Masthead Logo

University of Richmond
UR Scholarship Repository

Latin American, Latino and Iberian Studies Faculty Publications

Latin American, Latino and Iberian Studies

8-8-2014

Phonetics Instruction Improves Learners' Perception of L2 Sounds

Elizabeth M. Kissling
University of Richmond, ekisslin@richmond.edu

This is a pre-publication author manuscript of the final, published article.

Phonetics instruction improves learners' perception of L2 sounds

Elizabeth M. Kissling

James Madison University, United States

Abstract

Explicit phonetics instruction can help second language (L2) learners to moderately improve their pronunciation, but less is known about how the instruction affects learners' perception, even though there is evidence that perception and pronunciation are related. This study provided phonetics instruction to students (n=46) studying Spanish as a foreign language and measured the resulting change in their perception of eight target phones as compared with a control group (n=41). Perception was assessed with discrimination tests immediately following instruction and three weeks later. Results indicated that the instruction conferred a small advantage in the delayed posttest and that course level was not a significant factor, suggesting that phonetics instruction was effective for attuning the perception of learners at multiple stages of development.

Keywords

L2 perception, pronunciation instruction, phonetics instruction

I Introduction

Second language (L2) learners are often concerned about the “foreign” quality of their accent, and many express a desire to sound more like a “native speaker” of the L2 (Derwing & Rossiter, 2003; Timmis, 2002). Their concerns are not unfounded, because some listeners do indeed judge “foreign” accents negatively, even when those accents are perfectly intelligible (Derwing, 2003). Traditionally, pronunciation instruction aimed at reducing the differences between learners’ speech and the target speech employs explicit lessons in L2 phonetics and practice, with moderately successful results (Pardo, 2004). While instructors and learners have not been as concerned with *perception* skills per se, many theoretical accounts of L2 phonological acquisition posit that perception and production are intimately related (Colantoni & Steele, 2008). It might be necessary first to attune learners’ perception before one can make significant and long-lasting improvement in their pronunciation. Yet the training and testing paradigms prevalent in the perception literature bear little resemblance to the types of teaching and learning that typically occur in language classrooms. The present study was situated in a Spanish as a foreign language (FL) classroom environment and explored the effect of explicit phonetics instruction on learners’ perception of the FL under these typical classroom conditions.

I Second language speech perception

James Flege’s Speech Learning Model (SLM) claims that target-like perception of the L2 is a necessary precursor to target-like production. This claim is based on a wealth of empirical research linking perception to production in many different language pairs (Akerberg, 2005; Flege, 1988; Hwang, 2011; Munro, 2008; Newman, 1996; Rochet, 1995). The SLM posits that to fully acquire L2 sounds learners must first discern subtle phonetic differences between L2 sounds and analogous sounds in the L1 as they occur in a variety of phonetic environments, at which point learners can create new phonetic categories for L2 sounds and then finally produce the L2 sounds in target-like ways (Flege, 1988; Flege, 1995). Though there is some evidence inconsistent with the predictions of the SLM, namely that learners may produce contrasts they cannot perceive (Goto, 1971) and may demonstrate different acquisitional patterns in the perceptive and productive modes (de Jong, Hao, & Park, 2009; Eckman, Iverson, Fox, Jacewicz, & Lee, 2011), still the SLM is the model most commonly used in research on adult, instructed L2 speech, and it motivates the methodology of the present study. Other perception-based models also assert that perception underlies production (Munro, 2008). They characterize the nature of the link in different ways, but stated in general terms, some attentional detection (perception) is thought to be necessary for further processing and storage in long-term memory which, in turn, contains the phonological knowledge used for production.

Perception is not a monolithic construct but rather involves multiple levels of processing. Wode (1981) describes two processing modes, a relatively independent "continuous" mode that is used for non-speech sounds, among other things, and a "categorical" mode dependent on the phonemic categories built up over time with exposure to one's first language(s). Infants process phonetically and acoustically (in a "continuous" mode) and gradually transition to more phonemic ("categorical") processing (Werker & Tees, 1984), with the ambient language(s) serving as a magnet that warps perception around L1-informed phonemic categories. Adults maintain the ability to process speech using either perceptual mode, but under normal conditions the phonemic ("categorical") network tends to inhibit the acoustical ("continuous") network (Dehaene-Lambertz et al., 2005). The problem in L2 speech perception is that adults tend to use categorical/phonemic processing in the L2 even though their phonological knowledge of the L2 is imperfect. They use the same (ill-fitting) categories they have developed for their L1(s) when processing their L2(s). Models of L2 phonological acquisition such as the SLM imply that learners must switch to continuous/phonetic processing in order to discern subtle phonetic differences between L2 sounds and analogous L1 sounds. In other words, acquiring an L2 sound system begins first with detecting differences between native and non-native sounds and then developing the appropriate selective perceptual routines to "hear" them reliably (Strange & Shafer, 2008).

The two most commonly used experimental techniques for measuring L2 speech perception are phoneme discrimination and identification tasks (Boomershine, Hall, Hume, & Johnson, 2008; Logan & Pruit, 1995). Discrimination tasks present learners with two or three tokens of stimuli and learners determine whether they are identical. Identification tasks present learners with a minimal pair and learners identify the phone they hear. Both tasks can hypothetically tap any level of perception, depending on the task conditions. It is difficult, perhaps impossible, to completely separate the levels of processing in practice, because they occur automatically, rapidly, and in parallel fashion (Boomershine, Hall, Hume, & Johnson, 2008), but the timing of the presentation of stimuli is crucial to determining which level of processing participants will most likely use (Werker & Logan, 1985).

II Phonetics instruction and L2 speech production and perception

The majority of evidence as to whether learners can be taught to better perceive L2 sounds comes from experiments that provide participants with intensive exposure under relatively implicit conditions (see (Logan & Pruit, 1995). They find that learners can improve their ability to perceive L2 contrasts that are not relevant in their L1 (Golestani & Zatorre, 2009; Jamieson & Morosan, 1989), such as English /ɹ/-/l/ for Japanese learners (Bradlow, Yamada, Pisoni, & Tohkura, 1999; B. D. McCandless, Fiez, Protopapas, Conway, & McClelland, 2002). This research tends to focus solely on perception, though perception and production instruction might be mutually facilitative (Lacabex, García Lecumberri & Cooke, 2009). However, the relatively intense and implicit perceptual training used in these studies is unlike the less intensive, more explicit instruction traditionally utilized in FL classes. On the other hand, classroom-based research has focused primarily on production rather than perception. Typically, the core component of pronunciation instruction in the classroom is the explicit teaching of L2 phonetics, often in contrast with the L1 sound system. Though the format and duration of the instruction varies, instruction might include pronunciation practice, phoneme discrimination practice, and identification exercises, with feedback. Research suggests that while general language instruction is not related to global foreign accent (Piske, MacKay, & Flege, 2001), pronunciation instruction has a significant effect on L2 production accuracy in several FL contexts (Pardo, 2004), including English (Pennington, 1992), French (Walz, 1980), German (McCandless & Winitz, 1986; Moyer, 1999), and Spanish (Lord, 2005). Some adult L2 learners, particularly "fossilized" learners (Derwing, Munro, & Wiebe, 1997) might not achieve native-like pronunciation without the help of instruction (Fullana, 2006).

