44] One other aspect that they noted was knowing which syllable was stressed, which is a well-known challenge that L1 French speakers face when learning an L2 that has lexical stress (e.g., Dupoux, Sebastián-Gallés, Navarrete & Peperkamp, 2008).

that taps are much more widespread than trills in their L1 (Chitoran, 2002; Radu, 2016; Savu, 2012), and they are transferring this L1 knowledge into their L2.

Furthermore, the results specifically address the learning problems presented in Chapter 2. Recall that the learning problems were as follows: for the L1 Romanian-L2 Spanish group, the task was to move from two L1 allophones to a different distribution of the same sounds, where they are contrastive (intervocalically); for the L1 French-L2 Spanish group, the learning problem involved acquiring two new contrastive L2 sounds, neither of which has a direct L1 counterpart. The results show that both learner groups were generally unable to realize the tap-trill intervocalic contrast that exists in Spanish, and instead produced the tap in most cases. This suggests that learning a new L2 contrast when both of the sounds are absent from the L1 may be just as difficult as learning a new distribution of two sounds that already exist in the L1.

These findings corroborate previous research, such as studies on the acquisition of the English lateral-rhotic contrast (e.g., Borden et al., 1983 for Korean speakers; Riney & Flege, 1998 and Sheldon & Strange, 1982 for Japanese speakers) showing that, although recategorizing L1 allophones as L2 phonemes is not an impossible task, it is generally difficult for learners. Likewise, though some research has shown that it may be less difficult to recategorize L1 variants as L2 phonemes versus learning new segments (e.g., Derrick & Gick, 2005 for the English lateral-rhotic contrast by native speakers of Beijing and non-Beijing Mandarin), the results obtained in this study did not find support for this, and instead echo research showing that learning L2 contrasts that do not have L1 counterparts may be just as difficult as L2 contrasts that do have corresponding L1 sounds (e.g., Flege et al., 1996 for Italian speakers acquiring English /θ/ and /ð/; Vokic, 2010 for L1 English speakers

acquiring Spanish /ɾ/ and [ð]). Lastly, though both learner groups generally displayed non-native-like patterns regarding manner, when they did produce the target rhotics, both L1 Romanian and L1 French speakers mastered length and percentage of voicing, as they did not differ significantly from Spanish controls in these respects. This finding differs from Rafat (2008), who found cases of hyperarticulation (Lindblom, 1989) in trill production. Rather, the results in the present study corroborate the findings of Colantoni and Steele (2008), who showed that not all parameters of rhotics are mastered simultaneously. In their study, too, manner lagged behind length and percentage of voicing in contexts other than word-medially intervocalically (for taps). For trills, and to a lesser extent taps, manner seems to be a relatively more difficult parameter than length or percentage of voicing, possibly due to the precise aerodynamic constraints that must be met in order to produce a successful rhotic, as well as due to the new articulatory routines involved in their articulation (especially trills).

Additionally, an interesting finding was that there was a moderate positive correlation between L1 Romanian trilling and L2 Spanish trill rates. In other words, it was generally the case that, those participants who produced more trills in their L1 Romanian also had higher trill rates in their L2 Spanish. In fact, two of the three participants who produced exclusively taps in Romanian produced the fewest trills in L2 Spanish. This finding suggests that L1 articulatory routines contribute to shaping L2 production, providing support for Olsen's (2012, 2016) claims. Specifically, as mentioned in Chapter 2 Section 2.3, Olsen (2012, 2016) found that those speakers who had a more retroflex-like pronunciation (rather than more bunched-like) of their native English rhotic more accurately produced taps in their L2 Spanish, at least in the earlier stages of acquisition. Taken together,

these results indicate that the role of subtle differences in L1 articulatory routines should not be underestimated, and should instead be more carefully considered and integrated in L2 speech production theories and models. Moreover, the findings in this study show the importance of evaluating L1-based influence at the individual level in addition to overall group effects or tendencies.

5.2.2 The effect of position within the word on Spanish trill realizations

The second research question asked whether the two learner groups would produce trills in word-initial (WI) and word-medial, intervocalic (WM) positions with equal accuracy. The general hypothesis was that differences were expected between the two learner groups regarding the position in the word. Since trilling is more difficult word-initially than word-medially intervocalically due to the steady stream of air flow already present from the preceding vowel, (Johnson, 2008; Lewis, 2004; Major, 1986), the L1 French group was expected to exhibit lower rates of trill production WI than WM. Conversely, the L1 Romanian group was expected to target trills in WI before WM position based on the distribution of trills in the L1, a result of positive transfer from their L1 (Major, 1986), since trills occur more in WI than WM position in Romanian. It is important to note that, regarding the production of the Spanish rhotics with respect to position, only the trill productions will be discussed, since only one tap context was tested (i.e., word-medial intervocalic) and, as such, a positional comparison cannot be made for this segment.

The results presented in Chapter 4 only partially support the hypothesis regarding position. Unsurprisingly, in both WI and WM position, the Spanish controls were significantly more target-like than the two learner groups (RO: 47% WM and 30% WI; FR:

43% WM and 13% WI; SPA: 94% WM and 87% WI). Importantly, against the predictions in this study, both the L1 Romanian and L1 French groups were significantly more target-like in WM than WI position, with no differences between the two groups, supporting previous findings regarding L2 Spanish trills (e.g., Johnson, 2008). This also corroborates previous research on other aspects of L2 production, such as Rogers and Alvord (2014), who found that learners were significantly more target-like with spirants word-internally than word-initially. Additionally, though there were no statistical differences regarding the percentage of voicing, learners produced significantly longer trills word-medially than initially, which is an unexpected result. This could be due to the fact that the word-initial trills were not realized in absolute initial position. While all of the word-initial rhotics in the reading task occurred in non-absolute word-initial position (since the target stimuli were embedded in a carrier phrase and thus, the target rhotic followed /n/ in all cases), in the picture description task, controlling for absolute versus non-absolute word-initial position was not possible due to the more spontaneous nature of the task. In particular, some productions had the target word at the very beginning of the sentence, while others had the word in other positions. Moreover, though participants were told to produce complete sentences, they did not always follow this instruction. An additional explanation may be that learners realize that trills were needed in some of the word-medial contexts, and they were trying to draw attention to the fact that they can produce them (as noted in Johnson, 2008, p. 134).

The results for manner for the L1 French group confirm the hypothesis regarding position, and provide support for trilling being more difficult word-initially than medially. As concerns the L1 Romanian group, who exhibited a similar trend to the L1 French

speakers, this was an unexpected result, since, in Romanian, trilling occurs more in word-initial than medial position and, on this basis, it was predicted that this group would be more target-like in WI position. Though this cannot be said with certainty, it is possible that, for the L1 Romanian group, articulatory demands are more important than L1 phonetic context, causing these speakers to target word-medial trills as a result of basic articulatory principles. That is, since successful trill production involves a rapid stream of airflow, intervocalic position provides an advantage given that the steady airflow is already present from the previous vowel (Major, 1986). Another potential (or additional) explanation may be that the tap-trill contrast is transparent in the orthography word-medially, but word-initially, there is no orthographic indication that the <r> is a trill (this will be mentioned in more detail in the following section, when task differences are discussed). For this reason, learners may be producing more trills word-medially than -initially. The results or length also provide some support for this, since the learners' trills were significantly longer word-medially than -initially, suggesting that they are perhaps more aware that trills are sometimes needed medially but not initially. Lastly, that the Spanish controls also produced slightly more of the expected trills word-medially than word-initially possibly provides further support for the notion that there are more articulatory demands word-initially than there are medially.

5.2.3 The effect of task on Spanish rhotic realizations

Finally, the last research question investigated whether there would be differences between the two learner groups in the two tasks differing in formality, namely a sentence reading versus a picture description task. Overall, it was predicted that the two learner groups would

behave similarly across the two tasks. That is, more target-like rhotic productions were expected in the sentence reading than in the picture description task, because transfer errors are more common in more casual tasks (Major, 1986) and read speech may be more target-like due to reduced processing demands (Colantoni, et al., 2015, p. 110). Additionally, orthography was also expected to have some effect on rhotic accuracy. Recall that, in Spanish, orthography makes the tap-trill contrast clear word-medially (<r> versus <rr>), but not word-initially (only <r>), which was predicted to result in more target-like productions word-medially than -initially (Waltmunson, 2005, p. 37) for the L1 French group specifically, but not the L1 Romanian group, whose L1 has trills word-initially.

The results obtained in this study provide partial support for this hypothesis, specifically regarding trills. As was shown in Chapter 4, regarding tap productions, the L1 French group was slightly more target-like in the picture than in the reading task (77% versus 68%), while the L1 Romanian group was slightly more target-like in the reading than in the reading task (91% versus 87.7%). The native Spanish controls were only slightly more target-like in the picture than in the reading task (100% versus 97%). Despite these trends, none of these differences were statistically significant. With respect to the other parameters of taps, there were no differences between the two tasks for length or percentage of voicing.

For trills, although there were no task differences for either length or percentage of voicing (for the target realizations only), some interesting patterns emerged regarding manner. Overall, both learner groups were significantly more target-like in the reading than in the picture task, specifically word-medially (FR: 45% versus 32%; RO: 55% versus 13%). The same tendency was observed word-initially for the L1 Romanian group, although to a lesser degree (33% versus 21%). The L1 French group, however, had comparable

trilling rates word-initially across the two tasks (13% versus 13%). Lastly, the native Spanish controls produced more of the expected trills in the reading than in the picture task, both word-initially and medially (WI: 92% versus 78%; WM: 97% versus 85%). The between-task differences were statistically significant, with participants being significantly more target-like in the reading than in the picture task overall.

These findings were expected, since transfer errors are more prevalent in tasks that are more casual (Major, 1986, 1987; Tarone 1979, 1982, 1983). Although both of the tasks in this study were to some degree structured and performed in a laboratory setting, the picture description task is arguably less formal than the sentence reading, as read speech can often be deemed hyperformal (Colantoni, et al., 2015, p. 100). Additionally, read speech may also present more target-like variants (Tarone, 1979). This may be the case for Spanish the tap-trill contrast specifically, since word-medially, the orthography is transparent regarding this contrast and this contrast is absent from both French and Romanian. This is consistent with the notion that orthographic input may be beneficial for learners in cases where they are acquiring L2 contrasts that do not exist in the L1 (Young-Scholten, 2015, p. 96). Moreover, as mentioned in previous research (e.g., Colantoni & Steele, 2006; Face, 2006; Olsen, 2016), word-initially where <r> is realized as a trill, orthography may be detrimental, since all three languages under investigation use the same grapheme (i.e., <r>) to represent different sounds, leading to non-target-like productions. That participants were more target-like in the reading than in the picture task begs the question of whether trilling rates in previous studies may have overestimated learners' success with trills. That is, if there are generally higher trilling rates in reading tasks, and previous research on Spanish rhotics has used exclusively reading tasks, it may be the case that the same participants

would be significantly less successful with trills in more spontaneous speech, where task formality is diminished, and orthography is eliminated.

5.3 Contributions of the present study to the research on rhotics

Overall, this study has expanded the literature on rhotics, both as produced in a native and in a second language. One of the major contributions is a better characterization of Romanian rhotics, with acoustic analyses to support the findings of previous authors. To date, the literature on this topic has either been observational in nature, such as Chitoran (2002), or vaguely related to characterizing rhotics, such as Savu (2012), whose primary focus was epenthetic vowels. The present study has shown that, in the two linguistic contexts investigated here (i.e., word-initially before a vowel and word-medially intervocalically), while taps were by far the most common Romanian rhotic, trill productions were also attested, though much rarer. When trills did occur, they did so significantly more word-initially than medially, and they were significantly longer in read versus spoken speech.

An additional finding that is also valuable is that there is individual variability at the level of rhotic production. As we saw in Chapter 4, there were some participants in this study who produced only taps, regardless of task or position in the word, while others produced both taps and trills. Rhotic variation has also been attested in English, as shown in Derrick and Gick (2011), who deemed this "unconditioned, categorical, subphonemic variation" (p. 316). Variation may also occur in Spanish rhotic production (Blecua, 2001); however, the tap and trill are not used interchangeably the way they can be in Romanian,

since in Spanish, they are quasi-contrastive, while in Romanian, they do not affect meaning. Individual variation was also observed in the rhotic accuracy rates among the L2 speakers, specifically more variability with trill than tap accuracy (L1 French-L2 Spanish: 0%-64%; L1 Romanian-L2 Spanish: 7.7%-81%). That the three language groups exhibited different amounts of variability is an interesting finding. Specifically, compared to the L1 Romanian and native Spanish groups, the L1 French group displayed a large amount of variability not only regarding tap and trill accuracy rates (Section 4.7), but also in terms of the variants they produced in the unexpected realizations, as shown in Section 4.4. This may be due to several factors. First, although as shown in Section 3.2.2, the two learner groups did not differ with regard to the cloze test scores, they did differ with regard to their oral proficiency scores. The fact that the two groups received different accentedness ratings may explain the variability, at least in part. An additional reason could be due to the input the participants had received. Regarding the native speaker input, the L1 Romanian group is more likely to have been exposed to Spanish from Spain, since all participants reported having learned Spanish in Romania (at the Instituto Cervantes, or beginning in elementary school), where the main dialect is Castilian Spanish. On the other hand, the L1 French group may have been exposed to a variety of dialects. Some participants had begun learning Spanish in France, and the majority had also spent some time in Spain. One participant had been exposed to various dialects (e.g., Argentine Spanish) while living in Quebec, and the students that were on exchange at the University of Toronto had also been exposed to various dialects. Additionally, non-native input could have had an effect. Specifically, the only non-native input the L1 Romanian group would have received is from other native

Romanian speakers. Conversely, the L1 French group had received input from speakers of various L1s, ultimately resulting in more variability.

