

# PHONETIC TRAINING OF L2 SPANISH /u/\*

Andrew McCandless  
University of Toronto

## 1. Spanish and English vowels

Spanish contains five vowels in its inventory: a low central vowel, /a/, a medium front unrounded vowel, /e/, a high front unrounded vowel, /i/, a medium back rounded vowel, /o/, and a high back rounded vowel, /u/ (Schwegler et al. 2010, Hualde 2014). English contains many more vowels than Spanish. Standard American English contains 12 vowels: two low vowels, /æ/ and /a/, three medium open vowels, /ɛ/, /ʌ/ and /ɔ/, two medium closed vowels, /e/ and /o/, two high open vowels, /i/ and /u/, and a central medium vowel, /ə/ (Schwegler et al. 2010, Hualde 2014).

Five English vowels are similar to their Spanish counterparts (/æ/, /e/, /i/, /o/, /u/). These English vowel phonemes are articulated with longer duration than Spanish, and four of them (/e/, /i/, /o/, /u/) are often articulated as diphthongs. They are higher (lower F1, /e/, /i/) or lower (higher F1, /æ/, /o/, /u/) than their Spanish counterparts, and are more fronted (higher F2, /æ/, /e/, /i/, /o/, /u/) than their Spanish counterparts (Menke and Face 2010, Schwegler et al. 2010, Hualde 2014).

## 2. Spanish /u/

Of the five Spanish vowels, the vowel /u/ is one of the more difficult for L2 learners with L1 English to produce in a native-like manner. L2 Spanish learners with L1 English may substitute the Spanish shorter, higher, back vowel /u/ for the English longer, lower, fronted vowel /u/ in production, due to cross-linguistic category assimilation (Cobb and Simonet 2015).

The vowel /u/ can occur in unstressed syllables:

- (1) a. \*tribu ‘tribe’
- b. lucir ‘to shine, to wear, to show off’.

or in stressed syllables:

- (2) a. \*azúcar ‘sugar’
- b. lucha ‘fight, struggle’.

---

\* This paper was presented at the 2021 annual conference of the Canadian Linguistic Association, held virtually, June 5, 2021, and at the American Association of Teachers of Spanish and Portuguese (Ontario Chapter) annual conference, held at University of Toronto, November 9, 2019. The presentation was based on a course assignment for my PhD program in Hispanic Linguistics at University of Toronto, submitted December 19, 2016. I would like to thank Anabela Rato for her help and comments on this paper.

It can appear following a labial (bilabial) consonant (/p/ or /b/):

- (3) a. \**apuntes* ‘notes’  
 b. *buñuelo* ‘donut’.

or following a coronal (dental) consonant (/t/ or /d/):

- (4) a. \**atun* ‘tuna’  
 b. *dureza* ‘harshness’.

or following a dorsal (velar) consonant (/k/, /g/ or /x/):

- (5) a. \**ocupado/a* ‘busy’ [masculine/feminine]  
 b. *laguna* ‘lagoon’  
 c. *juntarse* ‘to get together’.

## 2.1 L2 Spanish perception and production of /u/

Following the Speech Learning Model (SLM) (Flege 1995), when an L2 sound and a corresponding L1 sound are perceptually similar, L2 learners may have relatively greater difficulty perceiving differences between the L2 sound and the L1 sound. In turn, they may have relatively greater difficulty differentiating the L2 and L1 sounds in production.

Previous studies (ex. Morrison 2003) have provided evidence that L2 Spanish learners with L1 English may perceive and produce Spanish /u/ as English /u/ or /ʊ/. However, with increased L2 proficiency, production of /u/ (F1 and F2) may become more native-like (Morrison 2003, Menke and Face 2010, Cobb and Simonet 2015).

Unstressed /u/ may be more centralized (higher F2) in L2 Spanish learners than in native speakers (Menke and Face 2010). Compared to L1 General American speakers, L1 Canadian English speakers may have more difficulty producing Spanish /u/ with native-like F2, since /u/ and /ʊ/ are more centralized in Canadian English than in General American (Hagiwara 2006).

## 3. Phonetic training<sup>1</sup>

Previous studies have provided evidence that phonetic training (perception and/or production training) may be effective for improving perception and production of a variety of L2 segments in many L1 groups, such as for L2 Spanish rhotics in L1 American English speakers (ex. Bradlow et al. 1997, 1999; Herd et al. 2013; Kartushina et al. 2015; Sakai 2016). The main objectives of phonetic training are to evaluate the following: L2 learners’ degree of pre- to post-test improvement, transfer of improvements from the trained modality to the opposite modality (perception to production, or vice versa), generalization of improvements from training to new contexts, new speakers, and

---

<sup>1</sup>This section is similar in content to my 2020 proceedings paper (McCandless 2020).

new stimuli, and long-term retention of these improvements with training.<sup>2</sup> Training can be short term (one session) or long term (multiple sessions), with long-term being much more common than short-term training (Logan and Pruitt 1995, Sakai and Moorman 2017).

Improvements with training are most commonly evaluated through a pre- and post-test. Tasks and stimuli at pre- and post-test are identical to each other, and are similar to those used in training. Results are most commonly measured based on percentage of correct responses at pre- and post-test. However, they can also be measured based on improvement and gain scores, response times, rating scales or evoked potentials (Logan and Pruitt 1995; Bradlow et al. 1997, 1999; Herd et al. 2013; Kartushina et al. 2015; Sakai 2016; Sakai and Moorman 2017).