It is still an empirical question whether explicit or implicit instruction is more beneficial for L2 perception in the FL setting. The few studies that explore perception changes after phonetics instruction have reported some positive evidence (Champagne-Muzar, Scheneiderman, & Bourdages, 1996; Matthews, 1997) and some negative (Yule, Hoffman, & Damico, 1987). Explicit knowledge about L2 phonetics may serve as an attention-orienting device to help learners change their L1-informed automatic processing routines (Guion-Anderson, 2013). Even with a relatively early start and relatively high L2 proficiency attained, learners do not necessarily acquire native-like phonetic categories implicitly through consistent and extensive contact with the L2, so instruction may be necessary to attune their L2 perception (Archila-Suerte, Zevin, Bunta, & Hernandez, 2011). Yet individual differences such as aptitude for speech sounds (Munro, 2008) and processing preferences (Markham, 1997; Thompson, 1991) can impact the ability of learners to attune their perception to the L2, and these individual differences might predict their ability to learn from particular types of instruction. Another complication is that the effect of instruction will depend on the particular target phones. Some require intensive, consistent, and repeated training to improve learners' perception (Polka, 1992), whereas others are simply easier to perceive because of their inherent perceptual saliency (Tees & Werker, 1984).

Several investigations of the effect of phonetics instruction in the context of the current study, adult Spanish FL, report generally positive results. These studies, detailed in Table 1, find that learners improve their production of some difficult Spanish phones after receiving phonetics instruction, but none addresses perception. One Spanish FL study to date has measured changes in perception after phonetics instruction and reports that learners are better able to detect non-target-like realizations of some Spanish phones, but not of the stops /p, t, k/ (Ausín & Sutton, 2010). Clearly much more research is needed to better understand the impact that phonetics instruction might have on Spanish FL learners' perception. The present study aimed to do just that. The research question was: Does instruction in L2 phonetics improve learners' ability to perceive the acoustic differences between L2 phones and their analogous L1 phones? Rather than investigate learners' perception of L2 contrasts, in this study a bilingual discrimination task paired Spanish phones with analogous English phones, because the Speech Learning Model predicts that a necessary first stage in L2 speech acquisition is being able to discern the subtle phonetic differences between these pairs.

Table 1. Empirical studies assessing Spanish FL learners' pronunciation of segments after phonetics instruction.

Author	Participants' Level	N	Targets	Instructional Time	Additional* Activities	Control condition	Results
Castino (1996)	6-8th semester, phonetics course	40	[β ð ɣ ɾ r x]	full semester course	Extensive drills. Transcription.	n/a	Significant improvement for all phones. Spontaneous speech more accurate than dialogue reading.
Elliott (1995, 1997)	3rd semester	66	[a e i o u p t k b d g β ð ɣ ɾ r ɲ s z w]	10-15 min/class, 21 periods	Repetition, jazz chants, rhymes, and tongue twisters. Focus on corrective feedback.	Section of same course taught by same instructor. Little corrective feedback.	Instruction significant predictor of aggregate post-test scores on all tasks but spontaneous speech and of post-test scores of liquids and stops only.
González –Bueno (1997)	4th semester. OPI determined all were intermediate	60	/ p t k b d g /	5-10 min/class, 3 times/week, full semester	Perceptual discrimination. Transcription. Sentence repetition and expanded dialogues.	Section of same course taught by same instructor.	Significant reduction of VOT of /p/ and /g/ only in spontaneous speech.
Lord (2005)	phonetics course	17	[p t k β ð ɣ r], dipthongs	full semester course	Acoustic analysis of spectrograms.	n/a	Significant improvement in producing [β ð ɣ r] and diphthongs when reading a passage. No reduction of VOT for / p t k /.
Lord (2010)	6-8th semester, study abroad	8	[b d g β ð ɣ]	8 weeks	Study abroad immersion following semester-long phonetics course.	Same program, but no prior phonetics instruction.	Experimental group improved 17% more than control (not tested for significance) in word list reading.

* Main instruction included explanations of Spanish/English contrasts, place and manner of articulation, and grapheme-phoneme correspondences.

II Method

Eight Spanish phones ([p, t, k, β̞, ð̞, ɣ̞, ɾ, r]) were chosen as the linguistic targets in this study because they are the segments most commonly investigated in Spanish FL phonetics instruction research (see Table 1) and are widely cited as being late acquired in native English speakers' pronunciation (Elliott, 1997; Face & Menke, 2010; Lord, 2005). Native speakers of English tend to aspirate /p, t, k/ in initial and stressed syllables when speaking Spanish (Hualde, 2005; Lord, 2005). They tend to produce stops for Spanish /b, d, g/ in many phonological environments that require their approximant allophones [β̞, ð̞, ɣ̞] (Hualde, 2005; Zampini, 1993), hereafter written without the diacritic [β, ð, ɣ] as in most other studies. Learners tend to produce /ɹ/ or other non-target sounds for the Spanish rhotics /ɾ/ and /r/ (Elliott, 1997; Face, 2006). Researchers have reported positive effects of instruction on learners' production of these phones, though results have been mixed as regards to which phones are actually amenable to instruction (see Table 1), and yet to date no systematic investigation of learners' *perception* of the phones has been published.

Learners (n=87) were enrolled in introductory, intermediate, or advanced Spanish FL courses at a large, public university in the southeastern United States. Seven classes participated, with five instructors. Instructors were Spanish-dominant until adolescence, were college educated, and had lived in the United States and taught Spanish for eight or more years. Learners were not tested for proficiency but will be referred to as first, second, and third year learners, in reference to the placement of their course in the curriculum. The learners were at least 18 years of age, had no Spanish exposure before age 10, and had no previous instruction in phonetics. There were 55 females and 32 males. Their average age was 22.06 (range 18-44, mode = 19). Their average age at outset of learning was 15.66 years (range 11–40, mode = 13).

The study was a pre-, post-, delayed posttest design implemented over six weeks during class hours. Between the experimental sessions, classes in each level followed similar syllabi, which did not contain any phonetics instruction. Instructors and learners were unaware of the research questions and target phones. They were told that the study was designed to develop instructional materials for listening and speaking. They were not compensated. During the first session half the learners in each class were randomly assigned to the experimental (+PI) group and half to the control (-PI) group.

The experimental, phonetics instruction group (+PI) completed four computer-delivered, interactive modules during sessions two and three of the study, with one week between sessions. The modules focused on: 1) an introduction to articulatory phonetics; 2) the voiceless stops /p, t, k/; 3) the voiced stops /b, d, g/ and their approximant allophones [β, ð, ɣ]; and 4) the rhotics /ɾ, r/. The modules were from Tal Como Suena, an online course created by Dr. Gillian Lord at the University of Florida (http://talcomsuena.spanish.ufl.edu/). All learners began with the introduction to articulatory phonetics. The three other modules were counterbalanced for order of presentation. Each module presented the following information and activities: an explanation of grapheme-phoneme correspondences; an explanation of the place and manner of articulation, with an animated diagram of the vocal tract illustrating how each sound is produced; an explanation of differences in the articulation of analogous Spanish/English sounds and the phonological environments in which the sounds are produced in each language; and identification activities which required learners to identify Spanish and English sounds played in isolation. Each section included a brief multiple-choice comprehension test that learners had to answer accurately before proceeding. Finally, each module contained a pronunciation practice activity that directed learners to listen to and repeat after a native speaker producing Spanish phrases until they thought their pronunciation approximated the native speaker. Learners received no additional feedback. A time limit of 25 minutes per module was suggested, but the modules were self-paced, and so learners actually spent between 15 and 40 minutes on each. The instructional time per phone was thus brief but comparable to that of other FL classes in similar studies (Elliott, 1995; González-Bueno, 1997). The instruction exposed learners minimally to ten unique tokens of each target phone, three of which were contained in the pronunciation practice section.

Learners in the control group (-PI) completed self-paced, computer-delivered, interactive online modules during sessions two and three of the study, with one week between sessions. These modules provided exposure, practice, and feedback like the +PI, but without the explicit instruction in phonetics. During each module learners watched a level-appropriate video vignette (from the University of Texas, Austin series found at http://laits.utexas.edu/spe/) featuring a Spanish speaker discussing a topic related to the topics being covered in their class. The featured speakers were from various regions but none produced the target phones in non-standard ways. Learners completed a dictation of the vignette (pausing the video as often as they liked), compared their dictation with a transcript, read an English translation for meaning, commented on the speaker's accent, and repeated aloud one sentence in the video until their pronunciation was like the speaker's. On average, learners in the -PI were exposed to the same number of unique tokens of the target phones as the +PI (ten unique tokens of each target phone, three of which were contained in the pronunciation practice section). However, learners in both groups could rewind as often as they liked, and so learners likely heard and/or pronounced more than these minimal tokens during instruction. Nonetheless, the dictation exercises were fairly comparable to the phonetics instruction in that they provided equal input in terms of target phones, number of native speakers and time on task, as well as similar pronunciation practice with identical feedback conditions.