Furthermore, this study contributes to the L2 phonology research, as it investigated two learning problems that speakers of the same L2 may face. The results revealed that, it may be just as difficult to recategorize already existing L1 allophones as L2 phonemes, as it is to learn two new phonemes that do not have L1 counterparts. In investigated two new language pairings (i.e., L1 Romanian-L2 Spanish and L1 French-L2 Spanish), the study also makes contributions to the L2 Spanish rhotic research, since previous research has focused primarily on speakers whose first language is one other than English. As mentioned, numerous studies have investigated rhotics in L2 Spanish, but it is not clear how applicable the findings are to speakers of other L1s. In this study, it has been suggested that, in addition to the articulatory difficulty that trills pose, the distribution of trills and taps in an L1 may also contribute to non-target-like productions when trills are expected. Specifically, it was shown that, for L1 French speakers, trills, and to an extent, even taps, are difficult from an articulatory point of view, which was also highlighted by some of the French speakers who expressed the challenge they face with "Spanish <r>s". For Romanian speakers, however, trills are not inherently difficult to produce; rather, the results here suggest that, because taps are the most prevalent variant in the L1, the Romanian speakers tend to overgeneralize this to trill contexts in L2 Spanish. That is, they have difficulty recategorizing two existing L1 sounds from allophones to (quasi)contrastive segments in the L2. Although the L1 Romanian speakers had slightly higher accuracy rates than the L1 French speakers, these differences were small, and they were not significant when position was taken into account for trills. To conclude, these findings suggest that, both learning problems are equally

challenging, at least in what concerns articulatorily complicated segments such as trills. That is, it is equally difficult to learn two new sounds and to learn a new distribution involving already existing L1 sounds.

From a methodological perspective, the contribution of the present study is that a reading task was used, as much research on rhotics has up to date, but, additionally, it also included a picture task in order to elicit more spontaneous data. Though a drawback of the picture task was that not all of the target stimuli were produced by the speakers, as was shown, there were indeed task effects, for the trill specifically. This finding highlights the importance of using more spontaneous tasks in L2 speech studies, in order to obtain a more complete picture of L2 production accuracy.

Lastly, the overall low rates of trilling in the learners' productions highlight the need for more explicit classroom instruction regarding Spanish rhotics. In working with learners of Spanish of different backgrounds, I have often heard students' frustration at their difficulty to "roll their <r>s". Evidently, Spanish does not have one, universal, "rolled <r>" and there are several subtleties that should be addressed in the classroom, from the articulatory aspects of rhotic production, to the phonological distribution, to the orthographic contrast that is realized word-medially intervocalically. For example, even though Romanian speakers can "roll their <r>s", this is not enough to approach a native-like rhotic pronunciation. Research on explicit instruction is growing, with some findings suggesting that perceptual abilities can be improved, but production ones cannot (e.g., Carlet & Souza, 2018), others showing that explicit classroom instruction can improve both comprehensibility and intelligibility in an L2 (Bouchhioua, 2017), and still others suggesting that explicit instruction may lead to categorical effects, but implicit instruction

may be necessary for more fine-grained phonetic characteristics to develop (Shea & Curtin, 2011). Regarding Spanish rhotics specifically, there is a possibility that less proficient speakers may not be aware of the difference between the two sounds at all, or that they are aware, but have difficulties producing the trill (Johnson, 2008, p. 87). While textbooks intended for pronunciation courses do make explicit reference to the contrast, as well as describe the complementary distribution of taps and trills in other contexts (e.g., Morgan, 2010), textbooks intended for Spanish language sequence courses do not (e.g., Dorwick, Pérez-Gironés, Becher, Elliott, Zapata, Rogers, & Santos, 2014). Rafat and Perry (2019) provide some practical strategies for teaching these sounds. They highlight the need or explicitly discussing the contexts in which the tap and trill occur and providing ample written and spoken examples, as well as showing students minimal pairs (e.g., *pero* 'but' versus *perro* 'dog') in order to further highlight the differences between the two sounds. The authors suggest an activity involving the written presentation of words with the rhotics in different positions in the word, as well as the instructor signalling which sound is produced in each instance. The students' task would be to identify any patterns they notice with regard to tap versus trill production. Rafat and Perry (2019) note that the goal of this activity is to raise students' awareness that there are inconsistent grapheme-phoneme mappings and that the same grapheme in different positions sometimes corresponds to different phonemes. In addition to the instructional suggestions made by Rafat and Perry (2019), learners should also be given sufficient instruction and practice regarding how to physically produce each rhotic, since, as previously discussed, these sounds are among the most difficult. Specifically, students should be given overt instruction regarding tongue

placement and movement. Importantly, learners would also benefit from explicit feedback from instructors (Borden et al., 1983).

The lack of findings regarding certain variables in the present study arguably also provide some insights into the nature of speech learning. Recall that, although not the focus of the study, variables such as AoA, hours/week of Spanish use, and Word Type (nonce versus real), were explored in order to investigate whether they had any effect on expected rhotic productions among the learners. As the results presented in Chapter 4 Section 4.2 (for rhotics in Romanian), as well as 4.3 (for rhotics in L2 Spanish) showed, there was no effect of Word Type on manner, length or percentage of voicing for either Romanian or Spanish, despite the fact that some previous research (e.g., Frost et al., 1988) has found an effect; however, the discrepancy could be due to the fact that most previous research on nonce words has been concerned with topics such as lexical access, phonological memory, and/or word learning, not with L2 production specifically.

With respect to the individual variables, though previous studies have shown that AoA (Munro et al., 1996; Ingvalson, et al., 2011) and L2 use (Freed et al., 2004; Tremblay, 2009) affect L2 speech, the present study failed to find this effect. Regarding the findings for AoA, which were presented in Section 4.6, it could be the case that, while it does not have an effect on the production of the particular segments investigated here (i.e., L2 Spanish rhotics), particularly for the learners involved in this study, it continues to affect overall speech and foreign accent, as previous research has shown. Though neither the AoA and tap correlation, nor the AoA and trill correlation, was statistically significant, the former correlation seemed to show a negative slope, while the latter was positive. That is, while tap accuracy tended to decrease as the AoA increased, trill accuracy seemed to increase as the

AoA increased. This merits some further discussion. Upon more careful analysis, it became clear that, for the L1 Romanian group, the three participants with the highest AoA (RS01, AoA=18; RS11, AoA=20; RS12, AoA=27) displayed the highest trill accuracy rates (RS01 = 81%; RS11 = 64%; RS12 = 65%). When these participants' data was removed from the correlation analyses, the slope for trill accuracy indeed became negative, showing that, as the AoA increased, participants' trill accuracy decreased. A possible explanation for these three learners' trends could be that the participants who began learning Spanish later in life, (and thus have a higher AoA) were made explicitly aware of the existence of the trill. If this is the case, it could be that, once they had learned this sound, they became more likely to pay more attention to producing it, while being less careful about producing the tap. Further, participant RS01 reported having spent 6 months in Spain, possibly receiving native speaker input that improved this participant's trill accuracy. As for the other participants, RS11 also spoke Bulgarian as a foreign language, and RS12 was taking independent courses (i.e., not through university) on Spanish. Regarding RS11, in learning both Bulgarian and Spanish, this participant could have had an increased level of metalinguistic awareness, perhaps allowing her to consciously reflect on aspects of her L2(s), including her production of the trill. As for RS12, it is possible that the Spanish language courses she was taking taught the trill in a more explicit manner in order to draw more attention to it.

Regarding the results for L2 use (Section 4.6), it is possible that the number of hours/week spent speaking, writing and listening simply are not indicative enough of L2 use. As previously mentioned, this particular variable was used because it was the only one for which participants were consistent in reporting their usage. However, instead of this measure, it might be more useful to focus on the proportion of the L2 used in social settings,

as well as at home, school and/or work separately (Borden et al., 1983). In addition, though the amount of weekly Spanish use was not significant for tap and trill accuracy specifically, as mentioned in Chapter 3 Section 3.2.2, this variable did significantly affect two of the proficiency scores used in this study. Specifically, those participants who reported using Spanish more per week also scored the highest on the cloze test, as well as received the highest oral proficiency scores from the native speaker judges. Thus, while the amount of weekly Spanish use did not seem to affect the accuracy with the segments investigated in this study, this individual variable may have consequences for more global aspects of L2 speech and learning.

5.4 Future work and conclusions

While this dissertation provides an in-depth characterization of Romanian rhotics and an analysis of the L2 production of Spanish taps and trills, it also has some limitations and therefore opens lines for future research. Firstly, the present study focused on only two environments (i.e., word-initial and word-medial intervocalic), but it would be worthwhile to conduct studies that include other environments where the tap and trill are in complementary distribution, such as after /n/, /l/ or /s/ for the trill, and after a tautosyllabic consonant, where the tap occurs. This would be revealing of whether learners know that certain environments require one rhotic versus the other, and that these environments do not overlap (as they do intervocalically). Second, for L1 Romanian learners specifically, a further study could involve an experiment with the Spanish trills in word-initial position (as in the present study), but in more emphatic speech in addition to non-emphatic speech. Since Romanian taps are more prevalent and generally occur in non-emphatic speech (Schulte,

2013), it is possible that, if Romanian speakers are transferring their L1 level of encoding, they would produce more of the target word-medial trills in emphatic versus non-emphatic contexts.

One limitation of the present study, as shown in Section 3.2.2, is that, though the L1 French and L1 Romanian groups did not differ with regard to their overall self-assessments, or their average cloze test scores, they did differ regarding their oral proficiency, with the L1 French group receiving lower average accentedness scores than the L1 Romanian group (1.74 versus 2.92). Every effort was made to match the learners on all levels; however, this was not possible for oral proficiency specifically, which was, in part, due to the difficulty of finding French participants, and in part due to Romanian overall being more phonologically similar to Spanish than French is. In addition, it could be the case that the judges' experience with L2 speech may have influenced their scores (e.g., Gass & Varonis, 1984). Moreover, though the accentedness scores used in this study are perhaps the most relevant measure of proficiency, since they target L2 phonology and phonetics (Colantoni et al., 2015, p. 89), it is unclear exactly which variables (or which combination of variables) affected native speakers' judgements of accentedness. Specifically, the judges' scores could have been influenced by a variety of factors, including segmental errors, prosodic accuracy, fluency (e.g., Riney, Takagi, & Inutsuka, 2005) and speech rate (e.g., Munro & Derwing, 2001), which may or may not be provide a complete picture of L2 proficiency. To address some of these issue, future studies could measure oral proficiency using additional tests, such as the American Council on the Teaching of Foreign Language's (ACTFL) Oral Proficiency Assessment (OPI), which, among other aspects, also measures phonetic and phonological features.

In addition, future studies could focus on learners of other proficiency levels in the same languages as the present study. This would provide a more comprehensive picture of the stages of rhotic development in L2 Spanish. Investigating developmental sequences involving Spanish rhotics among L1 Romanian and L1 French speakers may be beneficial for comparing their developmental patterns to those of L1 English speakers in order to evaluate the generalizability of previous findings. For example, Face's (2006) cross-sectional study of L1 English-L2 Spanish speakers showed that, while even the advanced learners had relatively low levels of accuracy, they overgeneralized the tap to trill contexts instead of transferring their L1 rhotic the way less proficient learners did. This would be an interesting point of comparison with the L1 French speakers in the current study, since these learners transferred their uvular French rhotic in addition to producing taps when trills were expected. Because Spanish trills pose such difficulty for learners, other L1 populations in addition to English, French and Romanian should also be investigated in order to gain a better understanding of how speakers of different L1s acquire these sounds and/or whether there are universal patterns of L2 development in rhotic acquisition. Lastly, it also remains unclear whether speakers at more advanced stages (for example, graduate students or students having lived abroad) have higher rhotic accuracy rates than those proficiency levels previously investigated.

