Journal of Second and Multiple Language Acquisition – JSMULA
Vol: 8 Issue: 4 79-110, 2020, December
ISSN:4147-9747

Production and perception of Spanish voiced stops and approximants by L2 learners

Received : 24.07.2020
Accepted : 18.10.2020
Published : 30.12.2020

Christina A. Mirisis[1]
St. Norbert College

Abstract

This study investigates second language (L2) Spanish learners' acquisition of word-initial voiced stops [b d g] and intervocalic voiced approximants [β ð ɣ] at different points in a university curriculum to explore the potential relationship between their production and perception. Twenty-three L1 English-speaking L2 learners of Spanish enrolled in first-, third-, and fourth-year courses, respectively, at a university in the United States and four L1 Spanish speakers participated in the study. Participants' production was assessed with a word list reading comprised of nonce words. For the perception task, participants listened twice to a recording of an L1 Spanish speaker reading the same words used in the production task. Each word contained a voiced stop in word-initial position and its corresponding approximant allophone in intervocalic position. Prior to completing the perception task, participants were informed that each word contained two tokens of the same consonant (e.g., two tokens of [b]) and they were to indicate on an answer sheet whether the two instances of the repeated consonant in each word were pronounced the same or differently. Acoustic analysis in Praat revealed that overall, learners produced word-initial [b d g] with significantly less prevoicing than L1 Spanish speakers and they infrequently produced target approximants as such. Acoustic analysis also indicated that first-, third-, and fourth-year learners achieved less spirantization of target intervocalic voiced approximants, when produced as such rather than as voiced stops, than L1 Spanish speakers. In addition, first-, third-, and fourth-year learners perceived a difference between target voiced stops and approximants 49%, 43%, and 58% of the time, respectively, following a U-shaped learning curve. Finally, Pearson's correlation analyses suggest that the relationship between learners' production and perception of word-initial voiced stops and intervocalic voiced approximants depends on their instruction level and the target sound/sound pair.

Keywords: L2 Spanish, voice onset time, spirantization, production, perception

1. Introduction

Many studies have documented L1 English-speaking second language (L2) Spanish learners' difficulty producing voiced stops and approximants in a target-like fashion, as they seldom produce Spanish voiced stops with prevoicing (Elliott, 1997; González-Bueno, 1994; Zampini, 1998) and generally produce target approximants as voiced stops (Díaz-Campos, 2004; Elliott, 1997; González-Bueno, 1995; Kissling, 2013; Shively, 2008; Zampini,

[1] Bio: Christina A. Mirisis is an Assistant Professor of Modern Languages and Literatures at St. Norbert College in De Pere, WI where she teaches Spanish language and linguistics courses. Her research interests include second language acquisition, phonetics and phonology, L2 Spanish phonology, L2 Italian phonology, and Spanish sociolinguistics. Corresponding author: christina.mirisis@snc.edu

1994). Previous studies have attributed learners' production difficulty to different factors, including first language (L1) transfer (Elliott, 1997; Face & Menke, 2009; González-Bueno, 1994, 1995; Shively, 2008; Zampini, 1994), negative influence of orthography (Elliott, 1997; Face & Menke, 2009; González-Bueno, 1995; Shively, 2008; Zampini, 1994), and markedness (Díaz-Campos, 2004; Elliott, 1997; González-Bueno, 1994). Several studies have found that in many cases, L1 transfer from English in L2 Spanish learners' production of target voiced stops and approximants is due to the negative influence of orthography. For example, a common L1 transfer error, which has been documented in second-semester (Shively, 2008), third-semester (Elliott, 1997), fourth-semester (Face & Menke, 2009), and second- and fourth-semester (Zampini, 1994) L2 Spanish learners' productions, respectively, is pronouncing [b] and [β] as a voiced labiodental fricative [v], as in English, in Spanish words containing orthographic "v". In addition, González-Bueno (1994) hypothesized that fourth-semester students who received targeted pronunciation instruction did not produce /d/ with prevoicing, on a conversational task, due to L1 transfer if they produced the target sound at the alveolar place of articulation, as in English. Moreover, González-Bueno (1995) attributed fourth-semester students' difficulty producing [ð], which they produced in a target-like fashion 26.7% of the time on a conversational task, to L1 transfer, as learners produced target [ð] as the English alveolar flap 32.2% of the time, and to a negative effect of orthography since [ð] is represented by "d" in Spanish but by "th" in English. More generally, Zampini (1994) and Elliott (1997) also found transfer errors in L2 learners' production of voiced stops in phonetic contexts that required their approximant allophones. Finally, researchers have also attributed L2 Spanish learners' difficulty producing voiced stops and approximants in a target-like fashion to markedness. Specifically, González-Bueno (1994) posited that markedness may have played a role in learners' difficulty since the contrast between voiceless unaspirated stops and prevoiced voiced stops is marked for L1 English-speaking L2 Spanish learners. In addition, citing Jakobson (1968) who claimed that "fricatives are more marked than their stop counterparts and consequently are more difficult to acquire" (p. 102), Elliott (1997) proposed that markedness may explain learners' overall difficulty with the voiced approximants. Finally, Díaz-Campos (2004) claimed that both study abroad (SA) and at-home (AH) L2 Spanish learners with an intermediate-low level of proficiency struggled to produce voiced approximants in a target-like fashion because voiced approximants are more marked than other sounds, such as stops, and are therefore more difficult to acquire (Eckman, 1987).

Most previous studies reviewed above describe participants referring to their semester or year of L2 Spanish instruction at the time of the study; however, other studies describe participants in terms of proficiency level (Díaz-Campos, 2004; González-Bueno & Quintana-Lara, 2011), making comparison of results across studies difficult. Further, there is variation among the latter studies in how L2 Spanish learners' proficiency level was determined. While Díaz-Campos (2004) used ACTFL's Oral Proficiency Interview to assess L2 learners' proficiency at the beginning and conclusion of his study, González-Bueno and Quintana-Lara (2011) used the course in which learners had been placed by the Spanish Department to determine their level of proficiency, a less objective method, as mentioned by the researchers themselves. Additionally, some of the earliest L2 acquisition studies on Spanish voiced stops and approximants (Elliott, 1997; Zampini 1994) relied on impressionistic data, while later studies on L2 acquisition of these sounds employed acoustic analysis (Face & Menke, 2009; González-Bueno, 1994, 1995; Kissling, 2013; Shively, 2008; Zampini, 1998) or a combination of impressionistic and acoustic analysis (González-Bueno & Quintana-Lara, 2011). However, it is important to note that previous acoustic studies on L2 acquisition of Spanish voiced approximants examined learners' production via visual inspection of waveforms and/or spectrograms rather than through more precise measurements. More recent studies (Bongiovanni et al., 2015; Nagle, 2017; Rogers & Alvord, 2014) provide a more fine-grained analysis, as they use relative intensity measurements to investigate L2 learners' production of Spanish voiced approximants. Specifically, Rogers and Alvord (2014) found that learners who spent two years abroad in a Spanish-speaking country as religious missionaries had a mean intensity difference of 6.24 dB and AH learners who had studied four semesters of university Spanish had a mean intensity difference of 13.53 dB, compared to L1 Spanish speakers whose mean intensity difference was 3.86 dB. Given these mean intensity differences, the SA learners in Rogers and Alvord (2014) more closely approximated a similar level of spirantization as L1 Spanish speakers. In a study on the effect of a short-term SA program on L2 Spanish phonetic development, Bongiovanni et al. (2015) report that the SA group's production of [β] was significantly less target-like (i.e., greater intensity difference) than the AH group's production of [β] at both times tested and the SA group's production of [ð] was also significantly less target-like than the AH group's but only at Time 1, while there were no differences between the groups in their production of [ɣ]. More recently, Nagle (2017) studied individual L1 English-speaking L2 Spanish learners' production of [β] across phonetic contexts in a longitudinal study, finding that collectively, learners did not produce significantly greater levels of spirantization over the course of the study, but individual development varied considerably. Focusing on individual patterns, Nagle reports that nine learners improved their production, five appeared to regress, and six did not exhibit either positive or negative development.

In addition to investigating the effect of context of learning on L2 Spanish learners' acquisition of voiced stops and approximants and learners' acquisition over time, some previous studies have also considered the role of explicit pronunciation instruction on these target sounds in the acquisition process. González-Bueno (1994) examined the effect of explicit pronunciation instruction on L2 Spanish learners' acquisition of voiceless and voiced stops, finding that learners who received such instruction improved their pronunciation of only /p/ and /g/ significantly more than learners who did not receive such instruction. In Elliott's (1997) study, learners in the experimental group (n = 43) who received instruction on the target sounds struggled to produce intervocalic voiced approximants in a target-like fashion, similar to the learners who did not receive such instruction (n = 23).

Zampini (1998) investigated production and perception of Spanish /p/ and /b/ in L2 learners enrolled in a Spanish phonetics course; however, even these learners did not produce Spanish /b/ in a target-like fashion. In a later study, some learners (*n* = 17) in Shively (2008) were similarly enrolled in a Spanish phonetics course. While Zampini (1998) did not find gains in learners' pronunciation of Spanish /b/, Shively (2008) found that learners enrolled in a Spanish phonetics course made greater gains in their pronunciation of intervocalic [β ð ɣ] than learners (*n* = 19) enrolled in a second-semester Spanish language course. Kissling (2013) examined the effect of phonetics instruction on L2 Spanish learners' perception and production of voiced approximants, among other sounds, finding that learners who received such instruction did not improve their production of target sounds significantly more than learners who did not receive such instruction. However, Kissling (2013) found that explicit instruction improved learners' ability to discriminate target sounds when analyzed collectively rather than individually. Kissling (2013) also found that explicit instruction led to an improvement in learners' accuracy identifying target L2 sounds, but only immediately following instruction. Finally, SA and AH learners in Bongiovanni et al. (2015) were enrolled in an Introduction to Hispanic Linguistics course, in which Spanish phonetics and phonology were covered, at the time of the study. However, since Spanish phonetics and phonology were covered differently in the SA and AH Introduction to Hispanic Linguistics courses, it is possible that differences in instruction influenced the results reported previously. In summary, the mixed results on the effect of explicit pronunciation instruction point to the difficulty with which Spanish voiced stops and approximants are acquired.