The present study was concerned with learners’ ability to discern subtle phonetic differences between L2 sounds and analogous sounds in the L1, which is the first stage of phonological acquisition according to the SLM. A discrimination task paired target-like Spanish phones with their English-accented counterparts, i.e. the non-target-like phones that learners tend to produce in specific phonological environments. The items, detailed in Table 2, presented the voiceless stops /p, t, k/ in initial position, the approximants [β, ð, ɣ] in intervocalic position, and the rhotics /ɾ, r/ in initial, medial, and final positions, all with a variety of vowels. So, for instance, an item might pair English-accented [p^{h}a] with Spanish target-like [pa]. There were five items for most pairs but only three items for [ɾ]/[ɹ] and [r]/[ɹ] because pilot testing suggested that participants would discriminate these pairs well. It was not expected that discrimination would be equally difficult across all pairs, and these predictions are discussed in the results section. A total of 36 items contrasted target phones with analogous English phones and will be discussed in the following analysis. Other test items not of interest were distracters of three types: target phones paired in XX pairs (e.g., [eðe] [eðe]) and non-target Spanish phones in XX pairs (e.g., [feɪ] [feɪ]) and AX pairs (e.g., [fe] [feɪ]). These same items, scrambled into different inter-pair and intra-pair orders, made up the pretest (one week prior to instruction), the immediate posttest (directly after instruction), and delayed posttest (three weeks after last instruction).

Table 2. Target phones in discrimination test.

Phone Pairs	Phonological Environments
[p]/[p^{h}], [t]/[t^{h}], [k]/[k^{h}]	[_a], [_e], [_i], [_o], [_u]
[β̞]/[b], [ð̞]/[d], [ɣ̞]/[g]	[a_a], [e_e], [u_o], [i_e], [o_i]
[ɾ]/[ɹ]	[o_o], [i_], [_a]
[r]/[ɹ]	[a_a], [e_], [_u]

The test stimuli consisted of sound recordings made by a male native Spanish speaker from Argentina who was trained in Spanish phonetics and had near-native proficiency in English. This speaker's dialect did not exhibit any non-standard realizations of the target phones. The recordings were made using the audio editing software SoundTrackPro at a 16-bit sampling rate of 48 KHz in a sound proof booth. A single recording was used for each of the token types, so, for instance, all tokens of [pa] were acoustically identical. While discrimination tests typically do not employ acoustically identical tokens, this method was chosen to suit the current study's linguistic targets, which were not contrastive and so could not be formed into minimal pairs. Volume, pitch, duration, and quality of the surrounding vowels were consistent between the two stimuli in the AX pairings. All these parameters sounded identical to both the researcher and the native speaker. The average duration of the stimulus exemplars was 1 second, with a 1500ms inter-stimulus interval (ISI). This long ISI should have precluded participants from using only their acoustic store to compare the stimuli, and therefore acoustically identical tokens could be used without compromising the objective of the task. These task conditions were appropriate for assessing to what degree learners would use Spanish-specific (target-like) phonetic categories while in a phonemic/categorical processing mode (Werker & Logan, 1985).

Learners completed the discrimination task while seated at individual computer stations in a language laboratory. Since pilot testing had indicated that participants tended to misinterpret the instructions *"If you hear a different sound,"* and underreport the phonetic differences they were able to discern, the instructions were revised to state *"If you hear any difference at all between the two recordings, choose 'different.'"* This choice of language in the task instructions also compelled the use of acoustically identical stimuli. Learners completed two practice items and received feedback to ensure they understood the instructions. Learners highlighted their chosen responses for the test items on a paper answer sheet. Learners' responses were assigned one point if correct and zero points if incorrect. Additionally, for each item, learners rated how confident they were in their choice, on a scale of one to four.

Several measures of learner differences were taken as well. Phonetic encoding ability was measured with the phonetic script learning subtest of the Modern Language Aptitude Test (Carroll, 1962). Phonological short-term memory was measured with two non-word repetition tasks, one in Spanish and one in English (taken from Lado, 2008). Attitude was measured with a Pronunciation Attitude Inventory (Elliott, 1995).

III Results

The +PI and -PI groups were comparable both in terms of demographic variables (sex, current age) and Spanish language experience (onset of Spanish learning, number of Spanish courses taken in high school and college, number of native speakers as teachers, time using Spanish outside class, and time abroad) as well as general language experience (formal learning, informal exposure, and proficiency in other languages). Independent samples T-tests compared the +PI and -PI groups at each course level (first, second and third year) with respect to each of these variables and found only one significant difference (all other $t < 2.27$, $p > .05$): the third year +PI learners spent more time abroad. This difference originated with three individuals (with two weeks, one month, and two months abroad, respectively), and they were kept in the analysis.

Table 3 lists learners' average scores for each of the target phones. The discrimination test was a forced-choice test with two possible responses, and thus a 50% accuracy rate represented a score that was at chance. On the pretest, learners were near chance on [k]/[k^h] (42%), [β]/[b] (48%), and [ɣ]/[g] (62%), more accurate on [t]/[t^h] (78%) and [ð]/[d] (62%), and highly accurate on [p]/[p^h] (87%), [ɾ]/[ɹ] (99%), and [r]/[ɹ] (90%). According to García Lecumberri's (1999) criteria for assessing perceptual difficulty, [k]/[k^h], [β]/[b] and [ɣ]/[g] presented a great deal of difficulty for these learners (below 65% accuracy), whereas the [t]/[t^h] and [ð]/[d] presented some difficulty for these learners (65-80% accuracy), and [p]/[p^h], [ɾ]/[ɹ], and [r]/[ɹ] presented little difficulty (80-99% accuracy). For sake of comparison, learners were highly accurate at discriminating the distractor items, both identical pairs (97%) and vowels (88% accuracy).

Most of these pretest results were expected. Learners should be able to discriminate the rhotics (English [ɹ] in contrast to Spanish [ɾ] and [r]) with a high degree of accuracy because these phones differ across multiple salient acoustic parameters, including duration and formant transitions. It was predicted that learners would have difficulty discriminating between the pairs [β]/[b] and [ɣ]/[g], since [β] and [ɣ] are not part of the English phonological inventory. What might seem surprising at first glance is that learners were not highly accurate in discriminating the [ð]/[d] items, which are ostensibly contrastive phonemes in English. However, one must remember that the phonetic realizations of these phonemic categories vary across languages. The English [ð] is typically produced as an interdental fricative. In contrast, the Spanish [ð], sometimes narrowly transcribed with an openness diacritic [ð̞], is typically produced as it was by the speaker in this study: not a fricative at all but rather as an approximant: a consonant characterized by low articulatory precision, lack of articulatory tension, and lack of turbulence in the airstream (Martínez-Celdrán, 2004). Historically Spanish [β, ð, ɣ] were often misclassified as fricatives, they are now more standardly described as approximants (Hualde, 2005; Martínez-Celdrán, 2004), none of which are phonemes of English.

Table 3. Mean scores by phone.