Two types of perception training and testing tasks used in previous studies are discrimination tasks, in which two or three stimuli are presented in each trial and the participant differentiates them auditorily, and identification tasks, in which one stimulus is presented in each trial and the participant identifies it from a closed set of responses (Logan and Pruitt 1995; Bradlow et al. 1997, 1999; Herd et al. 2013; Kartushina et al. 2015; Sakai 2016; Sakai and Moorman 2017). Three types of production training and testing tasks used in previous studies are elicited production tasks, in which participants speak segments/words out loud, passage reading tasks, and picture description tasks (Bradlow et al. 1997, 1999; Herd et al. 2013; Kartushina et al. 2015; Sakai 2016).

Feedback from training is very important to ensure that learning occurs, because the time period for training is limited. The most common (and perhaps, the most important) type of feedback for effective learning from training is immediate, or trial-by-trial, feedback. In addition, cumulative feedback (provided after each block of trials or at the end of each session) may also motivate participants to continue with training (although this may be less essential to learning from training than immediate feedback). Most frequently, feedback contains information on whether participants' responses are correct or incorrect, but it can also include information on their response times. It is only provided during training, not during the pre- or post-test (Logan and Pruitt 1995).

#### **4. Research question and hypotheses**

The research question for this study is the following: With phonetic training (i.e., perception and production training), do low-proficiency L2 Spanish learners with L1 Canadian English achieve more native-like perception and production of Spanish /u/?

The following hypotheses were made:

(1) Before phonetic training, low-proficiency L2 Spanish learners with L1 Canadian English would perceive and produce Spanish /u/ in a way that is more similar to the English /u/ or /ʊ/ than the Spanish /u/. However, with training, these learners would perceive and produce Spanish /u/ in a more native-like way.

---

<sup>2</sup> Typically, long-term retention of improvements with training is investigated several months after post-test (e.g. Bradlow et al. 1999). Due to the severe time limitations of this study (it was completed for a course assignment), it was not possible to assess long-term retention of pre- to post-test improvements.

(2) Before perception (discrimination) training, low-proficiency L2 Spanish learners with L1 Canadian English would perceive Spanish /u/ as the same vowel as English /u/ or /ʊ/, when these vowels are paired as different stimuli. However, with training, these learners would perceive Spanish /u/ and English /u/ and /ʊ/ more often as different vowels.

(3) Both before and after training, these learners will perceive Spanish /u/ as the same vowel as Spanish /u/, when these vowels are paired as same stimuli, produced by different native speakers.

(4) Before production (elicited imitation) training, low-proficiency L2 Spanish learners with L1 Canadian English would produce Spanish /u/ after a labial, coronal, or dorsal consonant with longer duration, higher F1 (i.e. as a lower vowel) and higher F2 (i.e. as a more fronted vowel) relative to native speaker norms for this vowel, similar to the English /u/. However, with training, these learners would produce Spanish /u/ with more native-like (decreased) duration, F1 and F2.

## **5. Methodology**

### **5.1 Participants**

Two L1 Canadian English speakers with beginner L2 Spanish, P001 (age 67, female) and P002 (age 71, female), were recruited for this study. Both participants were from a small town in Ontario, near the Greater Toronto Area. Neither participant had formal instruction in Spanish before adolescence, nor did either participant have proficiency above beginner level in another L2. P001 was assigned to the experimental condition (phonetic training), and P002 was assigned to the control condition (no training).

### **5.2 Stimuli**

Spanish stimuli for the perception test and training tasks in this study were extracted from the *University of Toronto Romance Phonetics Database* (Colantoni and Steele 2004). More specifically, they were extracted from the Dialect Atlas of Argentina (dialectological interviews with native speakers from different regions of Argentina) and the Romance Language Survey (speech samples of Spanish, Portuguese, French, Italian and Romanian, collected in Toronto from native speakers and L2 learners). Speakers of stimuli were native speakers of Spanish, ranging in age (19 to 65), in gender (male and female), and in variety (Mexico, Colombia, Argentina, Chile, El Salvador). Also, English stimuli for the perception test and training tasks, pre-recorded by the experimenter (a young adult male native Canadian English speaker) were used for the perception test and training tasks. Stimuli for all tasks were presented in pseudo-random order.

### **5.3 Tasks**

#### **5.3.1 Perception pre- and post-test**

The perception pre- and post-test task (administered to both P001 and P002) was an AX

discrimination task, in which participants listened to 54 real word stimuli pairs, presented using *TP* (Rauber et al. 2012/2016), and chose from two response options for each stimulus pair: Same or Different. 36 of the stimuli pairs contained Spanish words with /u/. Of these, 18 Spanish stimuli were preceded or followed by a different, but perceptually similar, stimulus in English, containing /u/, /ʊ/ or /ʌ/<sup>3</sup>. The other 18 were immediately preceded or followed by an identical Spanish stimulus, produced by a different native speaker. The remaining 18 stimuli pairs were distracters, containing other vowels. Of these, nine Spanish stimuli were preceded or followed by a different, but perceptually similar, English stimulus. The other nine were preceded or followed by an identical Spanish stimulus, produced by a different native speaker.

The 36 target stimuli were evenly distributed across three places of articulation for the preceding consonant, and two stress types (labial, coronal and dorsal, unstressed and stressed, six stimuli for each combination of these two variables). Stimuli for this task were nouns, verbs and adjectives, two to three syllables in length. Both unstressed and stressed /u/ appeared in the first or second syllable. Participants were allowed to listen to each stimulus three times before choosing a response.