Further, the present study focuses on production, but it does not reveal what occurs in the perception of Spanish rhotics. A large body of research has focused on the perception-production link (e.g., Riney & Flege, 1998; Sheldon & Strange, 1982), but, in the case of Spanish rhotics, it has yet to be established (or investigated, to the best of my knowledge) whether learners do in fact perceive a difference between taps and trills, especially in cases

where they are unaware that there is a contrast (i.e., intervocalically). Given that, to my knowledge, there is no perception research in this area, it is not obvious if learners simply cannot produce the contrast, or if they are oblivious to the fact that a contrast exists at all, and thus are not aiming to produce one. For example, as concerns the speakers in this study, in PAM-L2 (Best & Tyler, 2007) terms, there could be a case of Single Category Assimilation, where the learners are unable to perceive a distinction between the tap and trill, thus assimilating them both to the same native category, ultimately leading to poor discrimination. With the methodology and results of the present study, it is not possible to conclude this, but a perception component could provide some insight into this. Additionally, because the tap-trill contrast is generally difficult for learners, it would also be worthwhile to conduct further training studies (e.g., Herd, Jongman & Sereno, 2013), where learners could be explicitly trained on the contrast, since some previous research (e.g., Rato, 2014 for perceptual training) has shown that training can have positive effects on perceiving and producing contrastive L2 sounds.

5.5 Conclusions

The present study investigated the L2 production of Spanish taps and trills by native speakers of L1 Romanian and L1 French, arguing that, trill production is difficult for these two learner groups for different reasons. As previous research has shown, trills are generally difficult segments, and the difficulty of producing them in an L2 can be further exacerbated by the different distribution of rhotics in the L1. It was shown that, for L1 French speakers, the difficulty of producing the Spanish trills is due mainly to articulatory constraints; on the other hand, for L1 Romanian speakers the difficulty may primarily be due to the distribution

of taps and trills in the L1. That is, because the tap occurs more than the trill in Romanian, this learner group may overgeneralize taps to trill contexts in L2 Spanish. The contributions of this study are two-fold. First, it offers a fine-grained acoustic analysis and characterization of Romanian rhotics in two linguistic contexts, providing empirical evidence for observations that other authors have made in passing. That is, the results here support claims that the tap is the predominant variant in Romanian; when the trill does occur, it does so mainly word-initially. Second, this study is a comprehensive account of L2 Spanish tap and trill productions by L1s other than English, particularly expanding the literature on two learning scenarios: (1) both sounds exist in the L1, but have different distributions than they do in the L2, and the sounds are phonetically similar in both the L1 and L2; (2) neither sound exists in the L1, and instead, a highly different sound is used (i.e., the French uvular). To some extent, this study also makes a methodological contribution, having implemented a more spontaneous task when investigating L2 rhotics than other studies have previously used. This study creates a foundation for future research on L2 rhotics and learning problems involving L2 sounds more generally, with the hope of deepening our understanding of L2 phonology and expanding the language groups that are investigated.

References

Adler-Bock, M., Bernhardt, B. M., Gick, B., & Bacsfalvi, P. (2007). The use of ultrasound in remediation of North American English /r/ in 2 adolescents. *American Journal of Speech-Language Pathology*.

Amengual, M. (2016). Acoustic correlates of the Spanish tap-trill contrast: Heritage and L2 Spanish speakers. *Heritage Language Journal*, *13*(2), 88-112.

Archibald, J. (1998). Second language phonology, phonetics, and typology. *SSLA, 20,* 189-211.

Avram, A. (1993). Cercetări experimentale asupra consoanelor lichide din limba română [Experimental research on the liquid consonants in Romanian]. *Fonetică și dialectologie, 8,* 8-20.

Baltazani, M., & Nicolaidis, K. (2013). The many faces of /r/. In L. Spreafico & A. Vietti (Eds.), *Rhotics: New data and perspectives,* (pp. 125-245). Bozen-Bolzano University Press.

Bassetti, B. (2008). Orthographic input and second language phonology. *Input matters in SLA*, 191-206.

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

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.

Beebe, L. M. (1987). Sociolinguistic variation and style shifting in second language acquisition. *Interlanguage phonology: The acquisition of a second language sound system*, 378-388.

Best, C. T., Traill, A., Carter, A., Harrison, K. D., & Faber, A. (2003). ! Xσυ click perception by English, Isizulu, and Sesotho listeners. In *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona 3-9 August 2003*.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. *Language experience in second language speech learning: In honor of James Emil Flege*, *1334*, 1-47.

Blecua, J. M. (1999). *Filología e informática: nuevas tecnologías en los estudios filológicos* [Philology and information technology: new technologies in philological studies]. Seminario de filología e informática Dpto de filología española Univ. autónoma de Barcelona.

Blecua, B. (2001). Las vibrantes del español: manifestaciones acústicas y procesos fonéticos [The vibrants in Spanish: acoustic manifestations and phonetic processes]. *Barcelona: Universitat Autònoma de Barcelona dissertation*.

Boersma, P., & Weenink, D. (2016). Praat: Doing phonetics by computer (Version 6.0.14).

Bongaerts, T., Mennen, S., & Slik, F. V. D. (2000). Authenticity of pronunciation in naturalistic second language acquisition: The case of very advanced late learners of Dutch as a second language. *Studia linguistica*, *54*(2), 298-308.

Boomershine, A., Hall, K. C., Hume, E., & Johnson, K. (2008). The impact of allophony versus contrast on speech perception. *Contrast in phonology: Theory, perception, acquisition*, *13*, 145-172.

Borden, G., Gerber, A., & Milsark, G. (1983). Production and perception of the/r/-/l/contrast in Korean adults learning English. *Language Learning*, *33*(4), 499-526.

Bouchhioua, N. (2017). The effects of explicit pronunciation instruction on the comprehensibility and intelligibility of Tunisian EFL learners. *International Journal of Research*, *6*(3), 73-88.

Bradley, T. (2001). *The phonology and phonetics of rhotic duration contrast and neutralization* (Doctoral dissertation, Ph. D. Dissertation, The Pennsylvania State University).

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*(3-4), 155-180.

Bundgaard-Nielsen, R. L., Best, C. T., Kroos, C., & Tyler, M. D. (2012). Second language learners' vocabulary expansion is associated with improved second language vowel intelligibility. *Applied Psycholinguistics*, *33*(3), 643-664.

Carlet, A., & Souza, H. K. D. (2018). Improving L2 Pronunciation Inside and Outside the Classroom: Perception, Production and Autonomous Learning of L2 Vowels. *Ilha do Desterro*, *71*(3), 99-123.

Chan, A. Y., & Li, D. C. (2000). English and Cantonese phonology in contrast: Explaining Cantonese ESL learners' English pronunciation problems. *Language Culture and Curriculum*, *13*(1), 67-85.

Chan, A. Y. (2010). An investigation into Cantonese ESL learners' acquisition of English initial consonant clusters. *Linguistics*, *48*(1), 99-141.

Chitoran, I. (2002). *The phonology of Romanian: A constraint-based approach.* Berlin: Mouton de Grutyer.

Chitoran, I., & Hualde, J. I. (2007). From hiatus to diphthong: the evolution of vowel sequences in Romance. *Phonology*, *24*(1), 37-75.

Chitoran, I., Vasilescu, I., Vieru, B., & Lamel, L. (2014). Analyzing linguistic variation in a Romanian speech corpus through ASR errors. Talk given at Laboratory Approaches to Romance Phonology (LARP) 7, Aix-en Provence, September 3-5, 2014.

Cohen, E. G., & Ben-David, A. (2016). The role of allophony and frequency in the acquisition of the Hebrew rhotic. *Clinical linguistics & phonetics*, *30*(2), 101-118.

Colantoni, L. & Steele, J. (2005). Liquid asymmetries in French and Spanish. *Toronto Working Papers in Linguistics*, *24*, 1-14.

Colantoni, L., & Steele, J. (2006). Native-like attainment in the L2 acquisition of Spanish stop-liquid clusters. In *Selected proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 59-73). Cascadilla.

Colantoni, L., & Steele, J. (2007). Acquiring /ʁ/ in context. *Studies in Second Language Acquisition*, *29*(3), 381-406.

Colantoni, L., & Steele, J. (2008). Integrating articulatory constraints into models of second language phonological acquisition. *Applied Psycholinguistics*, *29*(3), 489-534.

Colantoni, L., Steele, J., Escudero, P., & Neyra, P. R. E. (2015). *Second language speech*. Cambridge University Press.

Crawley, M. J. (2013). *The R book*. Chichester, West Sussex, United Kingdom: Wiley.

Crewson, P. (2006). *Applied statistics handbook*. Retrieved from https://www.acastat.com/Pub/Docs/AppliedStatistics.pdf.

Derrick, D. J., & Gick, B. (2005, May). *Production quality of /r/ and /l/ liquids among Beijing and non-Beijing Mandarin ESL learners*. Paper presented at the Interdisciplinary Speech Research Lab, Vancouver, BC.

Derrick, D., & Gick, B. (2011). Individual variation in English flaps and taps: A case of categorical phonetics. *Canadian Journal of Linguistics/Revue Canadienne de linguistique*, *56*(3), 307-319.

Díaz-Campos, M. (2006). The effect of style in second language phonology: An analysis of segmental acquisition in study abroad and regular-classroom students. In *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 26-39). Somerville, MA: Cascadilla Proceedings Project.

Diehl, R., & Lindblom, B. (2000). Explaining the structure of feature and phoneme inventory: The role of auditory distinctiveness. In A. Greenberg & F. Popper (Eds.), *Speech processing in the auditory system* (pp. 101-162). New York: Springer-Verlag.

Dorwick, T., Pérez-Gironés, A.M., Becher, A. Elliott, A. R., Zapata, G., Rogers, D., & Santos, C. (2014). *Puntos de partida* (2nd Canadian ed.). McGraw-Hill Ryerson Higher Education.

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress 'deafness': The case of French learners of Spanish. *Cognition*, *106*(2), 682-706.

Eckman, F., & Iverson, G. K. (2013). The role of native language phonology in the production of L2 contrasts. *Studies in Second Language Acquisition*, *35*(1), 67-92.

Ellis, N. C., Natsume, M., Stavropoulou, K., Hoxhallari, L., Van Daal, V. H., Polyzoe, N., & Petalas, M. (2004). The effects of orthographic depth on learning to read alphabetic, syllabic, and logographic scripts. *Reading research quarterly*, *39*(4), 438-468.

Erdener, V. D., & Burnham, D. K. (2005). The role of audiovisual speech and orthographic information in nonnative speech production. *Language Learning*, *55*(2), 191-228.

Face, T. L. (2006). Intervocalic rhotic pronunciation by adult learners of Spanish as a second language. In *Selected proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 47-58). Somerville, MA: Cascadilla.

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* (pp. 39-52). Somerville, MA: Cascadilla Proceedings Project.

Féry, C. (2003). Markedness, faithfulness, vowel quality and syllable structure in French. *Journal of French Language Studies*, *13*(2), 247-280.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, *92*, 233-277.

Flege, J. E., Munro, M. J., & MacKay, I. R. (1996). Factors affecting the production of word-initial consonants in a second language. *Second language acquisition and linguistic variation*, *10*, 47-73.

Freed, B. F., Segalowitz, N., & Dewey, D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in second language acquisition*, *26*(2), 275-301.

Frost, R., Repp, B. H., & Katz, L. (1988). Can speech perception be influenced by simultaneous presentation of print?. *Journal of Memory and Language*, *27*(6), 741-755.

Gabriel, C., & Kireva, E. (2014). Prosodic transfer in learner and contact varieties: Speech rhythm and intonation of Buenos Aires Spanish and L2 Castilian Spanish produced by Italian native speakers. *Studies in Second Language Acquisition*, *36*(2), 257-281.

Gass, S., & Varonis, E. (1984). The effect of familiarity on the comprehensibility of non-native speech. *Language Learning*, 34, 65-89.

Gilbert, M.B., & Rohena-Madrazo, M. (2017). Revising the canon: Social and stylistic variation of coda (-ɾ) in Buenos Aires Spanish. In R. Lopes, J. Ornelas de Avelar, & S. Cyrino (Eds.), *Romance languages and linguistic theory 12: Selected papers from the 45th Linguistic Symposium on Romance Languages (LSRL), Campinas, Brazil* (pp. 63–78). Amsterdam, the Netherlands: Benjamins.

Goldstein, B. (2000). *Cultural and linguistic diversity*. San Diego: Thomson Delmar Learning.

Gudurić S., & Petrović D. (2005). O priordi galas [r] u srpskim jeziku [On the nature of the sound [r] in the Serbian language]. *Zbornik Matice za Filologiju i Lingvistiku 48*, 135- 50.

Guitart, J. M. (2004). *Sonido y sentido: teoría y práctica de la pronunciación del español contemporáneo con audio CD* [Sound and meaning: theory and practice of the pronunciation of contemporary Spanish]. Georgetown University Press.

Hammond, R. (1999). On the non-occurrence of the phone [r] in the Spanish sound system. In J. Gutiérrez-Rexach & F. Martínez-Gil (Eds.), *Advances in Hispanic Linguistics* (pp. 135-151). Somerville, MA: Cascadilla Press.