		+PI (n=46) *M* (SD)	-PI (n=41) *M* (SD)
[p] - [pʰ]	Pre	4.36 (0.75)	4.31 (0.90)
	Post	4.20 (1.07)	4.57 (0.83)
(5 items)	Delayed	4.29 (1.09)	4.33 (0.89)
[t] - [tʰ]	Pre	3.89 (1.01)	3.63 (1.33)
	Post	3.94 (1.15)	3.35 (1.46)
(5 items)	Delayed	3.94 (1.15)	3.35 (1.46)
[k] - [kʰ]	Pre	2.11 (1.12)	2.15 (1.23)
	Post	2.70 (1.23)	2.46 (1.39)
(5 items)	Delayed	2.70 (1.23)	2.46 (1.39)
[β] - [b]	Pre	2.38 (1.05)	2.83 (1.41)
	Post	2.87 (1.44)	2.15 (1.42)
(5 items)	Delayed	2.87 (1.44)	2.15 (1.42)
[ð] - [d]	Pre	3.09 (1.30)	3.30 (1.65)
	Post	4.02 (1.26)	4.08 (1.07)
(5 items)	Delayed	3.47 (1.20)	3.43 (1.77)
[ɣ] - [g]	Pre	2.43 (1.30)	2.38 (1.31)
	Post	2.96 (1.08)	2.80 (1.22)
(5 items)	Delayed	2.85 (1.35)	2.35 (1.37)
[ɾ] - [ɹ]	Pre	2.98 (0.14)	2.93 (0.25)
	Post	2.92 (0.28)	2.93 (0.25)
(3 items)	Delayed	2.88 (0.39)	2.88 (0.40)
[r] - [ɹ]	Pre	2.70 (0.51)	2.64 (0.65)
	Post	2.92 (0.28)	2.91 (0.35)
(3 items)	Delayed	2.96 (0.29)	2.98 (0.16)

It was predicted that learners' performance on the /p, t, k/ items would be homogeneous because all three are articulated in Spanish with a VOT that falls somewhere between the VOTs of English voiced and voiceless stops (Lisker & Abramson, 1964). For example, learners might perceive Spanish /p/ as something in between an English /b/ and /p/. Yet clearly learners had less difficulty discriminating [p] from [pʰ] than the other pairs. A more detailed acoustic analysis of the stimuli indicated that the [pʰ] tokens displayed some unexpected violent low frequency vibrations during the aspiration. Though a pop filter was used during the recording, it did not sufficiently dampen unwanted aperiodic activity. This subtle "pop" might have been an additional acoustic cue that helped learners discriminate between [p] and [pʰ] without having to rely on VOT. The fact that learners could discriminate [t]/[tʰ] better than [k]/[kʰ] is likewise likely explained by an additional cue, in this case place of articulation. Spanish /t/ is dental and English /t/ is alveolar, and the dental articulation leads to different formant transitions, particularly in the F2 of subsequent high vowels. Though VOT was the focus of analysis when creating the stimuli for the discrimination test in this study, clearly there were other articulatory and acoustic factors present that bore on the results.

Discrimination test scores were reliable; they passed tests of internal consistency, split-half homogeneity, and test/retest stability. However, a Shapiro-Wilk test indicated that the discrimination post and delayed posttest scores were not normally distributed ($p < .01$) but negatively skewed as a result of the high scores of some phones. A subsequent item analysis suggested that most /p, ɾ, r/ items were non-discriminatory and so they were eliminated from the analysis, thereby culling the discrimination test to 25 items, or 5 items per target phone [t, k, β, ð, ɣ].

Table 4 lists the pretest and gain scores for each learner on the 25-item discrimination test. Most learners had positive gain scores, though some learners did not change at all (gain score of 0), and a few learners actually scored lower on the posttests compared to the pretest. The majority of learners in the +PI group (78.26%) had positive pre- to posttest and pre- to delayed posttest (65.22%) gain scores, though a small proportion had negative gain scores (8.70% and 17.39%, respectively). In contrast, while the majority of learners in the -PI group (75.61%) also had positive pre- to posttest and pre- to delayed posttest (54.66%) gain scores, a relatively larger proportion of the learners in the -PI group had negative pre- to posttest gain scores (12.20%) and particularly pre- to delayed posttest gain scores (39.02%). Thus these frequency counts of positive and negative gain scores indicated that the vast majority of learners did demonstrate learning immediately after receiving either type of instruction (+PI or -PI) and that most learners demonstrated learning in the longer term as well, but the +PI group had an advantage in this regard.

Many of the negative gain scores were small and could have been due to normal variations in learner behavior on different test days and when experiencing slightly different conditions such as extraneous noises in the classroom, among other variables. However, a few of the negative gain scores were large and warranted further investigation. The learners whose gain scores were -10 or less were analyzed more closely, yet the analysis did not reveal any way in which this group of learners was much different from the rest. In their exit questionnaires they rated the helpfulness, usefulness, and difficulty of the instructional modules similarly to other learners, and they also reported feeling confident that their discrimination test scores were high (70% accurate or better). They were comparable to the rest of the learners in terms of Spanish language experience and general language experience we well. Age might have been a factor for two of the learners who began studying Spanish later in life (at 25 and 37 years of age).

To determine if particular individual difference factors predicted learners' ability to learn from instruction, learners' age, phonetic encoding ability, phonological short-term memory, and attitude towards pronunciation scores were entered into four standard linear multiple regression models with the dependent variables being the +PI and -PI groups' gain scores from pretest to posttest and from pretest to delayed posttest. Most of the models' explanatory power did not reach statistical significance ($p > .05$). Thus, these individual differences were not predictive of gain scores in general. One model was significant ($R^2 = .43$, $p = .006$) and showed that current age (coefficient -.61) contributed to -PI learners' pretest-to-delayed posttest gains. That is, the older a participant was at the time of the study, the less likely he was to improve and retain improvement in discrimination following exposure to the target phones through the focused listening with dictation activities. It was not surprising that individual difference factors would impact learning more under the more implicit of the learning conditions. A thorough discussion of the impact of individual difference measures on learning in the -PI group is beyond the scope of the current study, which aimed to examine the effect of explicit phonetics instruction (+PI) on L2 perception. However, sufficed to say that the age of the learners alone could not explain the advantage found for the +PI, because both groups included older learners (30 years of age or older), and in fact most of the older learners were randomly assigned to the +PI group (n= 5) rather than the -PI group (n= 2).

The +PI/-PI group differences that appeared to be present in the descriptive statistics of gain scores were tested with a repeated measures analysis of variance (RMANOVA). The within-groups factor was time of test (pre-, post-, and delayed posttest) and the between-groups factors were instructional condition (+PI and -PI) and course level (first, second and third year). Table 5 lists the results of the RMANOVA. There was a main effect of time, $F(1.75, 138) = 21.54$, $p < .001$, $\eta_p^2 = .21$, indicating that learners' scores changed after instruction. There was a significant time by condition interaction effect, $F(1.75, 138) = 6.17$, $p < .001$, $\eta_p^2 = .07$, indicating that 7% of the variance in scores could be explained by learners' instructional group (+PI or -PI). Figure 1 plots this interaction and indicates that the +PI group indeed did have an advantage in retention of learning. Independent samples T-tests, controlled for Type 1 error with the Bonferroni correction, confirmed that while the +PI did not gain significantly more than the -PI from pretest to posttest,

Table 4. Learners' individual scores.