While L2 Spanish learners' difficulty producing voiced stops and approximants has been attributed to different extralinguistic factors, the role of perception has been understudied. To the best of my knowledge, only González-Bueno and Quintana-Lara (2011) and Kissling (2013) have investigated L2 learners' production and perception of Spanish voiced approximants, while only Zampini (1998) has investigated L2 learners' production and perception of Spanish voiced stops. The results of these previous studies are discussed in Section 4.3 in light of the findings of the present study; however, it is important to note that González-Bueno and Quintana-Lara's (2011) results are not based on a correlation analysis, so it is unknown if there is indeed a relationship between learners' production and perception. In addition, since Zampini (1998) only investigated L2 Spanish learners' production and perception of [b], it is unknown if there is a relationship between learners' production and perception of the other voiced stops. Because the role of perception in L2 acquisition of Spanish voiced stops and their corresponding approximant allophones is not yet well understood, more research is needed. The present study aims to contribute to existing literature by further examining the potential relationship between L1 English-speaking L2 Spanish learners' production and perception of voiced stops and approximants.

Focusing on L2 learners' productive and perceptive abilities, this study tests some of the claims of Flege's (1995) Speech Learning Model (SLM), according to which accurate perception is a necessary precursor to accurate production in L2 phonological acquisition. Based on this claim, it can be hypothesized that the learners in the present study will have difficulty producing Spanish voiced approximants different from their voiced stop counterparts if they do not perceive a difference between the target sounds in different phonetic contexts. The third, fifth, and seventh hypotheses of the SLM are most relevant to the current study. The third hypothesis claims that greater phonetic dissimilarity between L1 and L2 sounds facilitates accurate perception because it is more likely that differences between dissimilar L1 and L2 sounds will be noticed. Certain allophonic/phonemic-graphemic relationships in English and Spanish highlight the relevance of the SLM's third hypothesis to the present study. For example, [ð] is represented orthographically by "d" in Spanish but /ð/ is represented orthographically by "th" in English and intervocalic "d" in English, as in "la<u>dd</u>er", is often realized as a similar sound to intervocalic "r" in Spanish, as in "pa<u>r</u>a" (Zampini, 1994). Although the graphemic representation of [ð]-/ð/ differs in Spanish and English, the perceptual similarity between intervocalic "d" in English and the Spanish tap /ɾ/ may not facilitate L1 English-speaking L2 Spanish learners' perception of Spanish intervocalic [ð] (Zampini, 1994). It is also possible that mismatches in phoneme/allophone-grapheme correspondences between English and Spanish will cause transfer from English, hindering learners' perception of target bilabial sounds in Spanish. For example, /b/ and /v/ are represented orthographically by "b" and "v", respectively, in English, while [b] and [β] are represented orthographically by "b" or "v" in Spanish (Face & Menke, 2009). In addition, similarities in voice onset time (VOT) between Spanish word-initial voiced stops, which are prevoiced, and English word-initial voiced stops, which may be produced with either prevoicing or short-lag VOT (Lisker & Abramson, 1964) may further contribute to L1 English-speaking L2 Spanish learners' perception difficulty of voiced stops. The fifth hypothesis proposes that a category may not be established for an L2 sound if it is perceived as equivalent to an L1 sound. Given the bimodal voicing distribution reported for word-initial voiced stops in American English by Lisker and Abramson, it is possible that L1 English-speaking L2 Spanish learners will not establish a separate category for word-initial voiced stops in Spanish. Finally, the seventh hypothesis claims that an L2 sound will be produced according to how it is perceived (i.e., how it is represented in the phonetic category). Accordingly, it may be predicted that Spanish voiced approximants will not be produced in a target-like fashion if they are not perceived as approximants, contextually dependent allophonic variants of voiced stops.

The present study addresses the following research questions.

1. Do L1 English-speaking L2 Spanish learners enrolled in courses at different points in a university curriculum show development, defined as movement toward target-like norms, in the production of Spanish voiced stops and approximants?
2. How does L1 Spanish speakers' and L2 Spanish learners' perception of the Spanish voiced stop-approximant alternation compare?
3. Are L1 English-speaking L2 Spanish learners' production and perception of Spanish voiced stops and approximants related?

2. Methodology

2.1. Participants

Twenty-three L1 English-speaking L2 Spanish learners enrolled in courses at different points in the Spanish curriculum at a university in the Midwestern United States participated in the present study. Learners were enrolled in the following Spanish courses: first-year (second-semester) language course that meets four times per week for three hours and 20 minutes each week, third-year (fifth-semester) composition and conversation course that meets two times per week for three hours and 50 minutes each week, and fourth-year (seventh-semester) upper-level elective literature course that meets two times per week for two hours and 30 minutes each week. None of these courses addressed pronunciation. Learners are classified according to the course in which they were enrolled at the time of the study to facilitate comparison of results, as most previous studies reviewed in Section 1 employed a similar classification.

Self-reported biographical data, summarized in Table 1, were obtained from learners through a questionnaire adapted from Kissling (2013), which included questions about previous and current L2 study. As reported in Table 1, L2 Spanish learners' mean age is similar across the three instruction levels. As expected, years of L2 Spanish study completed increase across instruction levels. Finally, only third- and fourth-year students have spent time abroad; however, third-year students spent very little time abroad compared to fourth-year students.

Table 1

L2 Spanish learners' background characteristics

Level of Spanish instruction	No. of subjects	Age		Gender	
		Range	Mean	Male	Female
First-year	6	18-32	22.7	3	3
Third-year	9	18-20	18.8	2	7
Fourth-year	8	19-22	20.5	4	4

Level of Spanish instruction	No. of subjects	Years of L2 study completed		Time spent abroad (in weeks)	
		Range	Mean	Range	Mean
First-year	6	.5-6.5	2	0	0
Third-year	9	3.5-8	5.8	0-3	.61
Fourth-year	8	5-15	9.8	0-24	7.4

The questionnaire adapted from Kissling (2013) also included questions about learners' Spanish language exposure and use outside of the classroom. These self-reported data are summarized in Table 2. As indicated in Table 2, all L2 learners had, on average, the greatest exposure to Spanish outside of the classroom through music. Conversely, all learners reported, on average, little exposure to Spanish outside of the classroom through TV and/or movies. While all learners infrequently used Spanish to speak with friends outside of the classroom, among learners who used Spanish to communicate at work, first-year learners, with an average of 1.5 hours per week, used it the most in that setting. Overall, L2 learners' total Spanish use outside of the classroom was quite similar across instruction levels, increasing from the first to the fourth years and from the third to the fourth years while decreasing slightly from the first to the third years.

Table 2
L2 Spanish learners' language exposure and use, reported in hours per week, outside of the classroom

Level of Spanish instruction	Watch Spanish TV/movies		Listen to Spanish music		Speak Spanish with friends		Speak Spanish at work		Total use outside of class	
	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean
First-year	0-1	.5	0-5	1.8	0-2	.8	0-4	1.5	1-12	4.7
Third-year	0-3	.9	1-6	2.4	0-2	.6	0-1	.1	1-10	4
Fourth-year	0-2	1.1	1-5	2.5	0-2	.8	0-2	.3	1-11	5.6

Four L1 Spanish speakers also participated in the study to provide a basis of comparison for the L2 learners' data. L1 Spanish-speaking participants were faculty and Ph.D. students teaching and studying at the same university as the L2 learners although they were not the instructors of any of the courses in which the learners were enrolled. Self-reported biographical data, summarized in Table 3, were obtained from L1 Spanish speakers through a questionnaire adapted from Kissling (2013). No participants, L1 Spanish speakers or L2 learners, had formal knowledge of linguistics or Spanish phonology and all participants were naïve as to the purpose of the study.

Table 3
L1 Spanish speakers' background characteristics

Speaker	Age	Gender	Birth Country	Time living in US (years)
1	58	Female	Puerto Rico	39
2	31	Female	Colombia	6
3	44	Male	Spain	20
4	40	Female	Chile	16

2.2. Materials

Given regional variation in the distribution of Spanish voiced stops and approximants, the current study examines learners' acquisition of Spanish voiced stops in absolute utterance-initial position and their corresponding approximant allophones in intervocalic position because of the lack of variation in these contexts (Alvord & Christiansen, 2012; Shively, 2008). To facilitate comparison of data by the same participants across tasks, the same words were used for the perception and production tasks. Nonce words rather than real Spanish words were used to avoid potential differences in production and/or perception due to learners' familiarity with certain words and unfamiliarity with others. The word list, which is included as an appendix, was comprised of 30 target bisyllabic nonce words (i.e., those containing a word-initial voiced stop and intervocalic voiced approximant, such as *bobe* [ˈbo.βe]) and 59 distractor nonce words that contained two to four syllables and Spanish sounds other than /b d g/ in word-initial and intervocalic positions (e.g., *treto*). All nonce words were created to respect Spanish phonotactics. In addition, stress was controlled in the creation of all nonce words such that stress fell on the penultimate or final syllable in accordance with the two general rules that govern stress assignment in Spanish. That is, stress falls on the penultimate syllable in words that end in a vowel, -n, or -s, and on the final syllable in words that end in a consonant other than -n or -s (Teschner, 2000). Exceptions to these rules are indicated with a written accent mark, as in words such as *así* 'like this/like that', *balón* 'ball', *exámenes* 'exams', and *ángel* 'angel'. Since all target nonce words, such as *bobe* [ˈbo.βe], end in a vowel, the penultimate syllable of each such word was stressed. Therefore, the intervocalic voiced approximant in all such words was in an unstressed syllable, which is a context found to favor spirantization in L1 Spanish (Cole et al., 1999; Eddington, 2011; Ortega-Llebaria, 2004) and L2 Spanish (Face & Menke, 2009; Nagle, 2017; Shively, 2008). Stress fell on either the penultimate or final syllable in distractor nonce words, such as *treto* and *lepal*, respectively. While Colantoni and Marinescu (2010) found that preceding vowel context had an effect on spirantization of intervocalic /b/ in L1 Argentine Spanish, Ortega-Llebaria (2004) did not find such an effect on spirantization of intervocalic /b/ in L1 Caribbean Spanish but she did find more lenition of intervocalic /g/ in /i/-/u/ contexts in L1 Caribbean Spanish. Given the variable results regarding preceding vowel context in L1 Spanish, this linguistic factor was not controlled in the present study. In addition, controlling vowel context in the present study would have limited the number of different target tokens at each place of articulation.