Harris, J. W. (1983). *Syllable structure and stress in Spanish: a nonlinear analysis*. Cambridge, MA: MIT Press.

Harris, J. (2001). Reflections on a phonological grammar of Spanish. *Amsterdam studies in the theory and history of linguistic science series 4,* 133-146.

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. (2005a). *The sounds of Spanish with audio CD*. Cambridge University Press.

Hualde, J. I. (2005b). Quasi-phonemic contrasts in Spanish. In B. Schmeiser, V. Chand, A. Kelleher, & A. Rodriguez (Eds.), *WCCFL 23 Proceedings,* 374-398. Somerville, MA: Cascadilla Press.

Hung, T. T. (2000). Towards a phonology of Hong Kong English. *World Englishes*, *19*(3), 337-356.

Ingvalson, E., McClelland, J., & Holt, L. (2011). Predicting English native-like performance by native Japanese speakers. *Journal of Phonetics*, 39, 571-594.

Jaskuła, K. (2018). The/r/ which dies hard–a diachronic look at the developments of the rhotic sound in selected Celtic, Germanic and Romance languages. *Lublin Studies in Modern Languages and Literature*, *42*(1), 4-17.

Jiménez, B.C. (1987). Acquisition of Spanish consonants in children aged 3-5 years, 7 months. *Language Speech and Hearing Services in Schools,* 18, 357–363.

Johnson, K. E. (2008). Second Language Acquisition of the Spanish Multiple Vibrant Consonant. Doctoral dissertation, University of Arizona, Tucson, AZ, USA.

Jun, S. A., & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus*, *14*(1), 147-172.

Jung, M-W. (1962). *A contrastive study of English and Korean segmental phonemes with some suggestions toward pedagogical application* (Unpublished master's thesis). Georgetown University, Washington, D.C.

Kohler, K. J. (1984). Phonetic explanation in phonology: the feature fortis/lenis. *Phonetica, 41*(3), 150-174.

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia, PA: University of Pennsylvania Press.

Ladefoged, P. (1993). *A course in phonetics*. Fort Worth, TX: Harcourt Brace.

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

Lenth, R. (2018). emmeans: Estimated Marginal Means, aka LeastSquares Means. R package version 1.1. Available at: https://CRAN.R-project.org/package=emmeans.

Léon, P. R. (1992). *Phonétisme et prononciations du français* [Phonetism and pronunciation of French]. Paris: Nathan.

Léon M., & Léon P.R. (2009). *La prononciation du français* [The pronunciation of French]. Paris: Nathan.

Levshina, N. (2015). *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins Publishing Company.

Levy, E. S., & Law, F. F. (2010). Production of French vowels by American-English learners of French: Language experience, consonantal context, and the perception-production relationship. *The Journal of the Acoustical Society of America*, *128*(3), 1290-1305.

Lewis, A. M. (2004). Coarticulatory effects on Spanish trill production. In A. Agwuele, W. Warren, & S. Park (Eds.), *Proceedings of the 2003 Texas Linguistics Society Conference* (pp. 116-127). Somerville, MA: Cascadilla Proceedings Project.

Lindau, M. (1985). The story of /r/. In V. A. Fromkin (Ed.), *Phonetic Linguistics: Essays in honor of Peter Ladefoged* (pp. 157-168). Orlando, FA: Academic Press.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In *Speech production and speech modelling* (pp. 403-439). Springer, Dordrecht.

Lipski, J. (1994). *Latin American Spanish*. New York: Longman.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English/r/and/l: A first report. *The Journal of the Acoustical Society of America*, *89*(2), 874-886.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Mafteiu, D. C. (n.d.). Rhotacism [Rotacismul]. Retrieved August 2018, from http://www.sfatulmedicului.ro/clinici/cabinet-logopedic-andu_177/articole/rotacismul_327.

Major, R. C. (1986). The ontogeny model: Evidence from L2 acquisition of Spanish r. *Language Learning*, *36*(4), 453-504.

Major, R.C. (1987). A model for interlanguage phonology. In G. Ioup & S.H. Weinberger (Eds.), *Interlanguage phonology: the acquisition of a second language sound system,* (pp. 101-124). Cambridge, MA: Newbury House.

Major, R. C. (2001). *Foreign accent: the ontogeny and phylogeny of second language phonology*. Mahwah, NJ: Lawrence Erlbaum.

Marin, S., & Pouplier, M. (2014). Articulatory synergies in the temporal organization of liquid clusters in Romanian. *Journal of Phonetics*, *42*, 24-36.

Martínez Celdrán, E. (1997). El mecanismo de producción de la vibrante apical múltiple [The mechanism of the production of the multimple apical vibrant]. *Estudios de fonética experimental*, *8*, 85-97.

McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of phonetics*, *30*(2), 229-258.

Mendoza, E., Carballo, G., Cruz, A., Fresneda, M. D., Muñoz, J., & Marrero, V. (2003). Temporal variability in speech segments of Spanish: Context and speaker related differences. *Speech Communication*, *40*(4), 431-447.

Montrul, S., & Slabakova, R. (2003). Competence similarities between native and near-native speakers: An investigation of the preterite-imperfect contrast in Spanish. *Studies in second language acquisition*, *25*(3), 351-398.

Montrul, S. (2005). Second language acquisition and first language loss in adult early bilinguals: Exploring some differences and similarities. *Second language research*, *21*(3), 199-249.

Montrul, S. (2010b). Dominant language transfer in adult second language learners and heritage speakers. *Second Language Research*, *26*(3), 293-327.

Morgan, T. A. (2015). *Sonidos en contexto: Una introducción a la fonetica del espanol con especial referencia a la vida real*. New Haven, CT: Yale University Press.

Munro, M. J., Flege, J. E., & MacKay, I. R. (1996). The effects of age of second language learning on the production of English vowels. *Applied psycholinguistics*, *17*(3), 313-334.

Munro, M. J., & Derwing, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech. *Studies in Second Language Acquisition*, 23, 451-468.

Navarro Tomás, T. (1971). Manual de pronunciación española [Manual of Spanish pronunciation]. Madrid: Consejo Superior de Investigaciones Científicas.

Ohala, J. J., & Kawasaki, H. (1984). Prosodic phonology and phonetics. *Phonology*, *1*, 113-127.

Ohala, J. J. (1997). Aerodynamics of phonology. *Proceedings of 4th Seoul international conference in linguistics* (pp. 92-97), Seoul: The Linguistic Society of Korea.

Olsen, M. K. (2012). The L2 acquisition of Spanish rhotics by L1 English speakers: The effect of L1 articulatory routines and phonetic context for allophonic variation. *Hispania* 95, 65-82.

Olsen, M. K. (2016). Limitations of the influence of English phonetics and phonology on L2 Spanish rhotics. *Borealis–An International Journal of Hispanic Linguistics, 5*(2), 313-331.

O'Shaughnessy, D. (1982). A study of French spectral patterns for synthesis. *Journal of Phonetics, 10*, 377-399.

Park, H., & de Jong, K. J. (2008). Perceptual category mapping between English and Korean prevocalic obstruents: Evidence from mapping effects in second language identification skills. *Journal of Phonetics*, *36*(4), 704-723.

Patin, C. (2013). /r/ in Waaashili Shingazidja. In L. Spreafico & A. Vietti (Eds.), *Rhotics: New data and perspectives*, (pp. 173-190). Bozen-Bolzano University Press.

Pérez-Leroux, A. T., Peterson, T., Castilla-Earls, A., Béjar, S., Massam, D., & Roberge, Y. (2018). The acquisition of modification in NPs. *Language 91*(2), 332-359.

Proctor, M. I. (2009). *Gestural characterization of a phonological class: The liquids.* Doctoral Dissertation, New Haven, CT: Yale University.

Puşcariu, S. (1959). *Limba română: Rostirea* (Vol. 2) [The Romanian language: Speech]. Editura academiei republicii populare romine [The edition of the Popular Republic of Romania].

Pyun, K. (1987). *Korean-Swedish Interlanguage Phonology.* Stockholm: Institute of Oriental Languages, University of Stockholm.

Quilis, A. (1981). *Fonética acústica de la lengua española* [Acoustic phonetics of the Spanish language]. Madrid: Gredos.

Quilis, A. (1993). *Tratado de fonología y fonética españolas* [Manual on Spanish phonology and phonetics]. Madrid: Arco.

Radu, M. (2016). *Conditioned variability in the realization of Romanian rhotics*. Poster session presented at the Canadian Linguistic Association (CLA), Calgary, Alberta.

Rafat, Y. (2008). The acquisition of allophonic variation in Spanish as a second language. In *Annual Conference of the Canadian Linguistic Association, Vancouver, BC*.

Rafat, Y. (2010). A socio-phonetic investigation of rhotics in Persian. *Iranian studies, 43*(5), 667-682.

Rafat, Y. (2011). *Orthography-induced transfer in the production of novice adult English-speaking learners of Spanish*. Doctoral dissertation, University of Toronto.

Rafat, Y., & Perry, S. (2019). Navigating orthographic issues in the teaching of Spanish pronunciation. In Rajiv Rao (Ed.), *Key Issues in the Teaching of Spanish Pronunciation: From Description to Pedagogy (Advances in Spanish Language Teaching Series).* New York/London: Routledge.

Rato, A. (2014). Effects of perceptual training on the identification of English vowels by native speakers of European Portuguese. In *Proceedings of the international symposium on the acquisition of second language speech*, 5 (pp. 529-546).

R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Recasens, D. (1999). Lingual coarticulation. In W. Hardcastle, & N. Hewlett (Eds.), *Coarticulation: Theory, data and techniques* (pp. 80–104). Cambridge: Cambridge University Press.

Recasens, D. (2002). Weakening and strengthening in Romance revisited. *Italian Journal of Linguistics*, *14*, 327-374.

Riney, T. J., & Flege, J. E. (1998). Changes over time in global foreign accent and liquid identifiability and accuracy. *Studies in Second Language Acquisition*, *20*(2), 213-243.

Riney, T. J., Takagi, N., & Inutsuka, K. (2015). Phonetic parameters and perceptual judgments of accent in English by American and Japanese listeners. *TESOL Quarterly*, 39, 441-446.

Rogers, B. M., & Alvord, S. M. (2014). The gradience of spirantization: Factors affecting L2 production of intervocalic Spanish [β̞, ð̞, ɣ̞]. *Spanish in Context*, *11*(3), 402-424.

Rose, M. (2012). Cross-Language Identification of Spanish Consonants in English. *Foreign Language Annals*, *45*(3), 415-429.

Sankoff, G., & Blondeau, H. (2007). Language change across the lifespan: /r/in Montreal French. *Language*, *83*, 560-588.

Savu, C. F. (2012). On the phonetic structure of the rhotic tap and its phonological implications. Master's Thesis, University of Bucharest.

Schulte, K. (2005). Vowel centralization in Romanian verbs of Slavic origin: deliberate exploitation of an indigenous sound change? *Amsterdam studies in the theory and history of linguistic science series 4, 270,* 311-325.

Seiciuc, L. (2014). From Romanian to "Rumañol": linguistic confusions in native Romanian spoken by children born in Spain. *Annals of „Ştefan cel Mare" University of Suceava, Philosophy, Social and Human Disciplines series, 1,* 77-87.

Shea, C. E., & Curtin, S. (2011). Experience, representations and the production of second language allophones. *Second Language Research,* 27, 229–250.

Shea, C. (2017). L1 English/L2 Spanish: Orthography–phonology activation without contrasts. *Second Language Research, 33*(2), 207-232.

Sheldon, A., & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied psycholinguistics*, *3*(3), 243-261.

Sheskin, D. J. (2011). *Handbook of Parametric and Nonparametric Statistical Procedures*. Boca Raton, FL: Chapman and Hall/CRC Press.

Simonet, M., Rohena-Madrazo, M., & Paz, M. (2008). Preliminary evidence for incomplete neutralization of coda liquids in Puerto Rican Spanish. In *Selected proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology* (pp. 72-86). Cascadilla Proceedings Project Somerville, Massachusets.

Solé, M. J. (1992). Experimental phonology: the case of rhoticism. In *Phonologica 1988. Proceedings of the 6th international phonology meeting* (W. U. Dressler, H. C.

Luschützky, O. E. Pfeiffer & J. R. Rennison, editors), pp. 259-271. Cambridge: Cambridge University Press.

Solé, M. J. (2002). Aerodynamic characteristics of trills and phonological patterning. *Journal of Phonetics, 30*(4), 655-688.

Steele, J. (2005). Assessing the role of orthographic versus uniquely auditory input in acquiring new L2 segments. *7èmes rencontres internationales du réseau français de phonologie*, 2-4.

Stibbard, R. (2004). The spoken English of Hong Kong: A study of co-occurring segmental errors. *Language, Culture and Curriculum, 17*(2), 127-142.