+PI (n=46)				-PI (n=41)			
Learner	Pretest Score (of 25)	Gain Pretest to Posttest	Gain Pretest to Delayed Posttest	Learner	Pretest Score (of 25)	Gain Pretest to Posttest	Gain Pretest to Delayed
First Year				First Year			
1003	9	6	0	1010	16	3	1
1005	20	3	2	1014	12	8	0
1009	17	1	-4	1016	7	4	5
1013	15	2	-5	1022	12	7	3
1015	10	6	1	1024	17	6	1
1017	19	-1	0	1026	7	3	-2
1021	17	2	1	1028	11	-3	-7
1023	19	1	1	1030	14	7	-2
1027	14	8	3	1032	13	6	3
1029	18	0	3	1034	15	3	1
1031	17	3	0	1036	16	4	5
1035	15	5	3	1038	8	10	-1
1043	14	0	-10	1040	7	5	5
1045	9	6	2	1042	12	1	2
1047	18	1	3	1044	7	0	1
1051	13	7	0	1046	20	0	-7
1053	16	8	8	1050	12	9	5
1057	16	0	1	1052	20	-1	-2
1059	13	1	1	1054	5	10	2
				1058	22	-1	-7
Second Year				Second Year			
2001	19	-1	-1	2004	10	0	-4
2003	17	2	1	2006	21	3	2
2005	14	8	6	2010	11	8	5
2009	9	6	-1	2012	21	2	-13
2013	10	9	8	2014	14	7	0
2017	20	-4	0	2016	19	3	1
2019	14	0	1	2018	18	2	1
2023	15	1	-2	2030	11	3	-6
2025	8	2	1	2032	13	0	-9
2027	9	8	2	2034	25	-3	0
2029	7	14	5	2040	22	3	2
2031	17	3	0	2050	18	1	1
2035	19	1	2				
2039	8	3	6				
2043	6	2	3				
2045	12	6	8				
2049	20	3	1				

Table 4. (Continued).

Learner	Pretest Score (of 25)	Gain Pretest to Posttest	Gain Pretest to Delayed Posttest	Learner	Pretest Score (of 25)	Gain Pretest to Posttest	Gain Pretest to Delayed Posttest
		Third Year				Third Year	
3001	9	1	9	3002	19	-3	-17
3005	8	-1	13	3008	10	3	-4
3007	19	4	0	3010	19	0	-5
3013	15	4	-2	3012	15	5	4
3015	12	0	5	3014	21	2	-8
3019	14	3	0	3018	13	4	-8
3021	8	9	14	3020	9	6	8
3023	18	0	-5	3027	15	1	2
3024	16	2	1	3028	11	5	12
3029	8	8	7				

$t(85) = -0.23$, $p = .82$, the +PI retained significantly more from posttest to delayed posttest, $t(85) = 2.64$, $p = .01$, $d = 0.56$, as well as from pretest to delayed posttest, $t(85) = 2.51$, $p = .01$, $d = 0.54$. The pretest scores were not significantly different ($t(85) = -0.55$, $p = .58$), so these changes over time could not be accounted for by preexisting group differences.

Table 5. RMANOVA results: Aggregate of [t, k, β, ð, ɣ].

Source	*SS*	*df*	*MS*	*F*	η_p^2	*Power*
Between subjects						
Condition	1.47	1.00	1.47	0.03	0.00	0.05
Level	80.36	2.00	40.18	0.87	0.02	0.20
Condition X Level	332.97	2.00	166.49	3.61*	0.08	0.65
Error	3645.42	79.00	46.14			
Within Subjects						
Time	429.07	1.75	244.89	21.54**	0.21	1.00
Time X Condition	122.96	1.75	70.18	6.17**	0.07	0.85
Time X Level	47.78	3.50	13.63	1.20	0.03	0.34
Time X Condition X Level	52.63	3.50	15.02	1.32	0.03	0.38
Error (Time)	1573.70	138.41	11.37			

Significant at *$p < .05$, **$p < .001$
With a Greenhouse-Geisser Correction

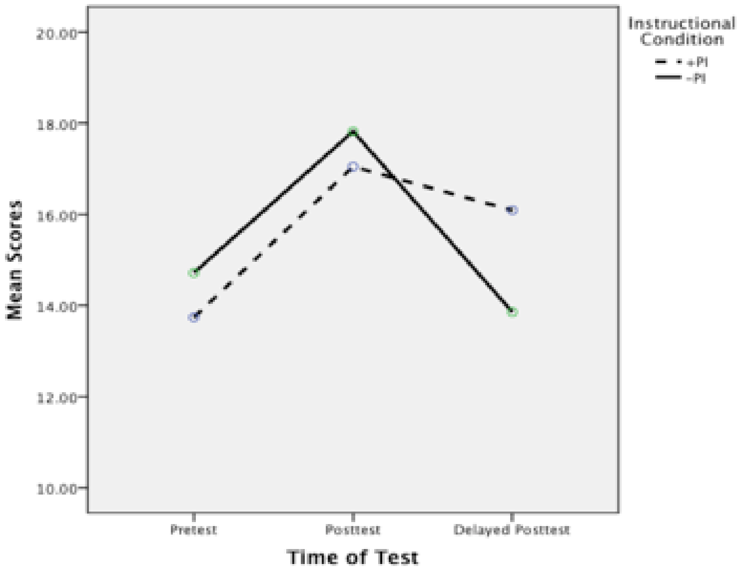
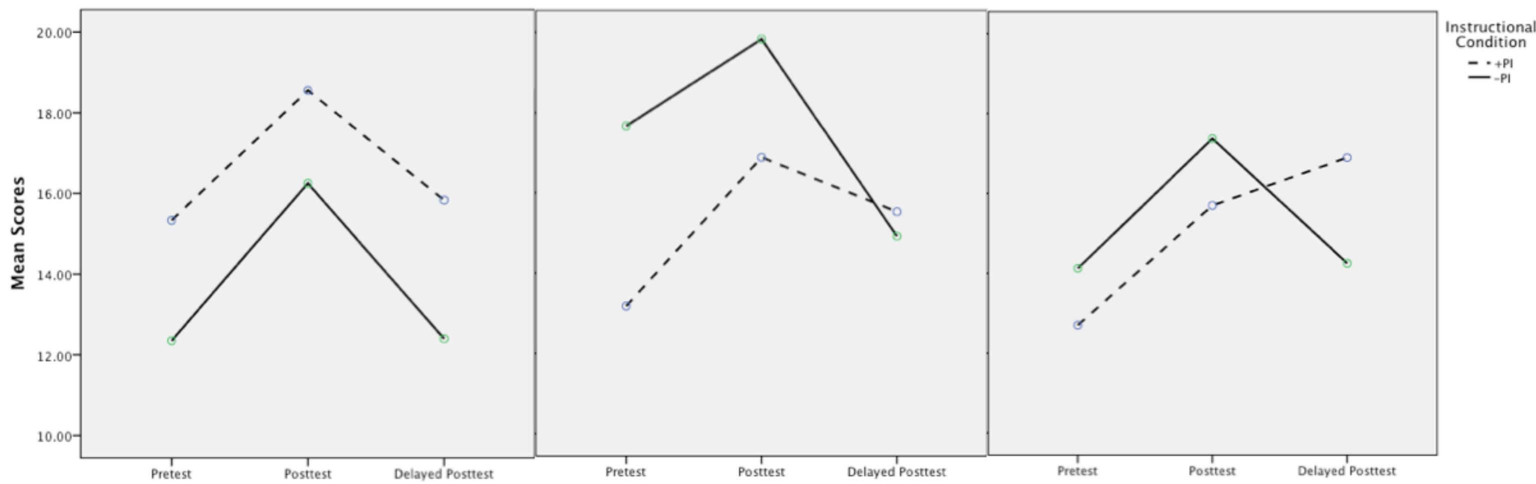


Figure 1. Discrimination test scores over time.

Figure 2 plots the discrimination test scores over time by course level and seems to indicate that the advantage of the phonetics instruction was conferred upon more advanced learners. However, neither the interaction of time by level nor the interaction of time by condition by level was significant, suggesting that course level did not influence learners' changes in scores. Since the observed power was below 80% for these tests, they should not be considered conclusive. It is possible that not enough participants were recruited, particularly at the more advanced levels.