### 5.3.2 Production pre- and post-test

The production pre- and post-test task (administered to both P001 and P002) was an elicited production task, in which participants read out loud 60 real word Spanish stimuli in the carrier phrase *Digo\_\_\_\_\_ otra vez* ('I say\_\_\_\_\_ again'). 36 of the stimuli contained /u/, and the other 24 were distracters, containing other vowels. At the beginning of both test times, five practice items were administered, to familiarize participants with the task.

The 36 target stimuli were evenly distributed across three places of articulation for the preceding consonant, and two stress types (labial, coronal and dorsal, unstressed and stressed, six stimuli for each combination of these two variables). Stimuli for this task were nouns, verbs and adjectives, three syllables in length. Unstressed /u/ occurred in the first (pretonic) syllable, while stressed /u/ always appeared in the second (medial) syllable. Recording of both participants for this task was achieved using a Zoom H4n Pro recorder, with a sample rate of 44 100 Hz.

### 5.3.3 Perception training

The perception training task (the first session of phonetic training, administered to P001 only), was a forced-choice identification/L1 perceptual assimilation task. 96 real word stimuli were presented auditorily, using *TP* (Rauber et al. 2012/2016). P001 chose one of 11 response options for each stimulus: *paso, peso, piso, poso, puso, had, hayed, head, who'd, hood, hut* (similar to Hagiwara 2006, Cobb and Simonet 2015). Of the 96 stimuli, there were 36 Spanish stimuli with unstressed and stressed /u/ after consonants with different places of articulation, 36 English stimuli with unstressed and stressed /u/, /ʊ/, /ʌ/

---

<sup>3</sup> English /ʌ/ was included, because the letter <u> can correspond to this vowel in English (ex. *buses, cuttung). It was hypothesized that the influence of orthography, especially in the production pre- and post-test reading tasks, may result in low-proficiency participants producing Spanish /u/ as English /ʌ/.*

after consonants with different places of articulation (12 with each vowel, many the same as in the AX discrimination task), and 24 distracters (12 English and 12 Spanish, many the same as in the AX discrimination task).

The 36 target Spanish stimuli and the 36 target English stimuli were evenly distributed across three places of articulation for the preceding consonant, and two stress types (labial, coronal, and dorsal, unstressed and stressed, six stimuli for each combination of these two variables). Stimuli for this task were nouns, verbs and adjectives, three syllables in length. Unstressed /u/ occurred in the first (pretonic) syllable, while stressed /u/ always appeared in the second (medial) syllable. Participants were allowed to listen to each stimulus three times before choosing a response.

Immediate feedback after each trial was presented in the following way: If an answer was correct, a green checkmark appeared and P001 was allowed to move to the next stimulus. If an answer was incorrect, a message in red appeared: “Incorrect answer! Click on Replay to listen again,” the correct response was shown, and P001 had to replay the stimulus and click on it before proceeding to the next trial. Cumulative feedback was presented at the end of the session, through a message in *TP* displaying total number of stimuli completed, total time elapsed, number of correct answers and number of errors.

### 5.3.4 Production training

The production training task (the second session of phonetic training, administered to P001 only), was an elicited imitation task. 50 real word Spanish stimuli were spoken, in the following context: (1) The experimenter spoke a Spanish sentence containing a stimulus (ex. *Luis dice curioso*). (2) The experimenter asked: *¿Qué dice X?* (“What does X say?”) (3) P001 answered the question, repeating the stimulus previously heard (ex. *Luis dice curioso*). 36 of the stimuli contained /u/, and the other 14 were distracters, containing other vowels. At both testing times, three practice items were administered before the target stimuli and distracters, in order to familiarize participants with the task.

The 36 target Spanish stimuli and the 36 target English stimuli were evenly distributed across three places of articulation for the preceding consonant, and two stress types (labial, coronal, and dorsal, unstressed and stressed, six stimuli for each combination of these two variables). Stimuli for this task were nouns, verbs and adjectives, three syllables in length. Unstressed /u/ occurred in the first (pretonic) syllable, while stressed /u/ always appeared in the second (medial) syllable.

Individual (trial-by-trial and cumulative) feedback was provided by the experimenter, through positive encouragement when correct productions of /u/ occurred and encouraging the participant to repeat (once) if a production error was detected (based on auditory judgment). Similar to the production pre-test and post-test tasks, recording for this task was done using a Zoom H4n Pro recorder, with a sample rate of 44 100 Hz.

## 5.4 Testing protocol

Testing took place at the participants’ homes, in a quiet room, over a period of approximately two weeks. The order of tasks is illustrated in the table below (Table 1).

**Table 1.** Testing protocol

Session	Session 1	Session 2	Session 3	Session 4
Tasks	Informed consent form, Background questionnaire, Pre-test: AX discrimination task, Elicited production task	Identification training task (five days after Session 1)	Elicited production training task (four days after Session 2)	Post-test: AX discrimination task, Elicited production task (four days after Session 3, approximately two weeks after Session 1)
Participants	P001, P002	P001	P001	P001, P002
Duration	45 minutes per participant	30 minutes	30 minutes	30 minutes per participant

## 6. Data analysis

The following variables were used to analyze data collected in this study.

For perception (pre- to post-test improvements): Mean accuracy, total accuracy (for different, Spanish-English, stimuli pairs) and total accuracy (for same, Spanish-Spanish, stimuli pairs). All of these variables were measured using the correct/incorrect pre- and post-test response scores in *TP* (Rauber et al. 2012/2016). They were compared between subjects (P001 and P002) and within subjects (pre-test versus post-test).