Tarone, E. (1979). Interlanguage as Chameleon. *Language learning*, *29*(1), 181-191.

Tarone, E. E. (1982). Systematicity and attention in interlanguage. *Language learning*, *32*(1), 69-84.

Tarone, E. (1983). On the variability of interlanguage systems. *Applied linguistics*, *4*(2), 142-164.

Thomas, E. (2011). *Sociophonetics: an introduction*. Basingstoke, England: Palgrave Macmillan.

Tousignant, C. (1987). Les variantes du /R/ montréalais: Contextes phonologiques favorisant leur apparition [The variants of Montreal /R/: Phonological contexts favoring their appearance]. *Revue québécoise de linguistique théorique et appliquée*, *6*(3), 73-113.

Tremblay, A. (2009). Phonetic variability and the variable perception of L2 word stress by French Canadian listeners. *International Journal of Bilingualism, 13*(1), 35-62.

Tremblay, A., & Garrison, M. D. (2010). Cloze tests: A tool for proficiency assessment in research on L2 French. In *Selected proceedings of the 2008 second language research forum* (pp. 73-88). Somerville, MA: Cascadilla Press.

Untu, A. (2011). *Trecere în revista a domeniului acusticii articulatorii cu aplicaţii medicale* (Raport de cercetare I) [An overview of the field of acoustic phonetics with medical applications] (Research report I). Iaşi: Universitatea Tehnică "Gheorghe Asachi" din Iaşi Facultatea de Electronică, Telecomunicaţii si Tehnologia Informaţiei.

Vago, R. M., & Gósy, M. (2007). Schwa vocalization in the realization of /r/. In *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 505-509).

van Lieshout, P., Merrick, G., & Goldstein, L. (2008). An articulatory phonology perspective on rhotic articulation problems: A descriptive case study. *Asia Pacific Journal of Speech, Language and Hearing*, *11*(4), 283-303.

Vokic, G. (2010). L1 allophones in L2 speech production: The case of English learners of Spanish. *Hispania*, *93*, 430-452.

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.

Wiese, R. (2001). The phonology of /r/. *Distinctive feature theory, 2*, 335-368.

Wong, H. (1953). Outline of the Mandarin phonemic system. *Word, 9*(3), 268-276.

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.

Yu, K., & Jamieson, D. G. (1993). Training of the English/r/and/l/speech contrasts in Korean listeners. *Canadian Acoustics, 21*(3), 107-108.

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

Appendix A

Background questionnaire

Subject Number :
Study :

Departments of Spanish & Portuguese

University of Toronto

A. Personal Information

- Sex: □ Male □ Female
- Year of Birth:

  _____
- Place of Birth: City _____ Country

  _____
- Occupation:

  _____
- Highest Level of Schooling: □ Secondary □ CEGEP/College/Professional □ University
- If you were not born in Canada, at what age did you move here?

  _____

B. First Language

What is your first language ?

_____

What is the first language of: your mother? _____ your father?

_____

Did you learn your first language from birth ? □ Yes □ No

- If you answered 'No' to the question above, please explain:

  _____

Which language(s) did you speak at home as a child ?

_____

Is your first language the language with which you are the most comfortable ? □ Yes □ No

- If you answered 'No' to the question above, please explain:

  _____

C. Education & Language Use

Which language(s) were you formally educated in ? Where (i.e. country)?

Primary/Elementary School

_____

High School

_____

CEGEP/College

_____

University

_____

Which language(s) do you use (Indicate approximate percentage, e.g. 0, 50, 100%):

At school

_____

At home

_____

At work

_____

In social situations

_____

D. Second Languages

	Second Languages	
	A.	**B.**
At what age did you begin to learn your 2nd language?		
Where did you learn your 2nd language? Give place and years.		
Were your teachers native speakers of this language?		
Did you learn this language as a subject or was it the principal medium of instruction?	□ Subject □ Medium of Instruction	□ Subject □ Medium of Instruction
Have you ever spent time in an area where this language was the native language?	Where? How long?	Where? How long?
Approximately how many hours a week do you use this language? Specify for each of speaking, listening and reading.	Speaking : _____ hrs Listening : _____ hrs Reading : _____ hrs	Speaking : _____ hrs Listening : _____ hrs Reading : _____ hrs

- *Please rate your linguistic ability in each of your second languages in the following areas by checking the appropriate answer.*

	Beginner	**Intermediate**	**Advanced**	**Near-Native**

READING				
Language A				
Language B				
WRITING				
Language A				
Language B				
SPEAKING				
Language A				
Language B				
LISTENING				
Language A				
Language B				
OVERALL COMPETENCE				
Language A				
Language B				

Do you know any other second languages? Please specify:

_____

D. Hearing and dentition

To your knowledge, do you have normal hearing? Yes No

Do you have any dentition problems (e.g. dentures)? Yes No

If you answered 'Yes' to either of the above, please elaborate:

Appendix B

Reading passage (Spanish): El viento del norte y el sol

El viento del norte y el sol discutían acerca de cuál de los dos sería el más fuerte, cuando, de repente, pasó un viajero envuelto en una amplia capa. Al verlo, convinieron en que el primero que consiguiera quitarle la capa sería el más fuerte. El viento del norte comenzó a soplar con mucha furia, pero, cuanto más soplaba, más se aferraba el viajante a su capa, hasta que el viento norte desistió. El sol brilló entonces con todo su esplendor, e inmediatamente, el viajante arrojó su capa. Así, el viento norte tuvo que reconocer la superioridad del sol.

Appendix C

Cloze test (adapted from DELE (Los Diplomas de Español como Lengua Extranjera [Diploma of Spanish as a Second Language])

Test de texto

Instructions: In the following text, some of the words have been replaced by spaces which are numbered from (1) to (20). First, read the complete text in order to understand it. Then re-read it and choose, from the list of words on the answer sheet, the correct word for each space. Mark your answers by circling your choice on the answer sheet, not on the text.

El sueño de Joan Miró

Hoy se inaugura en Palma de Mallorca la Fundación Pilar y Joan Miró, en el mismo lugar en donde el artista vivió sus últimos treinta y cinco años. El sueño de Joan Miró se ha (1) _____. Los fondos donados a la ciudad por el pintor y su esposa en 1981 permitieron que el sueño se (2) _____; más tarde, en 1986, el Ayuntamiento de Palma de Mallorca decidió (3) _____al arquitecto Rafael Moneo un edificio que (4) _____a la vez como sede de la fundación y como museo moderno. El proyecto ha tenido que (5) _____múltiples obstáculos de carácter administrativo. Miró, coincidiendo (6) _____los deseos de toda su familia, quiso que su obra no quedara expuesta en ampulosos panteones de arte o en (7) _____de coleccionistas acaudalados; por ello, en 1981, creó la fundación mallorquina. Y cuando estaba (8) _____punto de morir, donó terrenos y edificios, así como las obras de arte que en ellos (9) _____.

El edificio que ha construido Rafael Moneo se enmarca en (10) _____se denomina "Territorio Miró", espacio en el que se han (11) _____de situar los distintos edificios que constituyen la herencia del pintor.

El acceso a los mismos quedará (12) _____para evitar el deterioro de las obras. Por otra parte, se (13) _____, en los talleres de grabado y litografía, cursos (14) _____ las distintas técnicas de estampación. Estos talleres también se cederán periódicamente a distintos artistas contemporáneos; (15) _____se busca que el "Territorio Miró" (16) _____un centro vivo de creación y difusión del arte a todos los (17) _____.

La entrada costará 500 pesetas y las previsiones dadas a conocer ayer aspiran (18) _____ _____ que el centro acoja a unos 150.000 visitantes al año. Los responsables esperan que la institución funcione a (19) _____rendimiento a principios de la (20) _____ semana, si bien el catálogo completo de las obras de la Fundación Pilar y Joan Miró no estará listo hasta dentro de dos años.

Hoja de respuestas para el Test de texto

1.	a. cumplido	b. completado	c. terminado
2.	a. inició	b. iniciara	c. iniciaba
3.	a. encargar	b. pedir	c. mandar
4.	a. hubiera servido	b. haya servido	c. sirviera
5.	a. superar	b. enfrentarse	c. acabar
6.	a. por	b. en	c. con
7.	a. voluntad	b. poder	c. favor
8.	a. al	b. en	c. a

9.	a. habría	b. había	c. hubo
10.	a. que	b. el que	c. lo que
11.	a. pretendido	b. tratado	c. intentado
12.	a. disminuido	b. escaso	c. restringido
13.	a. pasarán	b. enseñarán	c. dirán
14.	a. sobre	b. en	c. para
15.	a. ya	b. así	c. para
16.	a. será	b. sea	c. es
17.	a. casos	b. aspectos	c. niveles
18.	a. a	b. de	c. para
19.	a. total	b. pleno	c. entero
20.	a. siguiente	b. próxima	c. pasada

Appendix D

Consent form

UNIVERSITY *of* TORONTO

DEPARTMENT *of* SPANISH & PORTUGUESE

Dear Participant,

I am currently conducting research on the second language learning of Spanish by native Romanian and French speakers. The main goal of this research is to increase our knowledge of how people understand and use second languages.

Participation in this study is voluntary, and you may withdraw at any time without consequence. The experiment involves the following tasks: reading a short text in your first language (Romanian or French) and/or in Spanish; describing some pictures, as well as reading some sentences in your first language (Romanian or French), and/or in Spanish; completing a questionnaire concerning your linguistic background. There are no known risks or particular benefits associated with your participation. The total time involved will be no more than 1 hour. You will receive $15 (47 RON) for your participation**.**

The speaking tasks that you do will be recorded and the audio will be saved for subsequent investigation. Please note that all materials will be treated confidentially. Only my supervisor and I will have access to the forms and data, both of which will be coded with a participant number. The questionnaire and data will be stored separately from any form containing your name and personal information and will contain no links to such forms. Moreover, your identity will not be revealed in any way in the written report of this study.

I thank you for your willingness to participate in this study. To indicate your consent, please sign the attached form. If you have any questions, please do not hesitate to ask them now. Should you require more information following your participation or wish to obtain a summary of the study results, I invite you to contact me by e-mail or telephone at the address and number at the bottom of this page. You may also contact the University of Toronto Ethics Review Office (416 946-3273; ethics.review@utoronto.ca). Should you change your mind following your participation, simply contact us at the e-mail addresses and phone numbers given at the end of this letter.

Yours sincerely,

Malina Radu
PhD Student
malina.radu@mail.utoronto.ca

Professor Jeffrey Steele
Associate Professor
jeffrey.steele@utoronto.ca

CONSENT FORM

I, the undersigned, have been informed of the nature of the present study, including the tasks to be undertaken, and agree to participate. I understand that my participation is voluntary and that I may withdraw at any time without consequence.

Name (*please print*):

Signature:

Date:

Appendix E

Reading passage (French): La bise et le soleil

La bise et le soleil se disputaient, chacun assurant qu'il était le plus fort. Quand ils ont vu un voyageur qui s'avançait, enveloppé dans son manteau, ils sont tombés d'accord que celui qui arriverait le premier à le lui faire ôter serait reconnu comme le plus fort. Alors, la bise s'est mise à souffler de toutes ses forces mais plus elle soufflait, plus le voyageur serrait son manteau autour de lui. Finalement, elle renonça à le lui faire ôter. Alors, le soleil commença à briller et au bout d'un moment le voyageur, réchauffé, ôta son manteau. Ainsi, la bise dut reconnaître que le soleil était le plus fort.

Appendix F

Reading passage (Romanian): Crivăţul și soarele

Crivatul si Soarele se certau sa vada care este mai puternic, cand un calator aparu infasurat intr-un pardesiu calduros. Crivatul si Soarele s-au pus de acord ca cel care reuseste sa il faca pe calator sa isi dea jos pardesiul sa fie considerat cel mai puternic. Atunci Crivatul a suflat cu toata puterea, dar cu cat mai puternic sufla, cu atat mai tare isi strangea calatorul pardesiul pe langa el. Pana la urma Crivatul a renuntat sa mai sufle. Atunci Soarele s-a apucat sa straluceasca, si imediat calatorul si-a dat jos pardesiul din cauza caldurii. Si asa, Crivatul a fost obligat sa recunoasca supremația Soarelui.