First year
+PI (n=19)
-PI (n=20)

Second year
+PI (n=17)
-PI (n=12)

Third year
+PI (n=10)
-PI (n=9)

Figure 2. Discrimination test scores over time, by course level

A subsequent RMANOVA compared scores across individual phones. There were significant main effects for time and phone as well as significant interactions of condition by level and time by phone. Comparisons of the gain scores of each target phone were conducted in order to explore this phone type by time of test interaction. The results indicated essentially that learners in both instructional conditions improved most on their discrimination of /k/ and least on /t/. This simply may have been a product of learners' pretest scores, which were lowest for /k/ and highest for /t/, and determined the amount of improvement possible across the restricted range of scores. Another series of RMANOVAs conducted on the individual phones resulted in just one significant time by instructional condition interaction, for [β]/[b]: ($F(2, 160) = 15.53$, $p < .001$, $\eta_p^2 = .16$). The +PI group improved in discrimination of [β]/[b] from pretest to posttest and from pretest to delayed posttest more than the -PI group. The time by condition interaction did not reach significance for any other phone in isolation (all $F \leq 1.61$, $p \geq .10$).

Finally, recall that for each item of the discrimination test, participants rated how confident they were in their choice, on a scale of one to four. The confidence report data from the 25-item test were submitted to a RMANOVA, and the only significant effect found was the main effect for time, $F(1.72, 138) = 6.50$, $p = .002$, $\eta_p^2 = .08$. Both groups grew more confident over time. It is interesting to note that learners in the +PI felt more confident in the delayed posttest than the pretest, and their accuracy did indeed improve significantly during that period. In contrast, -PI learners also felt more confident in the delayed posttest even as their accuracy plateaued. However, a series of correlation analyses did not unearth any correlations between confidence reports and scores. It seems learners in this study could not accurately assess their ability to discriminate between L2 and analogous L1 phones.

IV Discussion

The goal of the present study was to assess the effectiveness of explicit L2 phonetics instruction in terms of improving learners' ability to perceive L2 phones, specifically eight consonantal Spanish phones that are problematic for L1 English speakers. It has been argued that target-like perception is important for acquiring target-like pronunciation, yet most prior classroom studies have only reported on post-instructional improvement in learners' production of these phones. The main finding was that even brief instruction in Spanish phonetics did afford learners an advantage in discriminating L2 phones from their analogous L1 counterparts. Though leaners' discrimination of the [p, ɾ, r] items was already at ceiling levels on the pretest, for the other target phones ([t, k, β, ð, ɣ]) the group that received phonetics instruction (+PI) had an advantage over the group that did not (-PI). The advantage was small, about half a standard deviation with an interaction effect size of 7%, and the advantage was apparent not immediately following instruction but rather three weeks later. These results suggest that explicitly drawing learners' attention to particular acoustic features of the L2 system was more expedient for most learners than merely exposing them to L2 sounds in the hopes that they would discover those relevant acoustic features, at least in the short term for these FL students. The control group received the same if not more exposure to the target sounds through instruction, along with similar production practice and feedback, and while most participants in the control group demonstrated improvement immediately after instruction, the effect disappeared later. The explicit phonetics instruction in this study likely oriented learners to attend to relevant aspects of Spanish phones, helping them break out of their automatic processing routines and better take in relevant acoustic information (Guion-Anderson, 2013). They were able to retain and use this knowledge in the discrimination task administered three weeks after instruction. To be sure, not all students benefitted from instruction (of either type), as evidenced by a few negative gain scores. Some students may have been more confused than helped by the instruction or not be able to apply what they learned to the AX perception test, which was unlike tasks they performed during instruction. However, not only did the +PI group have an advantage in terms of mean group scores but also in terms of the number of individual learners who demonstrated improvement (positive gain scores) immediately following instruction, and this advantage was even larger three weeks later.

Previously, researchers working with similar populations of learners have reported that phonetics instruction has little effect on reducing learners' aspiration of /p, t, k/ (González-Bueno, 1997; Lord, 2005), that in some cases instruction improves learners' pronunciation of [ɣ] but not [β, ð] (González-Bueno, 1997), and that the Spanish trill (/r/) continues to be problematic for learners well through the most advanced levels (Face & Menke, 2010). The present study complements those studies by demonstrating that learners' poor discrimination of L2 and analogous L1 phones might be a factor limiting their productive abilities. Many models of second language phonological acquisition characterize target-like perception as a predictor of production accuracy (Munro, 2008). Specifically, the Speech Learning Model claims that the first stage of acquiring L2 speech is to learn to discern subtle phonetic differences between L2 sounds and analogous sounds in the L1 as they occur in a variety of phonetic environments (Flege, 1988; Flege, 1995). The Spanish FL learners in the present study had difficulty discriminating target-like productions of the Spanish phones [t, k, β, ð, ɣ] from their English-accented counterparts ([t^h, k^h, b, d, g]), even when acoustically identical tokens of the phones were presented in mono- and di-syllabic stimuli that allowed participants to concentrate solely on their phonetic form. If the predictions of the SLM are correct, then it is not likely these learners would improve their pronunciation of the phones [t, k, β, ð, ɣ] without first improving their perception. Explicit phonetics instruction was found to be advantageous for improving their perception compared to a more implicit instructional condition.

The phones included in this study were all consonantal segments but otherwise constituted a rather heterogeneous set of phonological targets: three voiceless plosives, three approximants, and two rhotics. They are the segments most commonly targeted in research on phonetics instruction with similar populations of learners (Spanish FL learners in the United States). As expected, learners could discriminate the English /ɹ/ from Spanish rhotics quite well based on their salient acoustic differences, but they found the other phones much more difficult to discriminate. One exception was /p/, which was found to have problematic stimuli and was eliminated from the analysis. When phones were analyzed in isolation, the discrimination advantage conferred by phonetics instruction (+PI) was confirmed for only one phone, [β], and the other results per phone were inconclusive due to low statistical power. However, the primary aim of the current study was to quantify the effect of explicit phonetics instruction on leaners' perception, and as such it was not as concerned with any one phone in isolation so much as the aggregate effect over a range of phones that are typically taught in phonetics. The fact that learners discriminated rhotics well in the pretest does not suggest that rhotics should not be taught in phonetics, however. Certainly L2 learners mispronounce some sounds because they cannot hear them, and they mispronounce sounds that they *can* hear because they cannot *articulate* them (Chastain, 1976). Such is the case of the Spanish trill /r/, and instruction is likely to help learners learn to articulate it. Anecdotally, several learners reported on their exit questionnaires that they had not previously known how to produce the trill and that the animated vocal tract was very helpful.

Even in the explicitly instructed group, there was substantial variation in scores, indicating that not all learners benefited equally from instruction of either type. To test whether the variation was related to learners’ prior experience with Spanish, current course level was entered as a between-groups factor in the RMANOVA. The descriptive data indicated a trend towards the phonetics instruction being most advantageous for the more advanced learners, but the trend was not statistically significant. Though not conclusive, these results suggest that phonetics instruction is appropriate and beneficial at several points along the introductory and intermediate stages but will not benefit all learners equally. Course level is a very coarse measure of language experience, so other measures were also collected, including prior coursework, number of native speaking instructors, time spent abroad and use of Spanish outside the classroom, yet none of these language experience measures surfaced as a significant predictor of learning in either instructional condition (+PI or -PI). The question of curricular sequencing did seem material, however, in the analysis of the post-instructional questionnaires. First year learners in the +PI rated their lessons as better than the -PI overall, $t(38) = 2.04$, $p = .05$. Thus if one believes that learners’ appraisals of instructional materials may influence their engagement and/or learning, one should note that first year learners preferred the more explicit phonetics lessons.