The word list was randomized so that participants did not hear and produce target sounds at the same place of articulation repeatedly with no intervening words containing different sounds from those under investigation. Ten of each of the following voiced stop/approximant pairs: [b]-[β], [d]-[ð], and [g]-[ɣ] for a total of 30 sound pairs were analyzed per L2 Spanish learner and L1 Spanish speaker, meaning that 60 sounds were analyzed per participant. However, nine tokens, between L1 Spanish speakers and L2 learners, of word-initial [b d g] were excluded because they lacked a stop burst, which made it difficult to measure VOT.

2.3. Tasks

Participants first completed the perception task followed by the production task because since the same word list was used for both tasks, if they had completed the reading task first, it is likely that their perception would have been influenced by having previously seen the words written. For the perception task, participants listened twice to a recording of an L1 Spanish speaker reading the same words used later in the production task. Each word contained a voiced stop in word-initial position and its corresponding approximant allophone in intervocalic position. Prior to completing the perception task, participants were informed that each word contained two tokens of the same consonant (e.g., two tokens of [b]) and they were to indicate on an answer sheet, which is included as an appendix, whether the two instances of the repeated consonant in each word were pronounced the same or differently.[2]

The L1 Spanish speaker that recorded the stimuli for the perception task was a professor at the same university the learners were attending but was not included in the group of L1 Spanish-speaking participants. Prior to carrying out the perception task with participants, visual inspection of waveforms and spectrograms ensured that the L1 Spanish speaker did indeed distinguish between the target sounds under investigation (i.e., Spanish voiced stops and approximants).

Similar to Kissling (2013), the production task was a word list reading. A word list reading rather than a passage reading, or a task designed to elicit spontaneous speech was used to ensure that learners at all instruction levels would be able to successfully complete the production task and to ensure that a sufficient number of tokens of target sounds would be elicited. Participants read the word list out loud while being audio-recorded, using a Marantz professional solid state recorder PMD660 and head-mounted microphone. All participants, L1 Spanish speakers and L2 learners, performed the same tasks in the same order in a quiet office with only the researcher present.

2.4. Production Data Analysis

Participants' productions were analyzed acoustically using Praat v.6.1.08 (Boersma & Weenik, 2019) signal-processing software. Following González-Bueno (1994) and Zampini (1998), VOT of word-initial [b d g] was measured in milliseconds (ms) as the interval between the release of the closure and the beginning of voicing. Since the vocal cords begin vibrating before the release burst during the production of Spanish voiced stops, the duration of this prevoicing was reported as a negative value. In tokens that did not exhibit prevoicing, VOT was measured as the duration in ms between the burst and the onset of voicing in the following vowel and reported as a positive value.

Tokens of target approximants were first coded for manner of articulation according to visual inspection of the spectrogram and waveform. Although all

[2] Participants did not see the words written when they completed the perception task so that orthography would not have affected their perception of target sounds.

target approximants appeared in intervocalic position, a context in which these sounds are consistently realized as approximants across Spanish dialects, there was variation in L2 learners', and to a much lesser extent L1 Spanish speakers', production of these sounds, as they were produced as stops, approximants, and voiceless fricatives. Following Alvord and Christiansen (2012), tokens of target approximants were coded as target-like if they "exhibited the continuation of formant structure throughout the production of the consonant accompanied by a reduction in intensity" (p. 249), as illustrated in Figure 1. Tokens of target approximants were coded as stops and hence non-target-like if they exhibited a closure evidenced by a lack of energy followed by a release burst. Tokens of target approximants were coded as fricatives and hence non-target-like if they lacked a visible closure and had a portion of glottal frication in the waveform and spectrogram. The frequency of these different productions for target approximants between L1 and L2 speakers is presented in Section 3 along with the other results.

Moreover, following Martínez Celdrán's (1991) finding that there are different degrees of approximants, as some are more open or closed than others, the degree of oral constriction of target approximants produced as such was measured using the intensity curve in Praat. Most previous L2 acquisition studies on the Spanish voiced stop/approximant alternation have analyzed learners' production from a binary perspective, reporting whether their production of approximants is target-like (Alvord & Christiansen, 2012; Díaz-Campos, 2004; Elliott, 1997; Face & Menke, 2009; González-Bueno, 1995; González-Bueno & Quintana-Lara, 2011; Kissling, 2013; Lord, 2010; Shively, 2008; Zampini, 1994). Few, more recent studies (Bongiovanni et al., 2015; Nagle, 2017; Rogers & Alvord, 2014) have examined the degree to which L2 Spanish learners spirantize intervocalic /b d g/, using relative intensity measurements. Following Rogers and Alvord (2014) and Bongiovanni et al. (2015), the intensity difference between the valley of the consonant and the peak of the following vowel was measured in the present study. "The difference in decibels (dB) between these two measurements was considered the degree of oral constriction achieved by each speaker in each specific instance measured" (Rogers & Alvord, 2014, p. 409). The greater the intensity difference between these two measurements, the less spirantization achieved by the speaker on a given token; conversely, the smaller the intensity difference, the greater degree of spirantization achieved by the speaker on a given token (Eddington, 2011; Martínez Celdrán & Regueira, 2008; Rogers & Alvord, 2014). For example, an L1 Spanish speaker's production of the target approximant [ɣ] in the nonce word *gogua* showed a high degree of spirantization with an intensity difference of 3.86 dB, while a first-year learner's production of the target approximant [ɣ] in the nonce word *gogua* showed less spirantization with an intensity difference of 14.46 dB. Figure 1 illustrates how target sounds were labeled in Praat.

Vol: 8 Issue: 4 79-110, 2020, December
ISSN:4147-9747

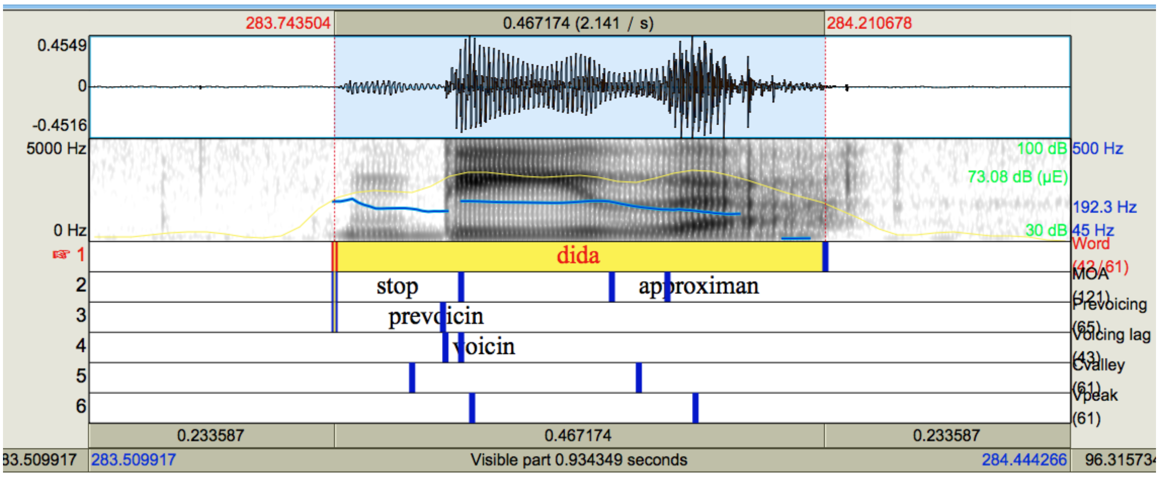

Figure 1. Example of how production of target sounds was labeled in Praat

As shown in Figure 1, which shows an L1 Spanish speaker's production of the nonce word *dida*, target sounds were coded for manner of articulation in tier 2 and VOT of target word-initial [b d g] was labeled in tier 3. The valley of target intervocalic [β ð ɣ] was marked in tier 5, while the peak of the following vowel was marked in tier 6.

2.5. Perception Data Analysis

Correct responses on the perception task were assigned 1 point (i.e., the participant chose "different") and incorrect responses were assigned 0 points (i.e., the participant chose "same"). Each group of participants' points were then added to determine the frequency with which they perceived a difference between target voiced stops and approximants.

2.6. Statistical Analysis

Descriptive statistics for the relevant phonetic cues of target sounds produced by L2 Spanish learners and L1 Spanish speakers were obtained. Specifically, the mean and standard deviation for the following measurements were obtained for the production data: VOT in ms for target word-initial [b d g] and intensity difference between the valley of target intervocalic [β ð ɣ] and the following vowel peak in dB. Mean accuracy percentages were also calculated for each group of participants' points on the perception task.