For production (pre- to post-test improvements): Duration, F1 and F2. Using Praat (Boersma and Weenink 1992/2015), all three parameters were measured for unstressed and stressed /u/, after a labial, coronal, and dorsal consonant. They were compared between subjects (P001 and P002) and within subjects (pre-test versus post-test).

Mean and standard deviation (rounded to two decimal places) were calculated for all of the ratio variables used to measure perception and production improvements.<sup>4</sup>

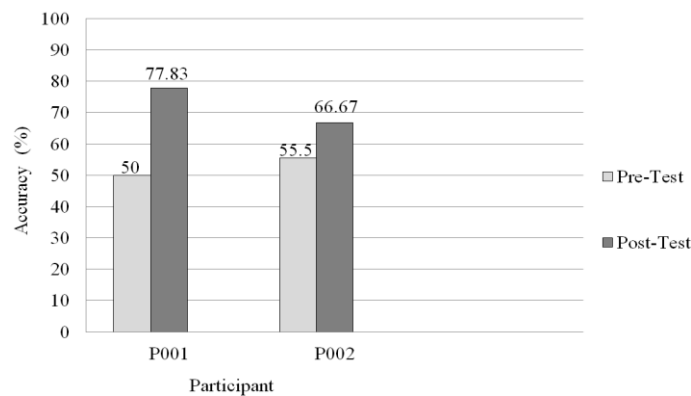
## 7. Results

<sup>4</sup> Because of the extremely small sample size for this pilot study, inferential statistical tests (ex. t-tests, ANOVAs) were not conducted. More participants are needed in order for this statistical testing to be conducted meaningfully. Target values for the production parameters were based on native Spanish speaker mean values in the *University of Toronto Romance Phonetics Database* (Colantoni and Steele 2004). They were: Duration: 71 ms (unstressed) and 74 ms (stressed), F1: 360 Hz (unstressed) and 429 Hz (stressed), and F2: 1039 Hz (unstressed) and 1490 Hz (stressed).

## 7.1 Perception

### 7.1.1 Mean accuracy

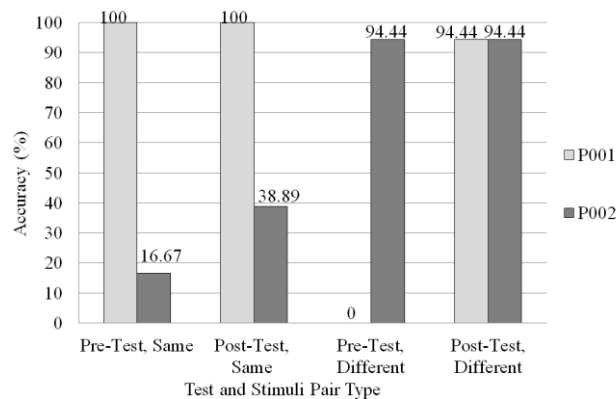
P001's mean accuracy (mean of each preceding consonant and stress combination combined, for the 36 target stimuli pairs) was found to increase considerably from pre- to post-test, from 3/6, or 50% (SD: 0), to 4.67/6, or 77.83% (SD: 0.52/6, or 8.67%). However, P002's mean accuracy also increased (to a lesser extent), from 3.33/6, or 55.5% (SD: 0.52/6, or 8.67%) to 4/6, or 66.67% (SD: 0.63/6, or 10.5%) (Figure 1).



**Figure 1.** Mean perception accuracy, by participant and test.

### 7.1.2 Total accuracy (different and same stimuli pairs)

For the 18 Spanish-English (different) target stimuli pairs: P001's total accuracy increased considerably from pre- to post-test, from zero to near ceiling (17/18, or 94.44%). P002's total accuracy was very high (17/18, or 94.44%) at pre- and post-test. For the 18 Spanish-Spanish (same) target stimuli pairs: P001's total accuracy was perfect at pre- and post-test. P002's total accuracy increased considerably from pre- to post-test, from 3/18 (16.67%) to 7/18 (38.89%), although remaining below chance level (Figure 2).



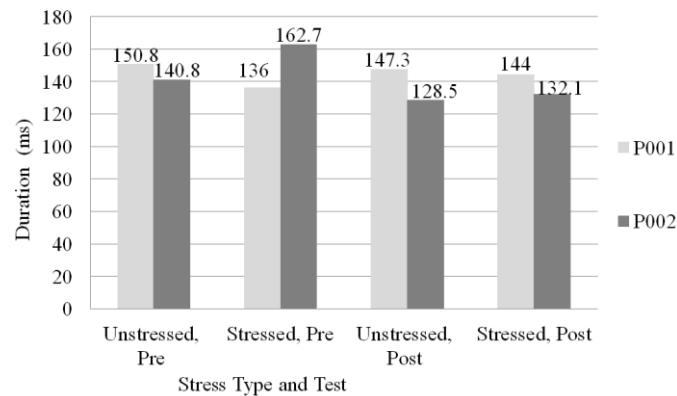
**Figure 2.** Total perception accuracy, by participant, test, and stimuli pair type.