Appendix G
Tables of Means

Romanian rhotics

Manner

Table G1

Overall Means and Counts of Romanian Rhotic Realizations Produced by L1 Romanian Speakers across the two tasks

Realization	% occurrence	Total (N=435)
Tap	84.83%	369
Trill	8.05%	35
Fricative	6.67%	29
Approximant	0.46%	2

Table G2

Overall Means and Counts of Romanian Rhotics Produced by L1 Romanian Speakers in Word-initial and Word-medial Positions

	WI (N=221)		WM (N=214)	
Realization	**%**	**Total**	**%**	**Total**
Tap	72.40%	160	97.66%	209
Trill	14.48%	32	1.40%	3
Fricative	12.22%	27	0.93%	2
Approximant	0.90%	2	0	0

Table G3

Overall Means and Counts of Romanian rhotics Produced by L1 Romanian Speakers in the Picture Description (P) and Carrier Sentence Reading (R) Tasks

	P (N=149)		R (N=286)	
Realization	**%**	**Total**	**%**	**Total**

Tap	87.25%	130	83.57%	239
Trill	10.07%	15	6.99%	20
Fricative	2.01%	3	9.09%	26
Approximant	0.67%	1	0.35%	1

Table G4

Overall Means and Counts of Romanian Rhotics Produced by L1 Romanian Speakers in the Nonce and Real Words

	Nonce (N=136)		**Real (N=299)**	
Realization	**%**	**Total**	**%**	**Total**
Tap	80.15%	109	86.96%	260
Trill	7.35%	10	8.36%	25
Fricative	12.50%	17	4.01%	12
Approximant	0%	0	0.67%	2

Length

Table G5

Overall Mean Length (ms) and SDs of Romanian Rhotics produced by L1 Romanian Speakers across both tasks

Realization	**Mean**	***SD***
Tap	30.76	9.77
Trill	74.06	25.67
Fricative	68.45	20.70
Approximant	59.5	9.19

Table G6

Overall mean length (ms) and SDs of Romanian rhotics produced by L1 Romanian speakers in Word-initial and Word-medial positions

	WI		**WM**	
Realization	**Mean**	***SD***	**Mean**	***SD***
Tap	33.86	10.2	28.39	8.73

Trill	75.75	26.21	56	4[45]
Fricative	70	20.60	47.5	4.95
Approximant	59.5	9.12	n/a	n/a

Table G7

Overall Mean Length (ms) and SDs of Romanian Rhotics Produced by L1 Romanian Speakers in the Picture Description (P) and Carrier Sentence Reading (R) tasks

	P		**R**	
Realization	**Mean**	***SD***	**Mean**	***SD***
Tap	30.88	8.89	30.69	10.23
Trill	65.4	25.49	80.55	24.44
Fricative	65.67	2.52	68.77	21.87
Approximant	66	n/a[46]	53	n/a[47]

Table G8

Overall Mean Length (ms) and SDs of Romanian Rhotics Produced by L1 Romanian Speakers in the Nonce and Real Words

	Nonce		**Real**	
Realization	**Mean**	***SD***	**Mean**	***SD***
Tap	30.77	10.42	30.75	9.50
Trill	88.1	29.34	68.44	22.27
Fricative	70.11	22.93	66.08	17.74
Approximant	n/a	n/a	59.5	9.19

[45] Low standard deviation because there were only 3 word-medial trills.
[46] Only one token.
[47] Only one token.

Percentage of voicing

Table G9

Overall Mean Percentage of Voicing and SDs of Romanian Rhotics Produced by L1 Romanian Speakers across both tasks

Realization	**Mean**	***SD***
Tap	87.62	30.55
Trill	89.00	28.39
Fricative	91.82	23.70
Approximant	100	0

Table G10

Overall Mean Percentage of Voicing and SDs of Romanian Rhotic Realizations Produced by L1 Romanian Speakers in Word-initial and Word-medial Positions

	WI		**WM**	
Realization	**Mean**	***SD***	**Mean**	***SD***
Tap	93.15	23.56	83.40	34.43
Trill	87.97	29.16	100	0
Fricative	91.21	24.48	100	0
Approximant	100	0	n/a	n/a

Table G11

Overall Mean Percentage of Voicing and SDs of Romanian Rhotic Realizations Produced by L1 Romanian Speakers in the Picture Description (P) and Carrier Sentence Reading (R) Tasks

	P		**R**	
Realization	**Mean**	***SD***	**Mean**	***SD***
Tap	89.82	34.89	90.24	27.64
Trill	78.39	38.68	96.96	13.58
Fricative	92.42	13.12	91.75	24.80
Approximant	100	n/a	100	n/a

Table G12

Overall Mean Percentage of Voicing and SDs of Romanian Rhotic Realizations Produced by L1 Romanian Speakers in the Nonce and Real words

	Nonce		**Real**	
Realization	**Mean**	***SD***	**Mean**	***SD***
Tap	94.61	19.57	84.69	30.55
Trill	100	0	84.51	32.73
Fricative	87.38	30.04	98.11	6.56
Approximant	n/a	n/a	100	0

Spanish rhotics

Manner

Table G13

Overall Manner of Rhotics Produced by the L1 French-L2 Spanish Speakers in Tap Contexts (left) and Trill Contexts (right) across both tasks

/ɾ/ (N=147)			**/r/ (N=265)**	
	% occurrence	**Total**	**% occurrence**	**Total**
Tap	70.75%	104	57.00%	151
Trill	14.29%	21	26.41%	70
Fricative	5.44%	8	4.91%	16
Approximant	4.08%	6	0.38%	2
Uvular	3.4%	5	7.92%	21
Other	2.04%	3	1.89%	5

Note. Expected realizations are underlined.

Table G14

Overall Manner of Rhotics Produced by the L1 Romanian-L2 Spanish Speakers in Tap Contexts (left) and Trill contexts (right) across both tasks

	/ɾ/ (N=157)		**/r/ (N=262)**	
	Mean	**Total**	**Mean**	**Total**
Tap	89.81%	141	57.63%	151
Trill	7.01%	11	37.79%	99
Fricative	3.18%	5	3.44%	9
Approximant	n/a	n/a	1.15%	3

Note. Expected realizations are underlined.

Table G15

Overall manner of rhotics produced by the Spanish controls in tap contexts (on the left) and trill contexts (on the right) across both tasks

	/ɾ/ (N=205)		**/r/ (N=385)**	
	% occurrence	**Total**	**% occurrence**	**Total**
Tap	98.05%	201	2.60%	10
Trill	1.46%	3	89.87%	346
Fricative	0%	0	7.27%	28
Approximant	0.49%	1	0.26%	1

Note. Expected realizations are underlined.

Table G16

Overall Manner of Unexpected Rhotic Realizations Produced by the L1 French-L2 Spanish Speakers in Word-initial and Word-medial Positions across both tasks

	WM tap (N=43)		**WI trill (N=127)**		**WM trill (N=68)**	
	%	**N**	**%**	**N**	**%**	**N**
Tap	[redacted]		77.95%	99	76.47%	52
Trill	48.83%	21	[redacted]			
Fricative	18.60%	8	8.66%	11	7.35%	5
Approximant	13.95%	6	0.79%	1	1.48%	1
Uvular	11.64%	5	12.60%	16	7.35%	5
	6.98%	3	0%	0	7.35%	5

Other						

Table G17

Overall Manner of Unexpected Realizations Produced by the L1 Romanian-L2 Spanish Speakers in Word-initial and Word-medial Positions across both tasks

	WM tap (N=16)		**WI trill (N=97)**		**WM trill (N=66)**	
	%	**N**	**%**	**N**	**%**	**N**
Tap	[redacted]		88.66%	86	98.48%	65
Trill	68.75%	11	[redacted]			
Fricative	31.25%	5	8.25%	8	1.52%	1
Approximant	0%	0	3.09%	3	0%	0

Table G18

Overall Manner of Unexpected Realizations Produced by the Native Spanish Speakers Word-initial and Word-medial Position across both tasks

	WM tap (N=4)		**WI trill (N=28)**		**WM trill (N=11)**	
	%	**N**	**%**	**N**	**%**	**N**
Tap	[redacted]		25%	7	27.27%	3
Trill	75.00%	3	[redacted]			
Fricative	0%	0	71.43%	20	72.73%	8
Approximant	25.00%	1	3.57%	1	0%	0

Table G19

Overall Manner of Unexpected Realizations Produced by the L1 French-L2 Spanish Speakers in the Picture Description and Carrier Sentence Reading Tasks

		WM tap (N=43)		**WI trill (N=127)**		**WM trill (N=68)**	
		%	**N**	**%**	**N**	**%**	**N**
Tap							
	Picture			22.83%	29	13.24%	9
	Reading			55.12%	70	63.24%	43
Trill							
	Picture	11.63%	5				
	Reading	37.21%	16				
Fricative							
	Picture	6.98%	3	2.36%	3	1.47%	1
	Reading	11.63%	5	6.30%	8	5.88%	4
Uvular							
	Picture	4.65%	2	5.51%	7	4.41%	3
	Reading	6.98%	3	7.09%	9	2.94%	2
Approximant							
	Picture	0	0	0.79%	1	0	0
	Reading	13.95%	6	0	0	1.47%	1
Other							
	Picture	2.32%	1	0	0	0	0
	Reading	4.65%	2	0	0	7.35%	5

Table G20

Overall Manner of Unexpected Realizations Produced by the L1 Romanian-L2 Spanish Speakers in the Picture Description and Carrier Sentence Reading Tasks

		WM tap (N=16)		**WI trill (N=97)**		**WM trill (N=66)**	
		%	**N**	**%**	**N**	**%**	**N**
Tap							
	Picture			1.03%	1	31.82%	21
	Reading			7.22%	7	66.67%	44
Trill							
	Picture	12.5%	2				
	Reading	18.75%	3				
Fricative							
	Picture	31.25%	5	29.90%	29	0	0
	Reading	37.50%	6	58.76%	57	1.51%	1

Approximant							
	Picture	0	0	0	0	0	0
	Reading	0	0	3.09%	3	0	0

Table G21

Overall Manner of Unexpected Realizations Produced by the Native Spanish Speakers in the Picture Description and Carrier Sentence Reading Tasks

		WM tap (N=4)		**WI trill (N=28)**		**WM trill (N=11)**	
		%	**N**	**%**	**N**	**%**	**N**
Tap							
	Picture			10.71%	3	9.09%	1
	Reading			14.29%	4	18.18%	2
Trill							
	Picture	0	0				
	Reading	75%	3				
Fricative							
	Picture	0	0	50%	14	54.55%	6
	Reading	0	0	21.43%	6	18.18%	2
Approximant							
	Picture	0	0	3.57%	1	0	0
	Reading	25%	1	0	0	0	0

Table G22

Overall Manner of Unexpected Realizations Produced by the L1 French-L2 Spanish Speakers in Nonce and Real words

		WM tap (N=43)		**WI trill (N=127)**		**WM trill (N=68)**	
		%	**N**	**%**	**N**	**%**	**N**
Tap							
	Nonce			28.35%	36	29.41%	20
	Real			49.61%	63	47.06%	32
Trill							
	Nonce	16.28%	7				
	Real	32.55%	14				
Fricative							
	Nonce	11.63%	5	3.95%	5	2.94%	2
	Real	6.98%	3	4.73%	6	4.41%	3
Uvular							
	Nonce	4.65%	2	3.94%	5	0	0
	Real	6.98%	3	8.66%	11	7.35%	5
Approximant							
	Nonce	6.98%	3	0	0	1.48%	1
	Real	6.98%	3	0.79%	1	0	0
Other							
	Nonce	4.65%	2	0	0	7.35%	5
	Real	2.32%	1	0	0	0	0

Table G23

Overall Manner of Unexpected Realizations Produced by the L1 Romanian-L2 Spanish Speakers in Nonce and Real Words

		WM tap (N=16)		**WI trill (N=97)**		**WM trill (N=66)**	
		%	**N**	**%**	**N**	**%**	**N**
Tap							
	Nonce			29.90%	29	33.33%	22
	Real			58.76%	57	65.15%	43
Trill							
	Nonce	18.75%	3				
	Real	50%	8				
Fricative							
	Nonce	12.5%	2	3.09%	3	0	0
	Real	18.75%	3	5.15%	5	1.52%	1
Approximant							
	Nonce	0	0	1.04%	1	0	0
	Real	0	0	2.06%	2	0	0

Table G24

Overall Manner of Unexpected Realizations Produced by the Native Spanish Speakers in Nonce and Real Words

		WM tap (N=4)		**WI trill (N=28)**		**WM trill (N=11)**	
		%	**N**	**%**	**N**	**%**	**N**
Tap							
	Nonce			0	0	18.18%	2
	Real			25%	7	9.09%	1
Trill							
	Nonce	50%	2				
	Real	25%	1				
Fricative							
	Nonce	0	0	10.71%	3	9.09%	1
	Real	0	0	60.71%	17	63.64%	7
Approximant							
	Nonce	25%	1	0	0	0	0
	Real	0	0	3.58%	1	0	0

Length

Table G25

Overall Mean Length (ms) and SDs of Expected Rhotics for the L1 French-L2 Spanish (FR), L1 Romanian-L2 Spanish (RO) and Native Spanish (SPA) across both tasks in the three linguistic contexts

	FR		**RO**		**SPA**	
Realization	**Mean**	***SD***	**Mean**	***SD***	**Mean**	***SD***
Medial tap	32.51	14.47	32.84	10.06	30.04	11.01
Initial trill	74.47	12.71	79.12	27.08	83.04	22.2
Medial trill	96.94	30.93	94.21	31.17	88.28	25.84