In addition to prior language experience, several other individual difference measures were taken in an attempt to understand the variation in how learners responded to instruction, because research suggests that individual difference factors often interact with different instructional conditions in complex ways (Robinson, 2002). Those factors were age, phonetic script learning aptitude (measured with a subtest of the MLAT), phonological short-term memory (PSTM) and attitude towards pronunciation. Only age surfaced as a significant predictor of learning, and only in the more implicit condition. The older a learner was, the less likely he was to improve discrimination of the L2/L1 phone pairs following focused listening with dictation exercises. One should note that age of onset of learning, generally one of the best predictors of foreign accent

(Flege & MacKay, 2011) was not evaluated in this study because all the participants were "later" learners, most having begun study of Spanish around the age of 15. Based on reports that PSTM impacts L2 speech learning (MacKay, Meador, & Flege, 2001) and formation of new phonetic categories for L2 sounds (Aliaga-García, Mora, & Cerviño-Povedano, 2011), it was expected that PSTM would be predictive of gains, and yet it was not. However, differences in PSTM are meaningful when tasks actually tax learners' memory store. In the present study learners could pause the speech stream when they desired. Learners likely broke the speech in the +PI and -PI instructional modules into chunks that were of a length that they could easily hold in their phonological store, and this instructional delivery model may have rendered some individual differences irrelevant. Other possible interpretations are that the instructional treatments may simply have been too brief for learner differences to make a difference, the wrong set of factors may have been considered, or the measurements of those differences may have been too coarse. Other individual difference factors not considered but potentially predictive of L2 speech learning are phonological awareness (Venkatagiri & Levis, 2007), working memory, inhibitory control (Trude & Tokowicz, 2011), and pre-existing knowledge of Spanish phonetics. Nonetheless, it was encouraging to conclude that the results and implications of the preceding analysis could likely be extended to learners with various aptitude, attitude, and language exposure backgrounds. It was not the case that only learners with particularly high memory capacity, for instance, could learn from either instructional condition. The *lack* of significant results regarding individual difference factors was indeed considered a positive outcome for this investigation of instructional effectiveness.

The relative advantage of the +PI should not be overstated. The effect size was small enough to make it clear that brief lessons in phonetics is no panacea for learners' problems in perceiving L2 phones, much as previous studies have found that even semester-long courses in phonetics do not result in learners' production becoming fully target-like. Though not the focus of this study, the data indicate that focused listening with dictation led to immediate improvement in perception, and so a case can be made for exploring other potentially beneficial types of instruction in the FL curriculum. The results should not be interpreted as representing the maximum learning possible from phonetics instruction. Learners spent between one and two hours interacting with the instructional modules, depending on their pace. Though this instructional time was brief, it is unlikely that FL courses would devote much more time to a similar inventory of phones, and thus the results of the study can be interpreted as estimating the learning effects of brief phonetics lessons added on to an existing FL curriculum. Importantly, though, the increased scores of the +PI were not merely a practice effect. Each module in the phonetics instruction contained one identification exercise, in which learners identified isolated syllables as being either Spanish-like or English-like. The practice exercise was exceedingly explicit, directing learners to pay attention to particular features and ordering the stimuli in a predictable sequence. Thus, it did not provide task-specific practice for the discrimination test because the task conditions were quite different.

The present study included a heterogeneous yet limited inventory of phonological targets in that they were all consonantal segments and thus represent just one part of learners' phonological knowledge. The range of learners' levels was also restricted; none were very advanced or true beginners. The discrimination test also had limitations. Its stimuli were mono- and disyllabic items lacking semantic content, and so the results cannot be extrapolated to infer the accuracy with which learners could have discriminated the same phone pairs in more extended or meaningful language. The choice to use acoustically identical stimuli was justified as a means to ensure that learners interpreted the task instructions appropriately, given that the target phones could not be represented in minimal pairs. The long inter-stimulus interval presumably ensured that learners were drawing on their phonetic and phonemic knowledge, not just their acoustic memory store, to complete the task. However, most prior research (on contrastive L2 phones) has employed discrimination tasks with phonetically identical but acoustically different stimuli (that is, several different tokens of the same phone), and using acoustically identical stimuli could have unduly influenced learners' responses. Much work is needed in the area of developing experimental tasks that can investigate learners' perception of non-contrastive L2 phones.

V Conclusion

The current study offered evidence that explicit phonetics instruction, compared to a balanced and viable instructional alternative that was more implicit in nature, improved learners' discrimination of target-like Spanish phones and their English-accented counterparts. Thus the study constitutes an endorsement of explicit phonetics instruction in the Spanish FL context for those who believe that improved discrimination is of benefit to learners because perception leads production. Yet others are not so convinced, citing evidence that learners may produce contrasts they cannot perceive (Goto, 1971) and may acquire perceptive and productive abilities quite differently (de Jong, Hao, & Park, 2009; Eckman, Iverson, Fox, Jacewicz, & Lee, 2011). Though Zampini's (1998) work suggests that learners' perceptual boundaries for the Spanish stops /p/ and /b/ do not correlate with their production of the same phones, very few studies have attempted to investigate perception alongside production in the instructed FL context. The extent to which perceptive abilities might predict or inhibit productive abilities in this context is still very much an empirical question.

Acknowledgments

I am grateful to Rusan Chen, Jeff Conor-Linton, Cristina Sanz and Elizabeth Zsiga for their guidance during the data collection and analysis stages of this research. I am also grateful to the two *Language Teaching Research* reviewers for their helpful contributions to previous drafts of this article. Any remaining shortcomings are my own.

Funding

This research was partially funded with a grant from the journal *Language Learning.*

References

Akerberg, M. (2005). La percepción auditiva como factor en la adquisición de sonidos de L2. In M. Akerberg (Ed.), *Adquisición de segundas lenguas: Estudios y perspectivas* (pp. 45-72). México, DF: UNAM.

Aliaga-García, C., Mora, J. C., & Cerviño-Povedano, E. (2011). L2 speech learning in adulthood and phonological short-term memory. *Poznań Studies in Contemporary Linguistics, 47*(1), 1-14.

Archila-Suerte, P., Zevin, J., Bunta, F., & Hernandez, A. E. (2011). Age of acquisition and proficiency in a second language independently influence the perception of non-native speech. *Bilingualism: Language and Cognition, 15*(1), 190-201.

Ausín, A., & Sutton, M. (2010). An L2 pronunciation judgment task. In C. Borgonovo, M. Español-Echevarría & P. Prévost (Eds.), *Selected proceedings of the twelfth Hispanic linguistics symposium* (pp. 234-245). Somerville, MA: Cascadilla Proceedings Project.

Boomershine, A., Hall, K. C., Hume, E., & Johnson, K. (2008). The impact of allophony versus contrast on speech perception. In P. Avery, E. Dresher & K. Rice (Eds.), *Contrast in phonology* (pp. 143-172). Berlin: Mouton de Gruyter.

Bradlow, A. R., Yamada, R. A., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception and Psychophysics, 61*(5), 977-985.

Carroll, J. B. (1962). The prediction of success in foreign language training. In R. Glaser (Ed.), *Training research and education* (pp. 87-136). Pittsburgh, PA: University of Pittsburgh Press.

Champagne-Muzar, C., Scheneiderman, E., & Bourdages, J. (1996). Second language accent: The role of the pedagogical environment. *International Review of Applied Linguistics, 31*(2), 143-160.

Chastain, K. (1976). *Developing second language skills: Theory to practice*. Chicago, IL: Rand McNally.

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

de Jong, K., Hao, Y. C., & Park, H. (2009). Evidence for featural units in the acquisition of speech production skills: Linguistic structure in foreign accent. *Journal of Phonetics, 37*(4), 357-373.

Dehaene-Lambertz, G., Pallier, C., Serniclaes, W., Sprenger-Charolles, L., Jobert, A., & Dehaene, S. (2005). Neural correlates of switching from auditory to speech mode. *Neuroimage, 24*(1), 21-33.

Derwing, T. M., Munro, M. J., & Wiebe, G. E. (1997). Pronunciation instruction for “fossilized” learners: Can it help? *Applied Language Learning, 8*(2), 217-235.

Derwing, T. M., & Rossiter, M. (2003). The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech. *Applied Language Learning, 13*(1), 1-17.

Derwing, T. M. (2003). What do ESL students say about their accents? *Canadian Modern Language Review, 59*(4), 547-566.