## 7.2 Production

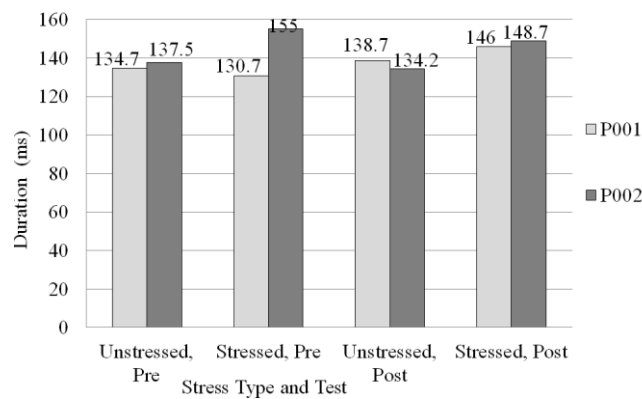
### 7.2.1 Duration

After labial consonants, for unstressed and stressed /u/, P001's mean duration changed little (unstressed: pre-test: 150.8 ms, SD: 15.8 ms; post-test: 147.3 ms, SD: 17.7 ms; stressed: pre-test: 136 ms, SD: 11.4 ms; post-test: 144 ms, SD: 22.4 ms). P002's mean duration decreased (moved toward target values) for unstressed and stressed /u/ (unstressed: pre-test: 140.8 ms, SD: 21.9 ms; post-test: 128.5 ms, SD: 13.9 ms; stressed: pre-test: 162.7 ms, SD: 25.9 ms; post-test: 132.1 ms, SD: 24.6 ms) (Figure 3).



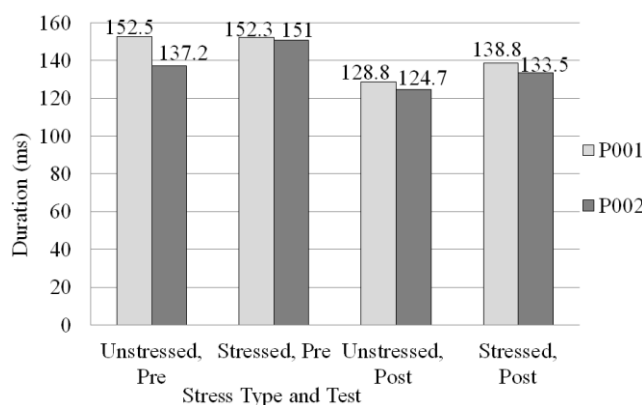
**Figure 3.** Mean duration of /u/ after labial consonants, by participant, stress, and test.

After coronal consonants, P001's mean duration changed little for unstressed /u/, but increased (moved away from target values) for stressed /u/ (unstressed: pre-test: 134.7 ms, SD: 12.6 ms; post-test: 138.7 ms, SD: 25.7 ms; stressed: pre-test: 130.7 ms, SD: 10.5 ms; post-test: 146 ms, SD: 27.4 ms). P002's mean duration showed little change for unstressed and stressed /u/ (unstressed: pre-test: 137.5 ms, SD: 25.2 ms; post-test: 134.2 ms, SD: 19.6 ms; stressed: pre-test: 155 ms, SD: 17.5 ms; post-test: 148.7 ms, SD: 22.4 ms) (Figure 4).



**Figure 4.** Mean duration of /u/ after coronal consonants, by participant, stress, and test.

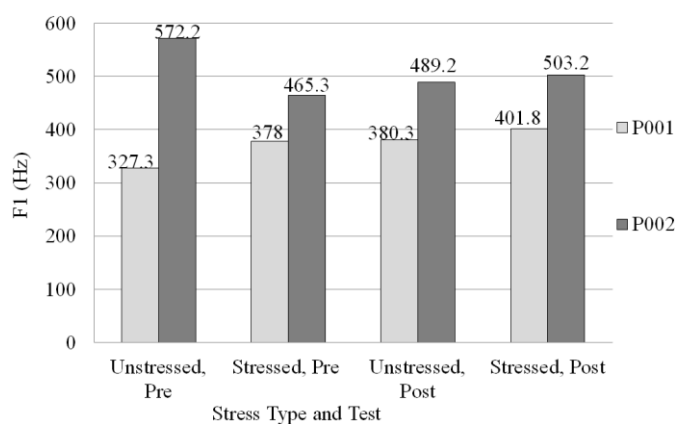
After dorsal consonants, P001's mean duration decreased (moved toward target values) for both unstressed and stressed /u/ (unstressed: pre-test: 152.5 ms, SD: 27.2 ms; post-test: 128.8 ms, SD: 25.3 ms; stressed: pre-test: 152.3 ms, SD: 18.9 ms; post-test: 138.8 ms, SD: 22.2 ms). P002's mean duration also decreased (moved toward target values) for unstressed and stressed /u/ (unstressed: pre-test: 137.2 ms, SD: 30.3 ms; post-test: 124.7 ms, SD: 10.0 ms; stressed: pre-test: 151 ms, SD: 21.8 ms; post-test: 133.5 ms, SD: 19.6 ms) (Figure 5).