Table G26

Overall Mean Length (ms) and SDs of Expected Spanish Rhotics in the Picture Description and Carrier Sentence Reading Tasks as Produced by L1 French-L2 Spanish (FR), L1 Romanian-L2 Spanish (RO) and Native Spanish (SPA) Groups

		FR		**RO**		**SPA**	
Realization		**Mean**	***SD***	**Mean**	***SD***	**Mean**	***SD***
Tap							
	Picture	36.47	16.19	32.44	9.31	33.11	12.45
	Reading	30.41	13.13	33.07	10.49	28.26	9.70
Trill							
	Picture	90.33	30.66	75.73	47.51	85.64	22.23
	Reading	90.95	28.83	89.49	27.49	85.48	24.90

Table G27

Overall Length (ms) and SDs of Expected Spanish Rhotic Realizations in the Nonce and Real Words as Produced by the L1 French-L2 Spanish (FR), L1 Romanian-L2 Spanish (RO) and Native Spanish (SPA) Groups

		FR		**RO**		**SPA**	
Realization		**Mean**	***SD***	**Mean**	***SD***	**Mean**	***SD***
Tap							
	Nonce	30.84	12.16	32.64	9.0	28.71	9.54
	Real	33.22	15.37	32.94	10.56	30.64	11.59
Trill							
	Nonce	88.27	28.12	91.18	23.67	87.80	26.66
	Real	92.36	29.61	85.28	34.93	84.27	22.55

Percentage of Voicing

Table G28

Overall Mean Percentage of Voicing and SDs of Expected Rhotic Realizations for the L1 French-L2 Spanish (FR), L1 Romanian-L2 Spanish (RO) and Native Spanish (SPA) Groups across both tasks

	FR		**RO**		**SPA**	
Realization	**Mean**	***SD***	**Mean**	***SD***	**Mean**	***SD***
Medial tap	95.46	16.73	84.72	32.22	92.57	23.92
Initial trill	91.73	19.68	87.22	25.95	75.28	34.84
Medial trill	88.56	24.08	78.35	34.04	85.27	29.02

Table G29

Overall Percentage of Voicing and SDs of Expected Spanish Rhotic Realizations in the Picture Description and Carrier Sentence Reading Tasks as Produced by the L1 French-L2 Spanish (FR), L1 Romanian-L2 Spanish (RO) and Native Spanish (SPA) Groups

		FR		**RO**		**SPA**	
Realization		**Mean**	***SD***	**Mean**	***SD***	**Mean**	***SD***
Tap							
	Picture	88.27	25.92	90.29	24.78	80.79	36.18
	Reading	99.26	6.06	80.97	35.36	99.42	4.60
Trill							
	Picture	94.97	13.73	82.46	31.18	76.23	34.49
	Reading	88.27	24.28	81.97	31.28	81.60	31.65

Table G30

Overall Percentage of Voicing of Expected Spanish Rhotic Realizations in the Nonce and Real Words as Produced by the L1 French-L2 Spanish (FR), L1 Romanian-L2 Spanish (RO) and Native Spanish (SPA) Groups

		FR		**RO**		**SPA**	
Realization		**Mean**	***SD***	**Mean**	***SD***	**Mean**	***SD***
Tap							
	Nonce	100	0	78.42	37.11	100	0
	Real	95.53	19.69	87.01	29.48	89.25	28.17
Trill							
	Nonce	87.10	24.76	83.12	29.31	78.95	33.32
	Real	90.80	21.87	81.11	32.79	80.60	32.19

Appendix H
Boxplots

Romanian rhotics

Length

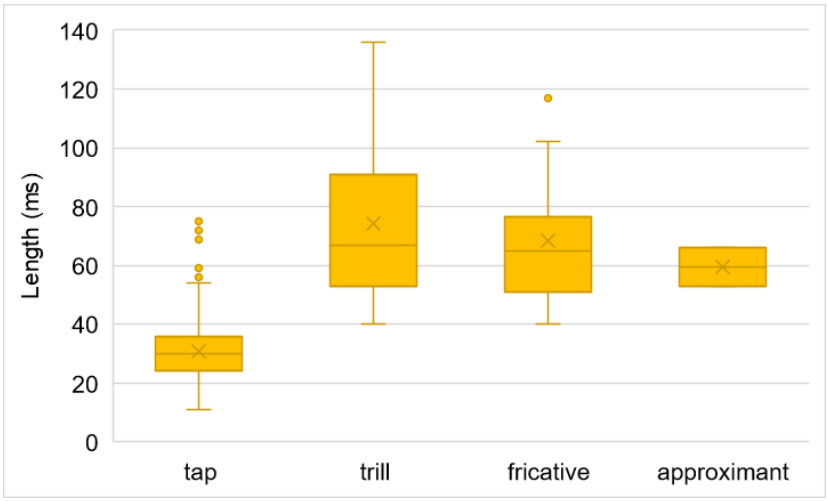

Figure H1. Boxplot representing overall length of Romanian rhotics produced by L1 Romanian speakers across both tasks.

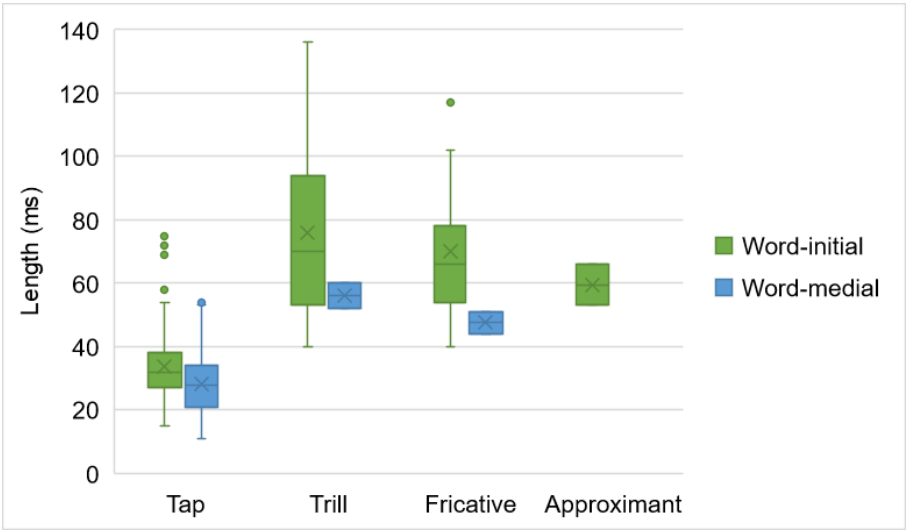

Figure H2. Boxplot representing length of Romanian rhotics produced by L1 Romanian speakers across both in word-initial and word-medial positions.

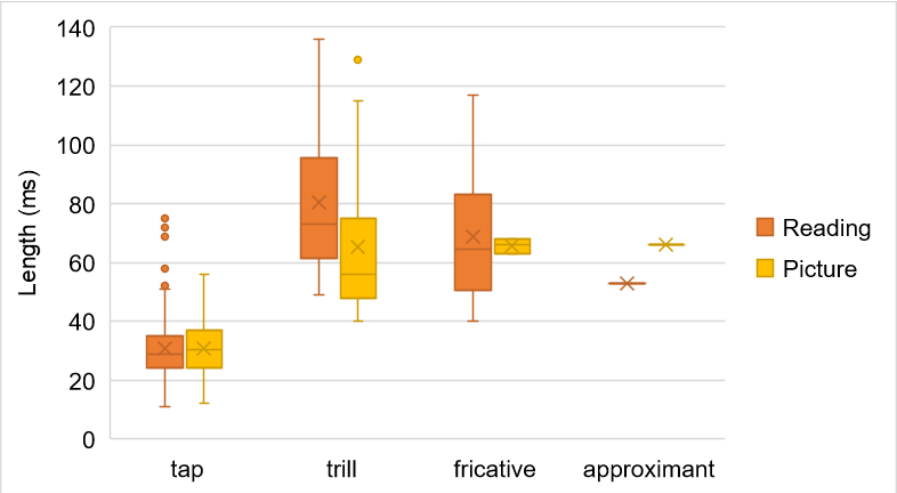

Figure H3. Boxplot representing length of Romanian rhotics produced by L1 Romanian speakers across both in the picture description and carrier reading tasks.

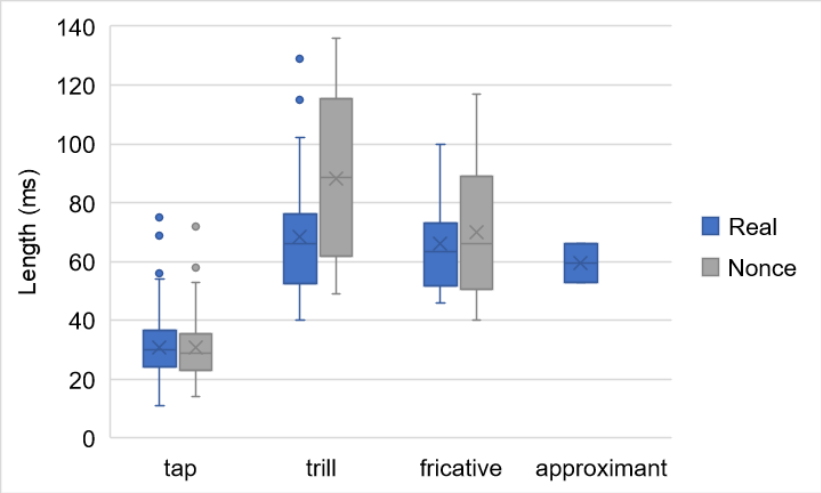

Figure H4. Boxplot representing length of Romanian rhotics produced by L1 Romanian speakers across both in the nonce and real words.

Percentage of voicing

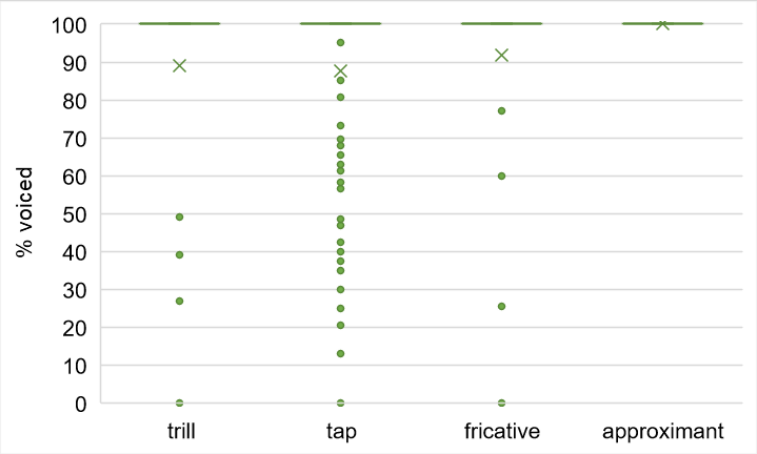

Figure H5. Boxplot representing the overall percentage of voicing of Romanian rhotics produced by L1 Romanian speakers across both tasks.

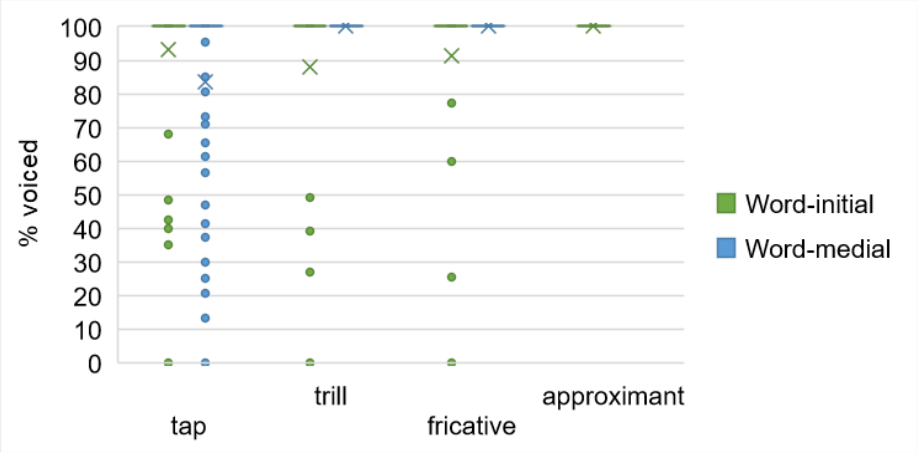

Figure H6. Boxplot representing the percentage of voicing of Romanian rhotics produced by L1 Romanian speakers in word-initial and word-medial positions.