Analyses of variance (ANOVA) were performed on the data using SPSS (Statistical Package for the Social Sciences, 2019) to answer the first research question about development, defined as movement toward target-like norms, in learners' production of target sounds. Specifically, separate two-factor (one between-subjects factor: instruction level and one within-subjects factor: sound) ANOVAs were performed on the production data, with mean VOT of word-initial [b d g] and mean intensity difference between the valley of intervocalic [β ð ɣ] and the following vowel peak, respectively, as the dependent variable.[3] A two-factor (one between-subjects factor: speaker level and one within-subjects factor: sound) ANOVA was also performed on participants' mean perception accuracy to answer the second research question about how L1 Spanish speakers' and L2 Spanish learners' perception of the Spanish voiced stop-approximant alternation compares.[4] Post-hoc Tukey HSD tests were subsequently carried out to determine the nature of significant findings revealed by the ANOVAs.

Finally, in order to determine whether L2 Spanish learners' production and perception of voiced stops and approximants are related, as asked in the third research question, Pearson correlation analyses were performed on the data. Specifically, learners' mean VOT of word-initial [b d g] and mean perception accuracy were correlated by instruction level (i.e., first year, third year, fourth year) and individual sound/sound pair. Similarly, learners' mean intensity difference for target intervocalic [β ð ɣ] and mean perception accuracy were correlated by instruction level and individual sound/sound pair. The level of significance was preset at 0.05 for all statistical analyses.

3. Findings

3.1. Production Results

Descriptive statistics for the relevant phonetic cues of target sounds produced by L2 Spanish learners and L1 speakers are reported in Tables 4 and 7, while the results of the ANOVAs on the production data are reported in Tables 5 and 8.

Table 4

Descriptive statistics for VOT of Spanish word-initial [b d g] in ms by speaker level

Speaker level	Sound	VOT Mean (ms)	N	Standard Deviation
L1 Spanish speakers	[b]	-73.93	40	35.43
	[d]	-73.73	40	34.49
	[g]	-37.10	39	46.35
	Total	-61.79	119	42.42
First-year learners	[b]	1.42	55	36.84
	[d]	-0.28	60	41.83
	[g]	12.41	59	44.07
	Total	4.56	174	41.28
	[b]	-6.38	90	44.87

[3] Since how small or large the intensity difference between the valley of [β ð ɣ] and the peak of the following vowel is of greater interest than the actual peak and valley measurements, the mean peak and valley measurements are not reported in Section 3 and were not included in the ANOVA.

[4] Speaker level refers to L1 Spanish speaker, first-year learner, third-year learner, fourth-year learner.

Third-year learners	[d]	-2.83	90	40.44
	[g]	12.81	90	39.17
	Total	1.20	270	42.24
Fourth-year learners	[b]	-26.44	78	53.14
	[d]	-24.99	80	54.72
	[g]	-12.98	80	50.86
	Total	-21.42	238	53.05

As observed in Table 4, L2 Spanish learners did not produce word-initial [b d g] with a similar degree of prevoicing as L1 Spanish speakers. First-year learners generally consistently produced word-initial voiced stops with voicing lag, while third-year learners produced some of these sounds – [b] and [d] – with prevoicing. Only fourth-year learners consistently produced word-initial [b d g] with prevoicing. However, both third- and fourth-year learners produced word-initial voiced stops with much less prevoicing than the L1 Spanish speakers in this study.

The results of the two-factor (one between-subjects factor: instruction level and one within-subjects factor: sound) ANOVA on participants' mean VOT of word-initial [b d g] are reported in Table 5.

Table 5
Results of the ANOVA on mean VOT of word-initial [b d g]

Source of Variation	Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Instruction level	42568.425	3	14189.475	13.269	.000*	.366
Sound	5942.073	2	2971.036	2.778	.069	.075
Instruction level x Sound	851.920	6	141.987	.133	.992	.011
Error	73784.306	69	1069.338			

$p < 0.05$

The ANOVA revealed that instruction level (F (3, 69) = 13.269, p < .000, η^2 = .366) was a significant factor, as participants' mean VOT of word-initial [b d g] varied according to their level of instruction.[5] However, participants' mean VOT of word-initial [b d g] did not differ significantly based on the individual sound (F (2, 69) = 2.778, p = .069, η^2 = .075). In addition, the interaction between instruction level and sound (F (6, 69) = .133, p = .992, η^2 = .011) was not significant, indicating that the effect of instruction level on

[5] Citing Cohen (1988), Eddington (2015) states that "for ANOVA, partial eta[2] values around .01 show a weak effect, those around .06, a medium effect, and values of about .14 and larger, a large effect" (p. 66). Based on this, the value of .366 indicates a large effect size.

participants' mean VOT does not depend on the individual target voiced stop. Post-hoc Tukey HSD tests revealed that L1 Spanish speakers' mean VOT was significantly different from first- ($p < .000$), third- ($p < .000$), and fourth-year learners' ($p < .003$), as L1 Spanish speakers produced word-initial voiced stops with significantly more prevoicing than any group of L2 learners. However, post-hoc Tukey HSD tests did not find significant differences among L2 Spanish learners at different instruction levels. Fourth-year learners did not produce word-initial [b d g] with significantly more prevoicing than third- ($p = .074$) or first-year learners ($p = .070$) and third-year learners did not produce word-initial [b d g] with significantly more prevoicing than first-year learners ($p = .992$). These findings will be discussed in the context of the first research question, presented in Section 1, and findings of previous studies in Section 4.

Having presented the results of the acoustic analysis on Spanish word-initial voiced stops, we will now turn our attention to the results of the acoustic analysis on Spanish intervocalic voiced approximants. As stated in Section 2, although all target approximants appeared in intervocalic position, a context in which these sounds are consistently realized as approximants across Spanish dialects, there was variation in L2 learners', and to a much lesser degree L1 Spanish speakers', production of these sounds, as they were produced as stops, approximants, and voiceless fricatives. Table 6 shows the frequency of different productions for target approximants produced by L1 Spanish speakers and L2 learners. While low, the overall frequency with which learners produced target approximants as such is comparable to or higher than the results of previous studies (Díaz-Campos, 2004; Face & Menke, 2009; Lord, 2010; Shively, 2008; Zampini, 1994) that examined L2 acquisition of Spanish voiced approximants by similar groups of learners.

Table 6
Frequency of different productions for target approximants

Participant level	Manner of articulation
L1 Spanish speakers	Approximant: 103/119; 87%
	Stop: 16/119; 13%
	Voiceless fricative: 0/119; 0%
First-year learners	Approximant: 40/174; 23%
	Stop:126/174; 72%
	Voiceless fricative: 8/174; 5%
Third-year learners	Approximant: 73/270; 27%
	Stop: 196/270; 73%
	Voiceless fricative: 1/270; 0.37%
Fourth-year learners	Approximant: 98/238; 41%
	Stop: 140/238; 59%
	Voiceless fricative: 0/238; 0%

In addition to the frequency with which L2 learners produced target approximants as such, it is also important to examine the degree to which learners' approximant productions were spirantized. Table 7 reports the mean intensity difference between the valley of [β ð ɣ] and the peak of the following vowel.

Table 7
Descriptive statistics for intensity difference between valley of Spanish intervocalic [β ð ɣ] and following vowel peak in dB by speaker level

Speaker level	Sound	Mean Intensity Difference (dB)	N	Standard Deviation
L1 Spanish speakers	[β]	11.62	40	3.64
	[ð]	9.96	40	4.12
	[ɣ]	12.24	39	5.09
	Total	11.26	119	4.39
First-year learners	[β]	15.39	55	5.11
	[ð]	9.78	60	3.88
	[ɣ]	16.18	59	6.44
	Total	13.72	174	5.96
Third-year learners	[β]	15.58	90	5.76
	[ð]	12.02	90	5.86
	[ɣ]	17.02	90	6.54
	Total	14.87	270	6.40
Fourth-year learners	[β]	12.33	78	6.09
	[ð]	10.74	80	4.41
	[ɣ]	15.32	80	5.53
	Total	12.80	238	5.69

As indicated in Table 7, among all participants, L1 Spanish speakers generally spirantized intervocalic /b d g/ the most, for all individual target approximants except for [ð] and for all target approximants grouped together. L1 Spanish speakers produced [β] (11.62 dB) and [ɣ] (12.24 dB) with the lowest mean intensity differences and thus a higher degree of spirantization than any group of L2 learners. Among the learner data, two interesting patterns emerge. First, first-year learners produced target approximants with more spirantization than third-year learners, as evidenced by first-year learners' lower mean intensity difference for each individual target approximant and for all approximants grouped together. Second, among all learners, fourth-year learners produced target approximants, with the exception of [ð], with the most spirantization given their lowest mean intensity differences. Although first-, third-, and fourth-year learners generally achieved less spirantization than L1 Spanish speakers, their overall mean intensity differences for [β ð ɣ] (13.72 dB, 14.87 dB, and 12.80 dB, respectively) were not considerably greater than the L1 Spanish speakers' overall mean intensity difference of 11.26 dB. These findings will be discussed in the context of the first research question, presented in Section 1, and findings of previous studies in Section 4.

The results of the two-factor (one between-subjects factor: instruction level and one within-subjects factor: sound) ANOVA on participants' mean intensity difference of intervocalic [β ð ɣ] are reported in Table 8.