**Figure 5.** Mean duration of /u/ after dorsal consonants, by participant, stress, and test.

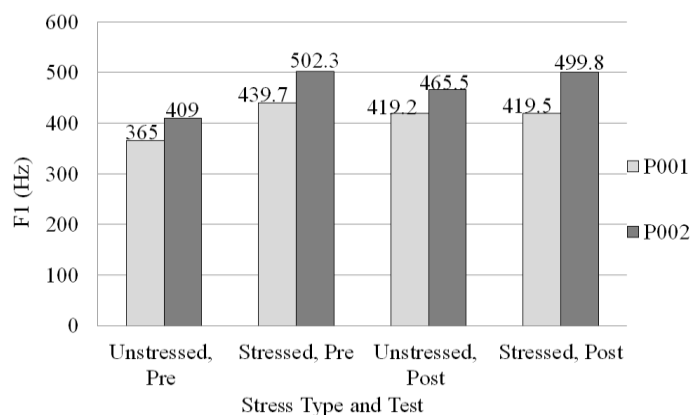
### 7.2.2 F1

After labial consonants, P001's mean F1 increased (moved toward target values) for unstressed and stressed /u/ (unstressed: pre-test: 327.3 Hz, SD: 44 Hz; post-test: 380.3 Hz, SD: 32.5 Hz; stressed: pre-test: 378 Hz, SD: 103 Hz; post-test: 401.8 Hz, SD: 93.4 Hz). P002's mean F1 decreased (moved toward target values) for unstressed /u/, but increased (moved away from target values) for stressed /u/ (unstressed: pre-test: 572.2 Hz, SD: 183.6 Hz, post-test: 489.2 Hz, SD: 82.5 Hz; stressed: pre-test: 465.3 Hz, SD: 64.4 Hz; post-test: 503.2 Hz; SD: 92.3 Hz) (Figure 6).



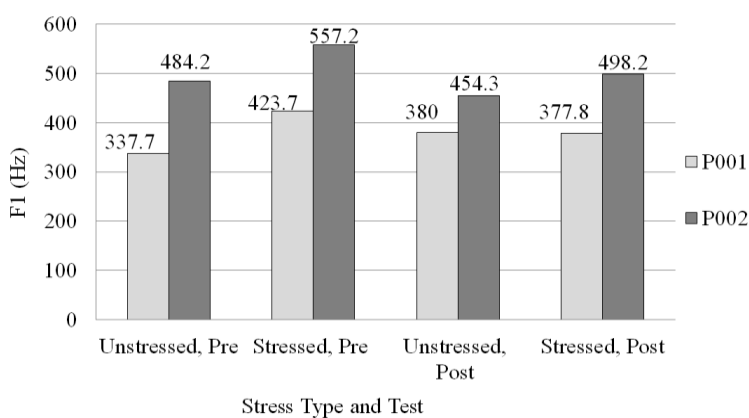
**Figure 6.** Mean F1 of /u/ after labial consonants, by participant, stress, and test.

After coronal consonants, P001's mean F1 increased (moved away from target values) for unstressed /u/, and decreased (moved from slightly above to slightly below target values) for stressed /u/ (unstressed: pre-test: 365 Hz, SD: 85.2 Hz; post-test: 419.2 Hz, SD: 19.1 Hz; stressed: pre-test: 439.7 Hz, SD: 131 Hz; post-test: 419.5 Hz, SD: 83.1 Hz). P002's mean F1 increased (moved away from the target) for unstressed /u/ and remained almost unchanged for stressed /u/ (unstressed: pre-test: 409 Hz, SD: 68.8 Hz; post-test: 465.5 Hz, SD: 43.9 Hz; stressed: pre-test: 502.3 Hz, SD: 96.4 Hz; post-test: 499.8 Hz, SD: 42 Hz) (Figure 7).



**Figure 7.** Mean F1 of /u/ after coronal consonants, by participant, stress, and test.

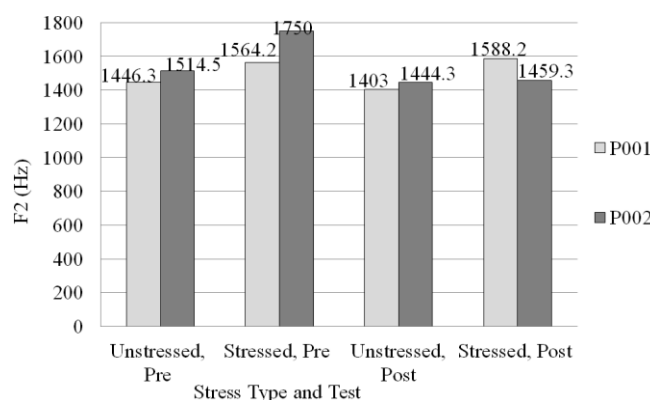
After dorsal consonants, P001's mean F1 increased (moved from below to above target values) for unstressed /u/, and decreased (moved away from target values) for stressed /u/ (unstressed: pre-test: 337.7 Hz, SD: 76.6 Hz; post-test: 380 Hz, SD: 75.8 Hz; stressed: pre-test: 423.7 Hz, SD: 64.9 Hz; post-test: 377.8 Hz, SD: 70.1 Hz). P002's mean F1 decreased (moved toward target values) for both unstressed and stressed /u/ (unstressed: pre-test: 484.2 Hz, SD: 135.5 Hz; post-test: 454.3 Hz, SD: 67.8 Hz; stressed: pre-test: 557.2 Hz; SD: 105.3 Hz; post-test: 498.2 Hz, SD: 76.1 Hz) (Figure 8).