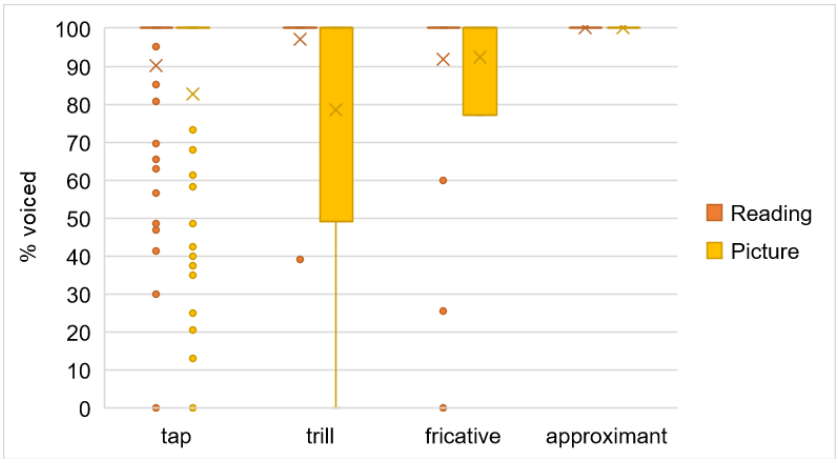

Figure H7. Boxplot representing the percentage of voicing of Romanian rhotics produced by L1 Romanian speakers in the picture description and carrier sentence reading tasks.

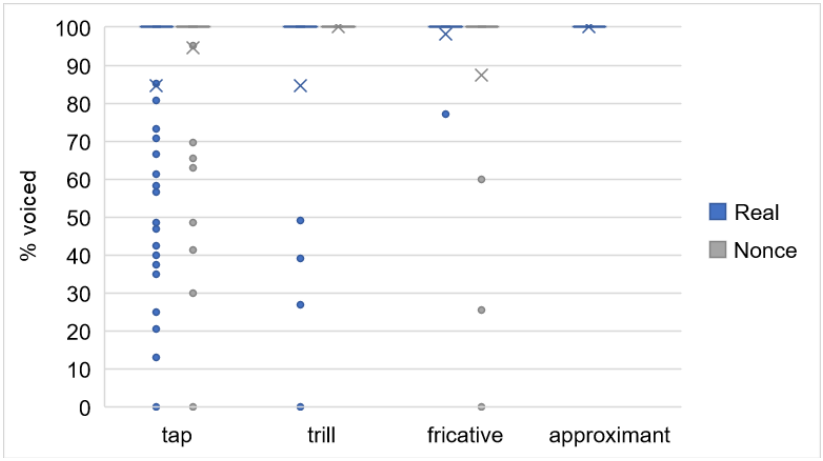

Figure H8. Boxplot representing the percentage of voicing of Romanian rhotics produced by L1 Romanian speakers in the nonce and real words.

Spanish rhotics

Length

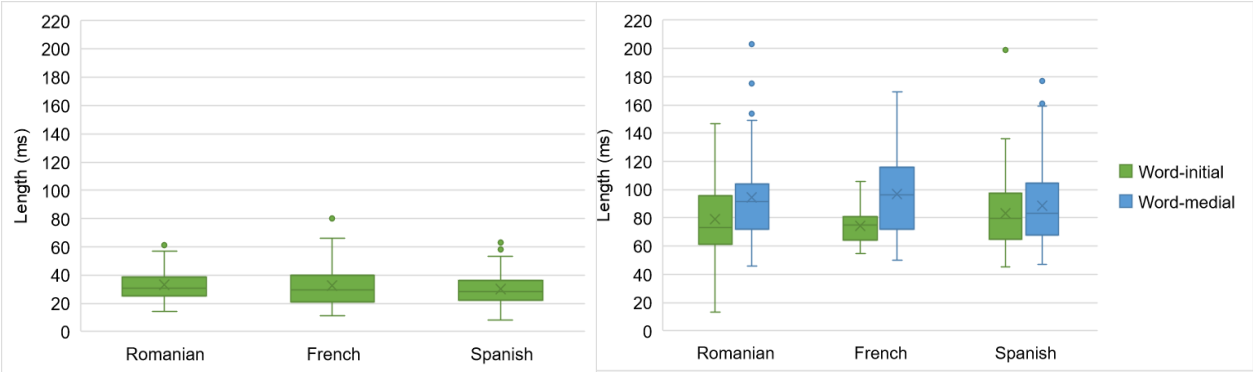

Figure H9. Boxplot representing the overall length of rhotics produced by the L1 French-L2 Spanish, L1 Romanian-L2 Spanish and native Spanish groups across both tasks in the three linguistic contexts (word-medial tap on the left; word-medial and initial trills on the right).

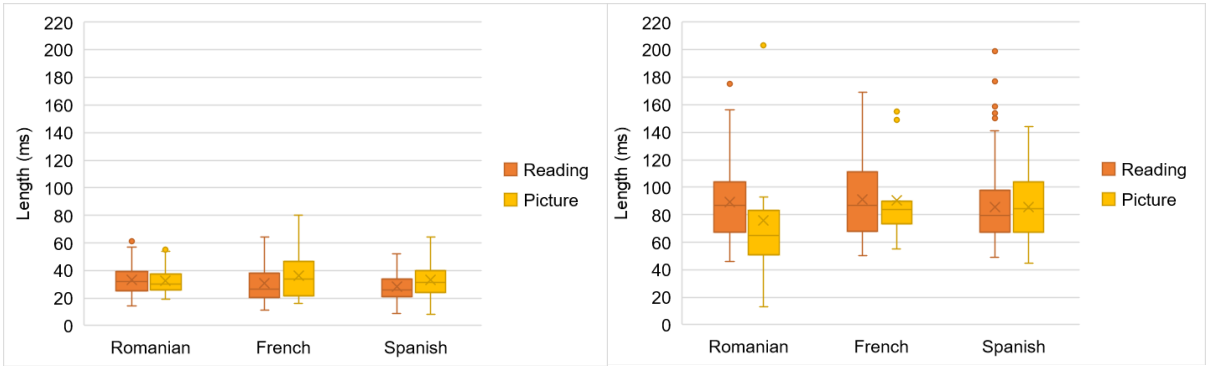

Figure H10. Boxplot representing the length of rhotics produced by the L1 French-L2 Spanish, L1 Romanian-L2 Spanish and native Spanish groups in the picture description and carrier sentence reading tasks in the three linguistic contexts (word-medial tap on the left; word-medial and initial trills on the right).

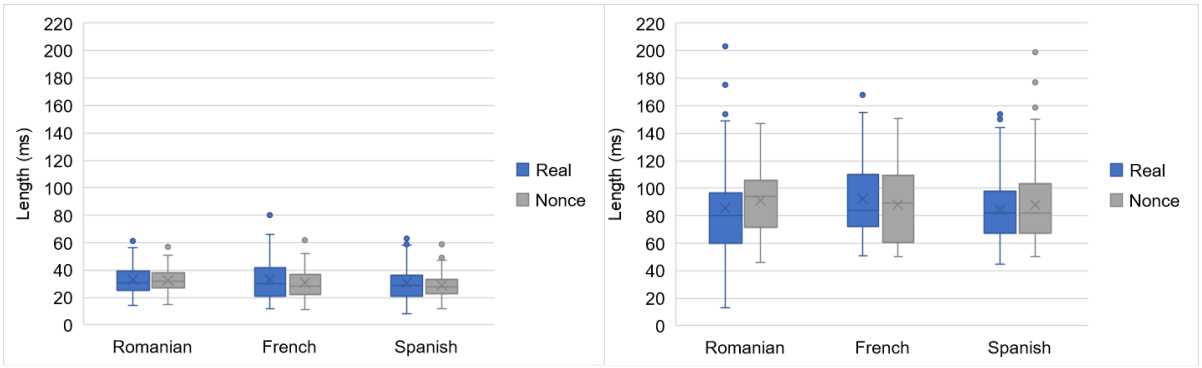

Figure H11. Boxplot representing the overall length of rhotics produced by the L1 French-L2 Spanish, L1 Romanian-L2 Spanish and native Spanish groups in the nonce and real words in the three linguistic contexts (word-medial tap on the left; word-medial and initial trills on the right).

Percentage of voicing

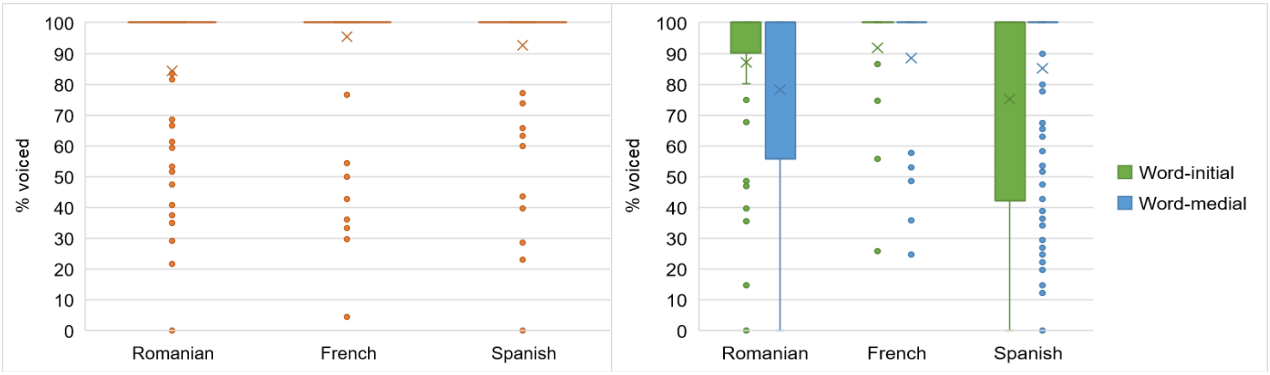

Figure H12. Boxplot representing the overall percentage of voicing of rhotics produced by the L1 French-L2 Spanish, L1 Romanian-L2 Spanish and native Spanish groups across both tasks in the three linguistic contexts (word-medial tap on the left; word-initial and medial trill on the right).

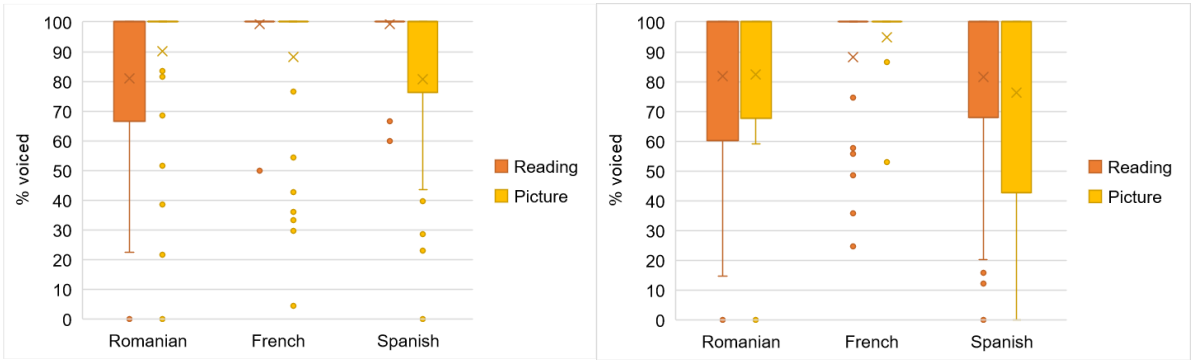

Figure H13. Boxplot representing the percentage of voicing of rhotics produced by the L1 French-L2 Spanish, L1 Romanian-L2 Spanish and native Spanish groups in the picture description and carrier sentence reading tasks in the three linguistic contexts (tap on the left; trill on the right).

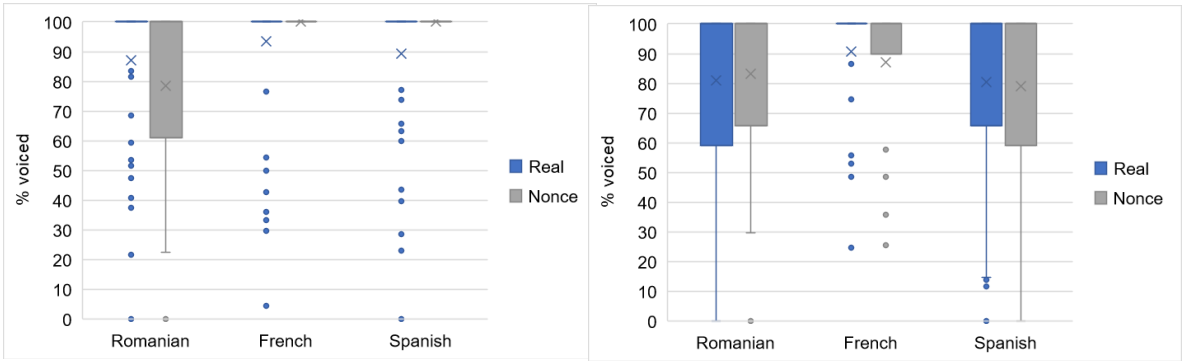

Figure H14. Boxplot representing the percentage of voicing of rhotics produced by the L1 French-L2 Spanish, L1 Romanian-L2 Spanish and native Spanish in the nonce and real words in the three linguistic contexts (tap on the left; trill on the right).