Table 8
Results of the ANOVA on mean intensity difference of intervocalic [β ð ɣ]

Source of Variation	Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Instruction level	123.264	3	41.088	2.847	.044*	.110
Sound	262.503	2	131.252	9.095	.000*	.209
Instruction level x Sound	40.376	6	6.729	.466	.831	.039
Error	995.795	69	14.432			

$^{*}p < 0.05$

The ANOVA revealed a significant main effect of instruction level (F (3, 69) = 2.847, p < .044, η^2 = .110) and of sound (F (2, 69) = 9.095, p < .000, η^2 = .209) on participants' mean intensity difference of intervocalic [β ð ɣ].[6] However, the interaction between instruction level and sound (F (6, 69) = .466, p = .831, η^2 = .039) was not significant, indicating that the effect of instruction level on participants' mean intensity difference does not depend on the individual target approximant. Post-hoc Tukey HSD tests revealed that only L1 Spanish speakers' mean intensity difference was significantly different from third-year learners' (p < .038). L1 Spanish speakers produced intervocalic [β ð ɣ] with a significantly lower mean intensity difference and hence a greater degree of spirantization than third-year learners but not than first- or fourth-year learners. Post-hoc Tukey HSD tests did not find significant differences between L1 Spanish speakers and first-year learners (p = .329) or between L1 speakers and fourth-year learners (p = .662). In addition, post-hoc Tukey HSD tests did not find significant differences among L2 Spanish learners at different instruction levels. Fourth-year learners did not produce intervocalic [β ð ɣ] with a significantly lower mean intensity difference and hence a greater degree of spirantization than third- (p = .219) or first-year learners (p = .882) and third-year learners did not produce intervocalic [β ð ɣ] with significantly more spirantization than first-year learners (p = .728). Finally, post-hoc Tukey HSD tests revealed a significant difference between participants' mean intensity difference of [β] and [ð] (p < .011) and between their mean intensity difference of [ð] and [ɣ] (p < .000), but not of [β] and [ɣ] (p = .239). Participants produced intervocalic [ð] with lower mean intensity differences and hence more spirantization than [β] and [ɣ]. These findings will be discussed in the context of the first research question, presented in Section 1, and findings of previous studies in Section 4.

[6] Citing Cohen (1988), Eddington (2015) states that "for ANOVA, partial eta[2] values around .01 show a weak effect, those around .06, a medium effect, and values of about .14 and larger, a large effect" (p. 66). Based on this, the values of .110 and .209, respectively, indicate a large effect size.

3.2. Perception Results

Figure 2 presents participants' mean accuracy in perceiving a difference between target voiced stops and approximants, while the results of the ANOVA on the perception data are reported in Table 9.

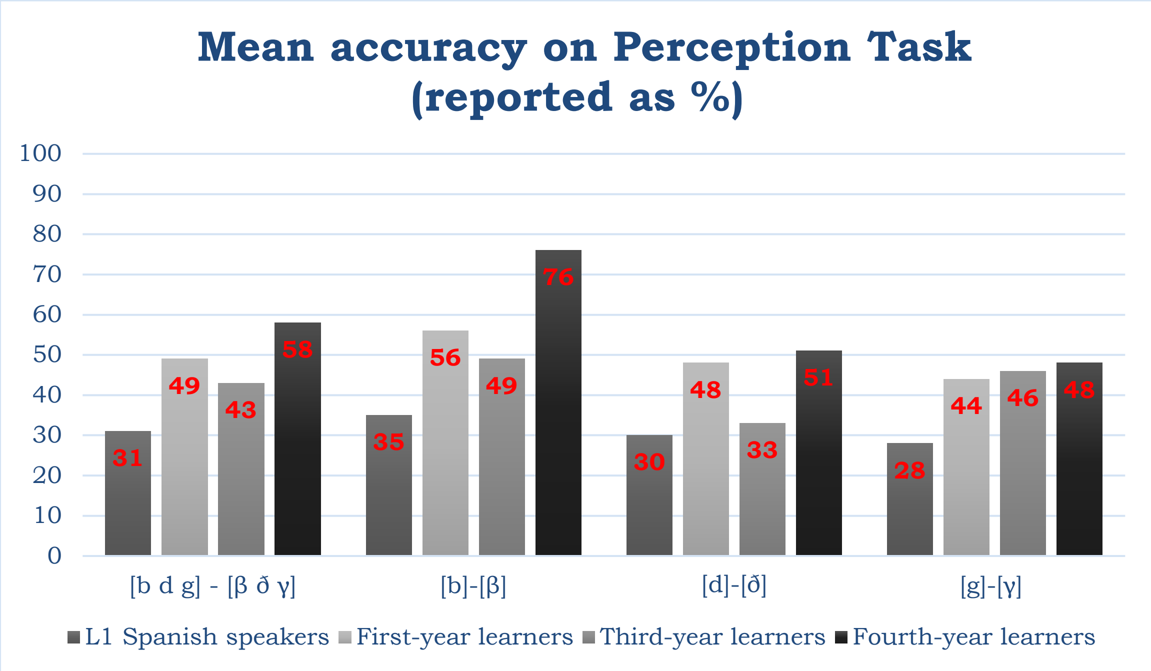

Figure 2. Mean accuracy on perception task

As observed in Figure 2, for each individual sound pair – except for [g]-[ɣ] – and for all target sounds grouped together, learners' perception accuracy decreased from the first to the third years of instruction but then increased from the third to the fourth years of instruction. In addition, L1 Spanish speakers' mean perception accuracy is notably lower than L2 learners' mean perception accuracy. These findings will be discussed in the context of the second research question, presented in Section 1, in Section 4.

The results of the two-factor (one between-subjects factor: speaker level and one within-subjects factor: sound) ANOVA on participants' mean perception accuracy are reported in Table 9.

Table 9
Results of the ANOVA on participants' mean perception accuracy

Source of Variation	Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Speaker level	.686	3	.229	3.160	.030*	.121
Sound	.318	2	.159	2.198	.119	.060
Speaker level x Sound	.187	6	.031	.430	.857	.036
Error	4.993	69	.072			

*$p < 0.05$

The ANOVA revealed a significant main effect of speaker level (F (3, 69) = 3.160, p < .030, η^2 = .121) on participants' mean perception accuracy, suggesting that L1 Spanish speakers' and L2 learners' mean perception accuracy differed significantly.[7] However, there was not a significant main effect of sound (F (2, 69) = 2.198, p = .119, η^2 = .060) on participants' mean perception accuracy and the interaction between speaker level and sound (F (6, 69) = .430, p = .857, η^2 = .036) was not significant, indicating that the effect of speaker level on participants' mean perception accuracy does not depend on the individual target voiced stop-approximant sound pair. Post-hoc Tukey HSD tests revealed that only L1 Spanish speakers' mean perception accuracy was significantly different from fourth-year learners' (p < .028). Fourth-year learners perceived a difference between target voiced stops and approximants significantly more accurately than L1 Spanish speakers but not than first- (p = .812) or third-year learners (p = .174). Post-hoc Tukey HSD tests did not find significant differences between first- and third-year learners (p = .750) or between L1 Spanish speakers and first-year learners (p = .208) or between L1 Spanish speakers and third-year learners (p = .603). Finally, post-hoc Tukey HSD tests did not reveal a significant difference between participants' mean perception accuracy of the [b]-[β] and [d]-[ð] sound pairs (p = .085), of the [b]-[β] and [g]-[ɣ] sound pairs (p =.139), or of the [d]-[ð] and [g]-[ɣ] sound pairs (p = .971).

3.3. Correlation Analyses on the Production and Perception Data

The results of the correlation analyses by learner instruction level are presented in Tables 10 and 11.

Table 10

Results of Pearson correlation analysis on L2 learners' mean VOT and mean perception accuracy

Level of instruction	Sound/Sound pair	Pearson Correlation	N	Sig.
First-year learners	[b]/[b]-[β]	-.924	6	.008*
	[d]/[d]-[ð]	-.430	6	.394
	[g]/[g]-[ɣ]	.201	6	.702
	Total	-.396	18	.104
Third-year learners	[b]/[b]-[β]	.319	9	.402
	[d]/[d]-[ð]	.598	9	.089
	[g]/[g]-[ɣ]	.644	9	.061
	Total	.456	27	.017*
Fourth-year learners	[b]/[b]-[β]	-.421	8	.298
	[d]/[d]-[ð]	-.482	8	.226
	[g]/[g]-[ɣ]	-.478	8	.231
	Total	-.435	24	.034*

**p < 0.05*

[7] Citing Cohen (1988), Eddington (2015) states that "for ANOVA, partial eta[2] values around .01 show a weak effect, those around .06, a medium effect, and values of about .14 and larger, a large effect" (p. 66). Based on this, the value of .121 indicates a large effect size.

As observed in Table 10, a negative relationship was found between first-year learners' mean VOT production of word-initial [b] and their mean perception accuracy of the sound pair [b]-[β] (*r* (4) = -.924, *p* < .008, two-tailed).[8] However, no relationship was found between first-year learners' mean VOT production of word-initial [d] and their mean perception accuracy of the sound pair [d]-[ð] nor between their mean VOT production of word-initial [g] and their mean perception accuracy of the sound pair [g]-[ɣ], respectively, as indicated by the *p* values over .05. Similarly, first-year learners' overall mean VOT production of word-initial [b d g] and their overall mean perception accuracy of all target sound pairs were not correlated, as the *p* value is above .05. The correlation analyses on the third- and fourth-year learners' data revealed a similar pattern. Specifically, no relationship was found between third- and fourth-year learners' mean VOT production of word-initial [b d g] individually and their mean perception accuracy of each corresponding sound pair, respectively, given the *p* values reported in Table 10 are over .05. However, both third-year and fourth-year learners' overall mean VOT production of word-initial [b d g] and their overall mean perception accuracy of all target sound pairs were correlated (*r* (25) = .456, *p* < .017, two-tailed) and (*r* (22) = -.435, *p* < .034, two-tailed), respectively.