**Figure 8.** Mean F1 of /u/ after dorsal consonants, by participant, stress, and test.

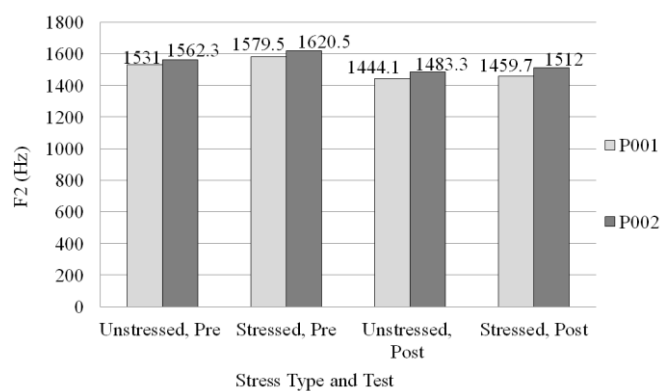
### 7.2.3 F2

After labial consonants, P001's mean F2 decreased (moved toward target values) for unstressed /u/, and increased (moved away from target values) for stressed /u/ (unstressed: pre-test: 1446.3 Hz, SD: 253.4 Hz; post-test: 1403 Hz, SD: 298.5 Hz; stressed: pre-test: 1564.2 Hz, SD: 335.5 Hz; post-test: 1588.2 Hz, SD: 227.9 Hz). P002's mean F2 decreased (moved toward target values) for unstressed and stressed /u/ (unstressed: pre-test: 1514.5 Hz, SD: 334.9 Hz; post-test: 1444.3 Hz, SD: 146.6 Hz; stressed: pre-test: 1750 Hz, SD: 349.3 Hz; post-test: 1459.3 Hz; SD: 354.2 Hz) (Figure 9).



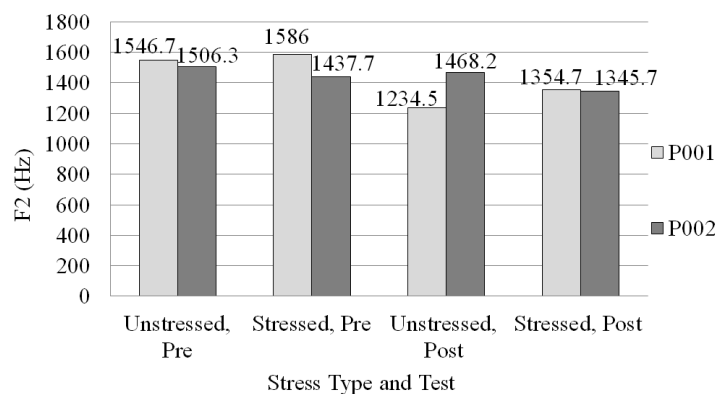
**Figure 9.** Mean F2 of /u/ after labial consonants, by participant, stress, and test.

After coronal consonants, P001's mean F2 decreased (moved toward target values) for unstressed and stressed /u/ (unstressed: pre-test: 1531 Hz, SD: 156.1 Hz; post-test: 1444.1 Hz, SD: 209.3 Hz; stressed: pre-test: 1579.5 Hz, SD: 161.5 Hz; post-test: 1459.7 Hz, SD: 291.9 Hz). P002's mean F2 also decreased (moved toward target values) for unstressed and stressed /u/ (unstressed: pre-test: 1562.3 Hz, SD: 286.5 Hz; post-test: 1483.3 Hz, SD: 307.2 Hz; stressed: pre-test: 1620.5 Hz, SD: 390.2 Hz; post-test: 1512 Hz, SD: 270.5 Hz) (Figure 10).



**Figure 10.** Mean F2 of /u/ after coronal consonants, by participant, stress, and test.

After dorsal consonants, P001's mean F2 decreased (moved toward target values) for unstressed /u/, and decreased (moved away from target values) for stressed /u/ (unstressed: pre-test: 1546.7 Hz; SD: 379.7 Hz; post-test: 1234.5 Hz, SD: 110.5 Hz; stressed: pre-test: 1586 Hz, SD: 87.9 Hz; post-test: 1354.7 Hz, SD: 84.9 Hz). P002's mean F2 decreased (moved toward target values) for unstressed /u/ and decreased (moved away from target values) for stressed /u/ (unstressed: pre-test: 1506.3 Hz; SD: 317.2 Hz; post-test: 1468.2 Hz, SD: 168.3 Hz; stressed: pre-test: 1437.7 Hz, SD: 208.2 Hz; post-test: 1345.7 Hz, SD: 291.4 Hz) (Figure 11).



**Figure 11.** Mean F2 of /u/ after dorsal consonants, by participant, stress, and test.

## 8. Discussion

### 8.1 Summary of results

Results for P001 in this study provide some evidence that phonetic training may be effective in improving perception and production of Spanish /u/ in low-proficiency L2 learners with L1 Canadian English. Perception (mean accuracy and total accuracy for different, Spanish /u/- English /u/, /ʊ/, or /ʌ/ stimuli pairs) may improve with training. In addition, production may improve with training on the parameters of: Duration (for unstressed and stressed /u/ after dorsal consonants), F1 (for unstressed and stressed /u/ after labial consonants), and F2 (unstressed /u/ after labial, coronal, and dorsal consonants, and stressed /u/ after coronal consonants). However, native speaker target values for these production parameters may be rarely reached or approximated with phonetic training alone.

Unexpectedly, results for P002 in this study also provide evidence that perception and production of Spanish /u/ in low-proficiency L2 learners with L1 Canadian English may be achievable without phonetic training. Specifically, perception (mean accuracy and total accuracy for same, Spanish /u/-Spanish /u/ stimuli pairs) may improve without training. And production may improve without training on: Duration (unstressed and stressed /u/ after labial and dorsal consonants), F1 (unstressed /u/ after labial and dorsal consonants, and stressed /u/ after dorsal consonants), and F2 (unstressed /u/ after labial, coronal, and dorsal consonants, and stressed /u/ after labial and coronal consonants). One

possible reason for this unexpected finding is that (as discovered in the background questionnaire data, after the experiment had been run), P002 was undergoing an informal Spanish learning program at the time of the experiment (Berlitz 2008). This program involved pronunciation (listening, speaking, and repeating) of words and phrases, some of which contained unstressed and stressed /u/. However, native speaker target values for the production parameters of duration, F1 and F2 may be rarely reached or approximated with informal L2 learning programs alone.