Table 11
Results of Pearson correlation analysis on L2 learners' mean voiced approximant intensity difference and mean perception accuracy

Level of instruction	Sound/Sound pair	Pearson Correlation	N	Sig.
First-year learners	[β]/[b]-[β]	-.720	6	.106
	[ð]/[d]-[ð]	-.907	6	.013*
	[ɣ]/[g]-[ɣ]	.518	6	.292
	Total	-.218	18	.385
Third-year learners	[β]/[b]-[β]	-.143	9	.713
	[ð]/[d]-[ð]	.252	9	.512
	[ɣ]/[g]-[ɣ]	.276	9	.472
	Total	.166	27	.407
Fourth-year learners	[β]/[b]-[β]	-.230	8	.584
	[ð]/[d]-[ð]	-.160	8	.705
	[ɣ]/[g]-[ɣ]	.156	8	.712
	Total	-.108	24	.614

**p < 0.05*

As observed in Table 11, a negative relationship was found between first-year learners' mean intensity difference for [ð] and their mean perception accuracy of the sound pair [d]-[ð] (*r* (4) = -.907, *p* < .013, two-tailed).

[8] Eddington (2015) states "if *r* is around .1 (or .-1), the correlation is weak. Correlations around .3 (or -.3) are considered moderate, and those around .5 and greater (or -.5 and smaller) are considered to indicate a strong relationship between the two variables" (p. 29).

However, no relationship was found between first-year learners' mean intensity difference for [β] and their mean perception accuracy of the sound pair [b]-[β] nor between their mean intensity difference for [ɣ] and their mean perception accuracy of the sound pair [g]-[ɣ], respectively, as indicated by the *p* values over .05. Similarly, first-year learners' overall mean intensity difference for intervocalic [β ð ɣ] and their overall mean perception accuracy of all target sound pairs were not correlated, as the *p* value is above .05. As was the case with the results of the correlation analyses on learners' mean VOT and mean perception accuracy reported in Table 10, the correlation analyses on third- and fourth-year learners' mean voiced approximant intensity difference and mean perception accuracy followed a similar pattern. Specifically, no relationship was found between third- and fourth-year learners' mean intensity difference for [β ð ɣ] individually and their mean perception accuracy of each corresponding sound pair, respectively, since the *p* values reported in Table 11 are over .05. In addition, both third-year and fourth-year learners' overall mean intensity difference for intervocalic [β ð ɣ] and their overall mean perception accuracy of all target sound pairs were not correlated (r (25) = .166, p = .407, two-tailed) and (r (22) = -.108, p = .614, two-tailed), respectively. The results of the correlation analyses are discussed in the context of the third research question and Flege's SLM in Section 4.

4. Discussion

4.1. Research Question #1

The first research question asked whether L1 English-speaking L2 Spanish learners enrolled in courses at different points in a university curriculum show development, defined as movement toward target-like norms, in the production of Spanish voiced stops and approximants. The results of the ANOVAs, reported in Section 3.1, reveal that participants' production varied significantly according to their level of instruction. However, the difference lies between L1 Spanish speakers and L2 learners, as L1 Spanish speakers produced word-initial [b d g] with significantly more prevoicing than any group of L2 learners. Given L2 Spanish learners' well-documented difficulty acquiring voiced stops, it is not unexpected that the L2 learners in the present study produced these sounds with less prevoicing, on average, than L1 Spanish speakers. Significant differences were not found among learners, as they produced word-initial [b d g] with a similar mean VOT across instruction levels.

Although the differences between learners in the present study are not large enough to reach statistical significance, they nevertheless show development in their production of word-initial [b d g]. As indicated in Table 4, learners' mean VOT of [b] and [d], respectively, decreased from the first to the third to the fourth years of instruction, hence showing movement, albeit not significant, toward target-like norms. While there was a slight increase in learners' mean VOT of [g] from the first to the third years of instruction, there was a considerable decrease in learners' mean VOT of [g] from the third to the fourth years of instruction, showing again movement – albeit non-significant – toward more target-like norms. Learners in the present study gradually moved from production of word-initial Spanish [b d g] with short-lag voicing to more prevoicing from the first to the fourth years of instruction.

It is noteworthy that third- and fourth-year learners in this study produced word-initial [b d g] with some degree of prevoicing, as previous studies (González-Bueno, 1994; Zampini, 1998) found that L2 Spanish learners, even at intermediate and advanced levels of instruction and many of whom received explicit pronunciation instruction, failed to produce these sounds with any voicing lead. It is possible that the difference in prevoicing observed in the learners in the present study and those in González-Bueno (1994) is due to differences in the tasks used in each study, as González-Bueno (1994) relied on spontaneous speech data, while the data in the present study come from a word list reading, which may favor prevoicing due to hyperarticulation in laboratory speech. Although Zampini (1998) similarly used a reading task to assess learners' production, she did not find prevoicing perhaps because her correlation data for [b] suggest that production and perception are independent processes.

For voiced approximants in the present study, the difference again lies between L1 Spanish speakers and L2 learners, as L1 Spanish speakers produced intervocalic [β ð ɣ] with a significantly greater degree of spirantization than third-year learners only. Significant differences were not found among learners, as they produced intervocalic [β ð ɣ] with a similar degree of spirantization across instruction levels. Although the differences between learners in the present study are not large enough to reach statistical significance, they nevertheless show development in their production of intervocalic [β ð ɣ]. As indicated in Table 6, the frequency with which learners produced target approximants as such – rather than as stops or fricatives – increased from the first to the third to the fourth years of instruction. In addition, the data reported in Table 7 show that learners' mean intensity difference for [β] and [ɣ] as well as their overall mean intensity difference decreased from the first to the fourth years of instruction, revealing that fourth-year learners produced these sounds with a greater degree of spirantization than first-year learners and hence showing movement, albeit non-significant, toward more target-like norms. While there was a slight increase in learners' mean intensity difference for [β ð ɣ] individually and collectively from the first to the third years of instruction, there was a decrease in learners' mean intensity difference for these sounds from the third to the fourth years of instruction, showing again movement – albeit not significant – toward more target-like norms.

Rogers and Alvord (2014), the first previous study, to the best of my knowledge, to examine the degree to which L2 Spanish learners spirantize intervocalic /b d g/, found higher levels of spirantization in their learners who had spent two years abroad in a Spanish-speaking country as religious missionaries. It is not surprising that the abroad learners in Rogers and Alvord (2014) achieved a greater degree of spirantization than the learners in the present study, as it has been previously reported that learners who spent an extended time abroad are among the few groups of L2 learners to have any success in acquiring Spanish voiced approximants (Alvord & Christiansen, 2012). However, it is noteworthy that the overall mean intensity difference of first-year learners in the present study (13.72 dB) is comparable to the overall mean intensity difference of the AH learners in Rogers and Alvord (2014) who had completed four semesters of Spanish (13.53 dB), indicating that both groups of learners achieved a similar degree of spirantization. In addition, the first- and fourth-year learners in the present study produced intervocalic [β ð ɣ] with similar or lower overall mean intensity differences than the intermediate-level AH and SA learners enrolled in an Introduction to Hispanic Linguistics course in Bongiovanni et al. (2015), as indicated in Tables 12 and 13. In fact, Bongiovanni et al. (2015) found that participation in a short-term SA program of four weeks did not lead to more target-like production of [β ð ɣ] in SA learners compared to AH learners at Time 2. These results alongside those of Alvord and Christiansen (2012) and of Rogers and Alvord (2014) further underscore the importance of longer-term stays abroad for L2 acquisition of Spanish voiced approximants. These results also suggest that prior experience rather than a short-term SA program may favor more target-like production of voiced approximants, as the fourth-year learners in the present study had studied Spanish for 9.8 years on average compared to the AH and SA learners in Bongiovanni et al. (2015) who had studied Spanish for 4.8 and 5.5 years on average, respectively.

Table 12
Overall mean intensity difference for AH and SA learners in Bongiovanni et al. (2015) at Time 1 and Time 2

Learners	Time 1	Time 2
AH learners	13.4 dB	13.83 dB
SA learners	16.63 dB	14.6 dB

Table 13
Overall mean intensity difference for L2 learners in the present study

Level of instruction	Overall mean intensity difference
First-year learners	13.72 dB
Third-year learners	14.87 dB
Fourth-year learners	12.80 dB

4.2. Research Question #2

The second research question asked how L1 Spanish speakers' and L2 Spanish learners' perception of the Spanish voiced stop-approximant alternation compares. As observed in Figure 2, for each individual sound pair and for all target sounds grouped together, L2 Spanish learners' mean perception accuracy was higher than L1 Spanish speakers' mean perception accuracy. Although it may generally be unexpected to find L1 Spanish speakers' perception of a given sound (e.g., /p/) to be lower than L2 Spanish learners' perception, in this case, the L1 Spanish speakers' low scores on the perception task may not be that surprising since this task asked them to distinguish two sounds that are allophonic variants (e.g., [b]-[β]) of the same category (/b/), not two different categories. In fact, it has been previously documented that L1 Spanish speakers often perceive two allophones of a given phoneme, such as /d/, to be the same (Hualde, Olarrea, Escobar, & Travis, 2010). It is possible that learners are more attuned to phonetic differences since their L2 phonological system is not yet fully developed, whereas L1 speakers' phonological system is fully developed and so they are only attuned to differences in phonemes, not to differences in allophones.

As observed in Figure 2, L2 Spanish learners' perception follows a U-shaped acquisition curve (Lightbown, 1983). The decrease in accuracy from the first to the third years of instruction for each individual sound pair – except for [g]-[ɣ] – and for all target sounds grouped together, may be surprising given third-year learners' greater L2 experience but may be due to differences in their studies and L2 use outside of the classroom. Some first-year learners were music majors, as indicated on their background questionnaire, and Alvord and Christiansen (2012) found that at least five years of music instruction favored target-like pronunciation of [β ð ɣ] because "it is likely that more music instruction leads to a better ability to hear differences in tonality, which increased their capability of hearing differences in pronunciation" (p. 263). Similar to music, speech sounds also differ in tones, pitch, and intensity. Therefore, it is possible that the first-year learners' music studies facilitated their ability to perceive differences between word-initial [b d g] and intervocalic [β ð ɣ], whereas third-year learners did not report studying music. In addition, as observed in Table 2, first-year learners reported using Spanish outside of the classroom for 4.7 hours per week on average, including using it to communicate at work for 1.5 hours per week on average, compared to third-year learners who reported using Spanish outside of the classroom for four hours per week on average, including using it to communicate at work for .1 hours per week on average.

4.3. Research Question #3

The third research question asked whether L1 English-speaking L2 Spanish learners' production and perception of Spanish voiced stops and approximants are related. The results of the correlation analyses reported in Section 3.3 revealed a relationship between learners' productive and perceptive abilities depending on their instruction level and the target sound/sound pair. Based on the data in Table 4 and Figure 2, the negative correlation between first-year learners' mean VOT production of word-initial [b] and their mean perception accuracy of the sound pair [b]-[β] suggests that as their mean VOT of word-initial [b] decreases and hence becomes more target-like, their perception accuracy of the sound pair [b]-[β] increases. Similarly, based on the data in Table 7 and Figure 2, the negative correlation between first-year learners' mean intensity difference for [ð] and their mean perception accuracy of the sound pair [d]-[ð] suggests that as their mean intensity difference for [ð] decreases and hence becomes more target-like, their perception accuracy of the sound pair [d]-[ð] increases. It is likely that these significant correlations are driven by the fact that first-year learners' mean perception accuracy was highest on the sound pair [b]-[β] followed by [d]-[ð] and their mean intensity difference was the lowest for [ð], while their mean VOT was second lowest for [b]. The positive correlation between third-year learners' overall mean VOT production of word-initial [b d g] and their overall mean perception accuracy of all target sound pairs is likely driven by their individual correlations for [d]/[d]-[ð] and for [g]/[g]-[ɣ], which are approaching significance with *p* values of .089 and .061, respectively. The positive correlation between third-year learners' overall mean VOT production and overall mean perception accuracy suggests that as their mean VOT increases, their mean perception accuracy increases. Finally, the negative correlation between fourth-year learners' overall mean VOT production of word-initial [b d g] and their overall mean perception accuracy of all target sound pairs suggests that their mean perception accuracy decreases as their mean VOT increases.

The results of the present study shed further light on the potential relationship between L2 Spanish learners' production and perception of voiced stops and approximants, corroborating, in part, findings of previous studies. Through a comparison of the percentage of learner errors on production and perception tasks, González-Bueno and Quintana-Lara (2011) reported similarity between production and perception accuracy of [β] and [ɣ] in beginning, intermediate, and advanced L2 Spanish learners but did not compare learners' production and perception through correlation analyses, so it is unknown if learners' productive and perceptive abilities were indeed related. Kissling (2013) empirically documented a relationship between learners' production and perception of Spanish voiced approximants, finding a negative correlation for second-year learners' perception and production of [β] and a positive correlation for third-year learners' perception and production of [β] and [ɣ]. A correlation was similarly found between bilabial production and perception in the present study, but for production of word-initial [b] rather than intervocalic [β] and perception of the sound pair [b]-[β] in first-year rather than second- or third-year learners. Conversely, Zampini (1998) did not find a correlation between production and perception of [b] in students enrolled in a Spanish phonetics course. While González-Bueno and Quintana-Lara (2011) found that production and perception accuracy of [ð] differed the most across learner levels, implying no relationship, Kissling (2013) found a positive correlation between first-year learners' production and perception of [ð]. A correlation, but negative rather than positive, was similarly found between first-year learners' production of [ð] and their mean perception accuracy of the sound pair [d]-[ð] in the present study. Although the nature of the relationship is different, this finding nevertheless corroborates Kissling's (2013) finding that first-year learners' production and perception of [ð] are related.

Although significant relationships were not found between production and perception for all learners and all sound pairs, the significant relationships for first- and fourth-year learners, reported in Section 3.3, that were found provide support for Flege's (1995) SLM, according to which accurate perception is a necessary precursor to accurate production in L2 phonological acquisition. This tenet is specifically supported by the negative correlation between first-year learners' mean VOT production of word-initial [b] and their mean perception accuracy of the sound pair [b]-[β] since their mean VOT of word-initial [b] decreases and hence becomes more target-like as their perception accuracy of the sound pair [b]-[β] increases. The negative correlation between first-year learners' mean intensity difference for [ð] and their mean perception accuracy of the sound pair [d]-[ð] provides further support for this claim since their perception accuracy of the sound pair [d]-[ð] increases as their mean intensity difference for [ð] decreases and hence becomes more target-like. In addition, the significant correlation for fourth-year learners also corroborates this claim because their mean perception accuracy decreases as their mean VOT increases and hence becomes less target-like. However, the significant correlation for third-year learners, reported in Section 3.3, does not provide support for the claim that accurate perception precedes accurate production because for them, more accurate perception is related to an increase in mean VOT, which is actually less target-like rather than more accurate production.

In addition, predictions were made in Section 1 based on the third, fifth, and seventh hypotheses of Flege's SLM. Based on the third hypothesis that greater phonetic dissimilarity between L1 and L2 sounds facilitates accurate perception, it was predicted that the perceptual similarity between intervocalic "d" in English and the Spanish tap /ɾ/ may not facilitate L1 English-speaking L2 Spanish learners' perception of Spanish intervocalic [ð]. This hypothesis is partially supported by the results reported in Figure 2, as third-year learners' mean perception accuracy was lowest on the [d]-[ð] sound pair, while first- and fourth-year learners' mean perception accuracy was second lowest on the [d]-[ð] sound pair. However, mismatches in phoneme/allophone-grapheme correspondences between English and Spanish and similarities in VOT between Spanish and English word-initial [b d g] did not hinder, contrary to the prediction in Section 1, learners' perception of target Spanish bilabial sounds, as their perception accuracy of [b]-[β] was the highest among all target sound pairs across instruction levels. The fifth hypothesis of the SLM proposes that a category may not be established for an L2 sound if it is perceived as equivalent to an L1 sound. The results of this study appear to provide support for this hypothesis, as first-year learners generally produced word-initial [b d g] with voicing lag, while third-year learners produced some target word-initial voiced stops with prevoicing and others with voicing lag, and fourth-year learners consistently produced word-initial [b d g] with prevoicing. Based on the seventh hypothesis of the SLM, it was predicted that Spanish voiced approximants will not be produced in a target-like fashion if they are not perceived as approximants, contextually dependent allophonic variants of voiced stops. The results of this study do not confirm this prediction, as learners generally perceived a difference between target voiced stops and approximants around 50% of the time, they produced target approximants as such rather than as stops or fricatives between 23% and 41% of the time, and when they produced target approximants as such, they generally achieved a similar degree of spirantization as L1 Spanish speakers.

5. Conclusion

The present study's investigation of L2 Spanish learners' perception of voiced stops and approximants contributes to our understanding of L2 acquisition of these sounds. While many previous studies have documented L2 Spanish learners' difficulty producing voiced stops and approximants in a target-like fashion (Díaz-Campos, 2004; Elliott, 1997; González-Bueno, 1994; Kissling, 2013; Lord, 2010; Zampini, 1994, 1998), the role of perception has been understudied and not well-understood. Perception studies in general have lagged behind production studies on this topic and only three previous studies (González-Bueno & Quintana-Lara, 2011; Kissling, 2013; Zampini, 1998) have investigated L2 Spanish learners' production and perception of these sounds. González-Bueno and Quintana-Lara (2011) and Kissling (2013) examined L2 learners' production and perception of Spanish voiced approximants, while Zampini (1998) investigated L2 learners' production and perception of Spanish /p/ and /b/. The present study is the first, to the best of my knowledge, to investigate L2 Spanish learners' production and perception of the entire class of voiced stops and their respective approximant allophones. The results of this study indicate that there is a relationship between L2 learners' production and perception of Spanish voiced stops and approximants depending on their instruction level and the target sound/sound pair. The correlations found in this study point to the important role of perception in L2 Spanish learners' acquisition of voiced stops and approximants.

The results of the present study also further point to the need for explicit pronunciation instruction on these sounds in the L2 Spanish classroom. The L2 learners in this study produced word-initial [b d g] with significantly less prevoicing than L1 Spanish speakers. In addition, learners produced target approximants as such, rather than as stops or fricatives, less than 50% of the time, compared to L1 Spanish speakers who produced target approximants as such 87% of the time. It is possible that these production difficulties may be attributed to learners' perception, as their overall mean perception accuracy was close to 50% and it was generally close to 50% for each individual sound pair across instruction levels. Sustained pedagogical intervention from first-, second-, and third-year Spanish language classes to upper-level Spanish linguistics or phonetics classes may help L2 learners move toward more target-like production of Spanish voiced stops and approximants and increase their ability to perceive a difference between these sounds in different phonetic contexts.

While the present study contributes to the L2 Spanish phonological acquisition literature, it is important to acknowledge that our understanding of the present findings may be limited given the limited number of L2 learners in each group. Our understanding of the present findings may also be limited by the way in which learners were grouped because although L2 Spanish learners in their first, third, and fourth years of university study differ in terms of years of L2 study completed, time spent abroad, and L2 exposure and use outside of the classroom, as reported in Section 2.1, these background data may not be considered objective measures of L2 experience, as they were self-reported. In addition, the background data and the course in which learners were enrolled at the time of the study are not intended to be measures of L2 proficiency, but rather of L2 experience. However, a future study should use a proficiency measure, such as an oral proficiency interview, to group learners to more clearly show objective differences between learner groups.