## 8.2 Future work

Several follow-up steps are needed for this study. First, it should be expanded to the other Spanish vowels: /a/, /e/, /i/, and /o/, to determine if there are differences in Spanish vowel improvement with training as a function of target vowel. In addition, the study should be conducted on more L1 Canadian English participants, to evaluate the hypotheses more thoroughly and have stronger evidence for its conclusions. Furthermore, this study needs to be conducted on participants from other L1 groups (ex. Canadian/European French, European/Brazilian Portuguese, Catalan, Italian, German, etc.) to investigate if there are differences in Spanish vowel improvement with training as a function of participants' L1.

## 9. Conclusion

This study has provided evidence that: To a similar extent, both phonetic training and informal L2 pronunciation learning programs may be effective in improving perception and production of Spanish /u/ in low-proficiency L2 learners with L1 Canadian English. Both types of L2 pronunciation learning may improve perception, on the parameter of mean accuracy. Phonetic training may be more effective in improving total perception accuracy for different (Spanish /u/-English /u/, /ʊ/, or /ʌ/) stimuli pairs, while informal L2 pronunciation learning programs may be more effective in improving total perception accuracy for same (Spanish /u/-Spanish /u/) stimuli pairs. Regarding production, both types of L2 pronunciation learning may improve duration, F1 and F2. An important implication of this study for L2 Spanish learning is that L2 Spanish phonology may be developed and improved in multiple ways.

## References

- Berlitz. 2008. *Spanish phrase book & CD*. Singapore: Berlitz.
- Boersma, Paul, and David Weenink. 1992. Praat: doing phonetics by computer [Computer program]. Version 5.4.04, retrieved 12 January 2015 from <http://www.praat.org/>.
- Bradlow, Ann R., David B. Pisoni, Reiko Akahane-Yamada, and Yoh'ichi Tohkura. 1997. Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America* 101(4): 2299–2310.
- Bradlow, Ann R., Reiko Akahane-Yamada, David B. Pisoni, and Yoh'ichi Tohkura. 1999. Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics* 61(5): 977–985.
- Cobb, Katherine, and Miquel Simonet. 2015. Adult second language learning of Spanish vowels. *Hispania* 98(1): 47-60.

- Colantoni, Laura, and Jeffrey Steele. 2004. *The University of Toronto Romance Phonetics Database*. <http://rpd.chass.utoronto.ca/>
- Flege, James E. 1995. Second language speech learning: Theory, findings and problems. In *Speech perception and linguistic experience: Issues in cross-language research*, ed. Winifred Strange, 233–277. Timonium, MD: York Press.
- Hagiwara, Robert E. 2006. Vowel production in Winnipeg. *The Canadian Journal of Linguistics/La revue canadienne de linguistique* 51(2/3): 127-141.
- Herd, Wendy, Allard Jongman, and Joan Sereno. 2013. Perceptual and production training of intervocalic /d, r, r/ in American English learners of Spanish. *The Journal of the Acoustical Society of America* 133(6): 4247–4254.
- Hualde, José I. 2014. *Los sonidos del español*. Cambridge, UK: Cambridge University Press.
- Kartushina, Natalia, Alexis Hervais-Adelman, Ulrich H. Frauenfelder, and Narly Golestani. 2015. The effect of phonetic production training with visual feedback on the perception and production of foreign speech sounds. *The Journal of the Acoustical Society of America* 138(2): 817–832.
- Logan, John S., and John S. Pruitt. 1995. Methodological issues in training listeners to perceive non-native phonemes. In *Speech perception and linguistic experience: Issues in cross-language research*, ed. Winifred Strange, 233–277. Timonium, MD: York Press.
- McCandless, Andrew. 2020. Phonetic training of L2 Spanish rhotics. In *Actes du congrès annuel de l'Association canadienne de linguistique 2020/Proceedings of the 2020 Annual Conference of the Canadian Linguistic Association*, ed. Angélica Hernández and M. Emma Butterworth. Association canadienne de linguistique/Canadian Linguistic Association.
- Menke, Mandy R., and Timothy L. Face. 2010. Second language Spanish vowel production: An acoustic analysis. *Studies in Hispanic and Lusophone Linguistics* 3(1): 181-214.
- Morrison, Geoffrey S. 2003. Perception and production of Spanish vowels by English speakers. In *Proceedings of the 15<sup>th</sup> International Congress of Phonetic Sciences: Barcelona 2003*, ed. Maria-Josep Solé, Daniel Recasens, and Joaquin Romero, 1533-1536. Adelaide, South Australia: Causal Productions.
- Rauber, Andréia S., Anabela Rato, Denise C. Kluge, Giane R. dos Santos, and Marcos Figueiredo. 2012. *TP* [Application software]. Version 3.1, retrieved 31 October 2016 from [http://www.worken.com/br/tp\\_regfree.php](http://www.worken.com/br/tp_regfree.php)
- Sakai, Mari. 2016. (Dis)Connecting perception and production: Training adult native speakers of Spanish on the English /i/-/ɪ/ distinction. Doctoral dissertation, Georgetown University.
- Sakai, Mari, and Colleen Moorman. 2017. Can perception training improve the production of second language phonemes? A meta-analytic review of 25 years of perception training research. *Applied Psycholinguistics* 39(1): 1–38.
- Schwegler, Armin, Juergen Kempff, and Ana Ameal-Guerra. 2010. *Fonética y fonología españolas*. 4th ed. Hoboken, NJ: John Wiley & Sons.