In conclusion, given the contribution of the present study and despite its limitations, the results suggest that this topic is ripe for further research. Given the lack of correlation data on L2 Spanish learners' production and perception of voiced stops and the different nature of the correlations found between L2 Spanish learners' production and perception of voiced approximants in the present study and Kissling (2013), further research is needed to better understand this potential relationship. In addition, in light of Alvord and Christiansen's (2012) and Rogers and Alvord's (2014) findings, a future study should examine the effect of a long-term academic SA program of one semester or one year on L2 learners' production and perception of Spanish voiced stops and approximants. The effect of explicit pronunciation instruction on Spanish voiced stops and approximants and a comparison of different pronunciation pedagogical approaches (e.g., formal instruction in a Spanish language course vs. formal instruction in an Introduction to Hispanic Linguistics course or Spanish phonetics course) similarly merit further attention in future studies given the mixed results of previous studies that examined the effect of pronunciation instruction on L2 acquisition of these sounds. Another topic that warrants investigation in the future is whether differences observed acoustically in the learner data – such as significantly less prevoicing of Spanish word-initial [b d g] and the high frequency of target Spanish voiced approximants produced as voiced stops – impact how native speakers perceive learner speech. While there is a long line of SLA research on foreign accent in general (cf. Flege, Munro, & MacKay, 1995; Munro, 1993; Munro & Derwing, 1994, 1995), perception of foreign accent by native speakers has not been widely studied with respect to L2 acquisition of Spanish stops, as only one such study, to the best of my knowledge, by González-Bueno (1997) – focusing on the voiceless velar stop /k/ – has been carried out. González-Bueno found that English-accented productions of /k/, the result of long VOT and aspiration, noticeably affected native speakers' perception of the learners' productions. Given this finding and Kissling's (2013) claim that "it remains an empirical question to what extent accuracy in production and perception of … L2 … segments impacts accent, comprehensibility, and/or intelligibility" (p. 172), future studies on L2 acquisition of Spanish voiced stops and approximants as well as of other sounds, should investigate native speakers' perception of learner speech.

References

Alvord, S. M., & Christiansen, D. E. (2012). Factors influencing the acquisition of Spanish voiced stop spirantization during an extended stay abroad. *Studies in Hispanic and Lusophone Linguistics, 5*(2), 239-276.

Boersma, P., & Weenik, D. (2019). Praat: Doing phonetics by computer [Computer program]. Version 6.1.08, http://www.praat.org/.

Bongiovanni, S., Long, A. Y., Solon, M. E., & Willis, E. W. (2015). The effect of short-term study abroad on second language Spanish phonetic development. *Studies in Hispanic and Lusophone Linguistics, 8*(2), 243-283.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences.* 2nd ed. Hillsdale: Lawrence Erlbaum.

Colantoni, L., & Marinescu, I. (2010). The Scope of Stop Weakening in Argentine Spanish. In M. Ortega-Llebaria (Ed.), *Selected Proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology*, pp. 100-114.

Cole, J., Hualde, J. I., & Iskarous, K. (1999). Effects of Prosodic and Segmental Context on /g/-Lenition in Spanish. In O. Fujimura, B. D. Joseph, & B. Palek (Eds.), *Proceedings of the Fourth International Linguistics and Phonetics Conference*, pp. 575-589.

Díaz-Campos, M. (2004). Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition, 26*(2), 249-273.

Eckman, F. R. (1987). Markedness and the contrastive analysis hypothesis. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage Phonology: The acquisition of a second language sound system* (pp. 55-69). New York: Newbury House/Harper & Row.

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

Eddington, D. (2015). *Statistics for Linguists*: *A Step-by-Step Guide for Novices.* Newcastle upon Tyne: Cambridge Scholars Publishing.

Elliott, R. A. (1997). On the teaching and acquisition of pronunciation within a communicative approach. *Hispania, 80*(1), 95-108.

Face, T. L., & Menke, M. R. (2009). Acquisition of Spanish voiced spirants by second language learners. In J. Collentine, M. García, B. Lafford, & F. Marcos Marín (Eds.) *Selected Proceedings of the 11th Hispanic Linguistics Symposium* (pp. 39-52). Somerville: MA: Cascadilla Proceedings Project.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 233-277). Timonium, MD: York Press.

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

González-Bueno, M. (1994). *Effects of formal instruction on the improvement in the pronunciation of Spanish stops by second language learners: Changes in voice onset time in initial stops /p, t, k/ and /b, d, g/* (Doctoral Dissertation). Retrieved from ProQuest Dissertations and Theses. (9504216)

González-Bueno, M. (1995). Adquisición de los alófonos fricativos de las oclusivas sonoras españolas por aprendices de español como segunda lengua. *Estudios de lingüística aplicada, 13*, 64-79.

González-Bueno, M. (1997). Voice-onset-time in the perception of foreign accent by native listeners of Spanish. *International Review of Applied Linguistics, 35*(4), 251-267.

González-Bueno, M., & Quintana-Lara, M. (2011). Students' awareness of Spanish spirantization allophonic rule. In *Proceedings of the 2nd Annual Pronunciation in Second Language Learning and Teaching Conference,* 179-197.

Hualde, J. I., Olarrea, A., Escobar, A. M., & Travis, C. (2010). *Introducción a la lingüística hispánica* (2ª ed.). New York: Cambridge University Press.

IBM Corp. (2019). IBM SPSS Statistics for Macintosh, Version 26.0. Armonk, NY: IBM Corp.

Jakobson, R. (1968). *Child Language Aphasia and Phonological Universals.* Netherlands: Mouton.

Kissling, E. M. (2013). *The effect of phonetics instruction on adult learners' perception and production of L2 sounds* (Doctoral Dissertation). Retrieved from ProQuest Dissertations and Theses. (201303826)

Lightbown, P. M. (1983). Exploring relationships between developmental and instructional sequences in L2 acquisition. In H. Seliger & M. Long (Eds.), *Classroom-oriented research in language acquisition* (pp. 217-243). Rowley, MA: Newbury House.

Lisker, L., & Abramson, A. S. (1964). A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *Word, 20*(3), 384-422.

Lord, G. (2010). The combined effects of immersion and instruction on second language pronunciation. *Foreign Language Annals, 43*(3), 488-503.

Martínez Celdrán, E. (1991). Sobre la naturaleza fonética de los alófonos de /b, d, g/ en español y sus distintas denominaciones. *Verba, 18*(1), 235-253.

Martínez Celdrán, E., & Regueira, X. L. (2008). Spirant approximants in Galician. *Journal of the International Phonetic Association, 38*(1), 51-68.

Munro, M. J. (1993). Productions of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. *Language and Speech, 36,* 39-66.

Munro, M., & Derwing, T. (1994). Evaluations of foreign accent in extemporaneous and read material. *Language Testing, 11*, 253-266.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45,* 73-97.

Nagle, C. L. (2017). Individual developmental trajectories in the L2 acquisition of Spanish spirantization. *Journal of Second Language Pronunciation, 3*(2), 218-241.

Ortega-Llebaria, M. (2004). Interplay between phonetic and inventory constraints in the degree of spirantization of voiced stops: Comparing intervocalic /b/ and intervocalic /g/ in Spanish and English. In T. L. Face (Ed.), *Laboratory Approaches to Spanish Phonology*, pp. 237-253.

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

Shively, R. L. (2008). L2 Acquisition of [β], [ð], [ɣ] in Spanish: Impact of Experience, Linguistic Environment, and Learner Variables. *Southwest Journal of Linguistics,* 27(2), 79-114.
Teschner, R. V. (2000). *Camino oral: Fonética, fonología y práctica de los sonidos del español* (2nd ed.). Boston, MA: McGraw-Hill.
Zampini, M. L. (1994). The role of native language transfer and task formality in the acquisition of Spanish spirantization. *Hispania,* 77(3), 470-481.
Zampini, M. L. (1998). The relationship between the production and perception of L2 Spanish stops. *Texas Papers in Foreign Language Education, 3*(3), 85-100.

Appendix I
Word list used for the production and perception tasks

Table 14
Test items organized by sound pair

Target sound pair [b]-[β]	Target sound pair [d]-[ð]	Target sound pair [g]-[ɣ]
bobe	dadu	gogu
bobi	dido	gagu
bibo	dade	gugu
bube	dode	gugo
biba	dudi	gogo
bubu	dedi	goga
babo	dida	gogua
babe	deda	gagua
babu	dudu	guga
bibe	dadi	gugua

Table 15
Distractor words

nenada	mastito
chechu	tator
pepo	copica
papen	naneco
preposa	nuna
nenato	tometo
refunfusa	tutila
trita	memer
treto	titela
memiso	pepeto
mamir	pechacha
pepato	lala
caqueta	tecaco
copaco	pipato
lepal	sistre
fufe	tatara
lecal	fefa
capaco	tetal
nonuco	titubo
tetor	copeca
compuco	sesmo
muma	papeto
titera	farfula
pripesa	nineco
sustre	tital
sesar	popato

cantaco	cocao
cinasa	mumar
sosta	popila
rarota	

Appendix II
Answer sheet used for the perception task

You will hear 89 words in Spanish two times. Each word contains two tokens of the same consonant (e.g., two tokens of “b”). Please indicate if you think the two instances of the repeated consonant (e.g., “b”) in each word were pronounced the same or differently.

1. Same Different

2. Same Different

3. Same Different

4. Same Different

5. Same Different

6. Same Different

7. Same Different

8. Same Different

9. Same Different

10. Same Different

Continued in the same way until 89.