Eckman, F. R., Iverson, G. K., Fox, R. A., Jacewicz, E., & Lee, S. (2011). Explicit training and implicit learning of L2 phonemic contrasts. In C. Sanz, & R. P. Leow (Eds.), *Implicit and explicit conditions, processes, and knowledge in SLA and bilingualism* (pp. 159-174). Washington, DC: Georgetown University Press.

Elliott, A. R. (1995). Foreign language phonology: Field independence, attitude, and the success of formal instruction in Spanish pronunciation. *The Modern Language Journal, 79*(4), 530-542.

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

Face, T. L. (2006). Intervocalic rhotic pronunciation by adult learners of Spanish as a second langauge . In C. A. Klee, & T. L. Face (Eds.), *Selected proceedings of the seventh conference on the acquisition of Spanish and Portuguese as first and second languages* (pp. 47-58). Somerville, MA: Cascadilla Proceedings Project.

Face, T. L., & Menke, M. R. (2010). Acquisition of the Spanish rhotics by L2 learners. *Current approaches to Spanish and Portuguese L2 phonology*, Gainesville, Florida.

Flege, J. E. (1988). The production and perception of speech sounds in a foreign language. In H. Winitz (Ed.), *Human communication and its disorders: A review* (pp. 224-401). Norwood, NJ: Ablex.

Flege, J. E., & MacKay, I. (2011). What accounts for "age" effects on overall degree of foreign accent? In M. Wrembel, M. Kul & K. Dziubalska-Kołaczyk (Eds.), *Achievements and perspectives in the acquisition of second language speech: New sounds 2010, vol 2* (pp. 65-82). Bern, Switzerland: Peter Lang.

Flege, J. E. (1995). Second-language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 229-273). Timonium, MD: York Press Inc.

Fullana, N. (2006). The development of English (FL) perception and production skills: Starting age and exposure effects. In C. Muñoz (Ed.), *Age and the rate of foreign language learning* (pp. 41-64). Clevedon: Multilingual Matters.

García Lecumberri, M. (1999). Influencia del tratamiento fonético en la percepción de vocales inglesas en una situación de L2. In J. Cuevas, & D. Fasla (Eds.), *Contribuciones al estudio de la lingüística aplicada* (pp. 181-188). Logroño, Spain: Gráficas Ochoa.

Golestani, N., & Zatorre, R. J. (2009). Individual differences in the acquisition of second language phonology. *Brain and Language, 109*(2-3), 55-67.

González-Bueno, M. (1997). *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/.* (Unpublished Doctoral Dissertation). The Pennsylvania State University, State College, Pennsylvania.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds 'L' and 'R'. *Neuropsychologia, 9*(3), 317-323.

Guion-Anderson, S. (2013). Speech perception. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 5348-5356). Oxford, UK: Blackwell Publishing.

Hualde, J. I. (2005). *The sounds of Spanish*. Cambridge, UK: Cambridge University Press.

Hwang, J. (2011). *Non-native perception and production of foreign sequences.* (Unpublished Doctoral Dissertatoin). State University of New York, Stony Brook, NY.

Jamieson, D. G., & Morosan, D. E. (1989). Training new, nonnative speech contrasts: A comparison of the prototype and perceptual fading techniques. *Canadian Journal of Psychology, 43*(1), 88-96.

Lado, B. (2008). *The role of bilingualism, type of feedback, and cognitive capacity in the acquisition of non-primary languages: A computer-based study*. (Unpublished Doctoral Dissertation). Georgetown University, Washington, DC.

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*, 384-382.

Logan, J. S., & Pruit, J. S. (1995). Methodological issues in training listeners to perceive non- native phonemes. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 351-377). Timonium, MD: York Press.

Lord, G. (2005). (How) can we teach foreign language pronunciation?: On the effects of a Spanish phonetics course. *Hispania, 88*(3), 557-567.

MacKay, I. R. A., Meador, D., & Flege, J. E. (2001). The identification of English consonants by native speakers of Italian. *Phonetica*, *58*, 103-125.

Markham, D. J. (1997). *Phonetic imitation, accent, and the learner*. Lund: Lund University Press.

Martínez-Celdrán, E. (2004). Problems in the classification of approximants. *Journal of the International Phonetic Association*, *34*(2), 201-210.

Matthews, J. (1997). The influence of pronunciation training on the perception of second language contrasts. In J. Leather, & A. James (Eds.), *New sounds 97: Proceedings of the third international symposium on the acquisition of second-language speech* (pp. 223-229). Klagenfurt, Austria: University of Klagenfurt.

McCandless, B. D., Fiez, J. A., Protopapas, A., Conway, M., & McClelland, J. L. (2002). Success and failure in teaching the r-l contrast to Japanese adults: Predictions of a hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective, and Behavioral Neuroscience, 2*(2), 89-108.

McCandless, P., & Winitz, H. (1986). Test of pronunciation following one year of comprehension instruction in college German. *Modern Language Journal, 70*(4), 355-362.

Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age, motivation and instruction. *Studies in Second Language Acquisition, 21*(1), 81-108.

Munro, M. N. (2008). *Aptitude for novel speech sounds and sound sequences: Implications for second-language pronunciation.* (Unpublished Doctoral Dissertation). University of California, Irvine, Irvine, CA.

Newman, R. S. (1996). Individual differences and the perception-production link. *Journal of the Acoustical Society of America, 99*(4), 2592-2603.

Pardo, D. B. (2004). Can pronunciation be taught?: A review of research and implications for teaching. *Revista Alicantina De Estudios Ingleses, 17*(1), 6-38.

Pennington, M. (1992). Reflecting on teaching and learning: A development focus for the second language classroom. In J. Flowerdew, M. Brock & S. Hsia (Eds.), *Perspectives on second language classroom teacher education* (pp. 47-65). Kowloon: City Polythenic of Hong Kong.

Piske, T., MacKay, J. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29*(2), 191-215.

Polka, L. (1992). Characterizing the influence of native language experience on adult speech perception. *Perception and Psychophysics, 52*(1), 37-52.

Robinson, P. (2002). *Individual differences and instructed language learning*. Amsterdam: John Benjamins.

Rochet, B. L. (1995). In Strange W. (Ed.), *Perception and production of second-language speech sounds by adults*. Baltimore: York Press.

Strange, W., & Shafer, V. L. (2008). Speech perception in second language learners: The re- education of selective perception. In G. J. Hansen Edwards, & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 153-191). Amsterdam: John Benjamins.

Tees, R. C., & Werker, J. F. (1984). Perceptual flexibility: Maintenance or recovery of the ability to discriminate non-native speech sounds. *Canadian Journal of Psychology, 38*(5), 579-590.

Thompson, I. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants. *Language Learning, 41*(2), 177-204.

Timmis, I. (2002). Native-speaker norms and international English: A classroom view. *English Language Teaching Journal, 56*(3), 240-249.

Trude, A. M., & Tokowicz, N. (2011). Negative transfer from Spanish and English to Portuguese pronunciation: The roles of inhibition and working memory. *Language Learning, 61*(1), 259-280.

Venkatagiri, H. S., & Levis, J. M. (2007). Phonological awareness and speech comprehensibility: An exploratory study. *Language Awareness, 16*(4), 263-277.

Werker, J. F., & Logan, J. S. (1985). Cross-language evidence for three factors in the perception of non-native speech sounds. *Perception and Psychophysics*, *37*(1), 35-44.

Werker, J. F., & Tees, R. C. (1984). Phonemic and phonetic factors in adult cross language speech perception. *Journal of the Acoustical Society of America*, *75*(6), 1866-1878.

Wode, H. (1981). *Learning a second language: An integrated view of language acquisition*. Tubingen: Narr.

Yule, G., Hoffman, P., & Damico, J. (1987). Paying attention to pronunciation: The role of self-monitoring in perception. *TESOL Quarterly, 21*(4), 765-768.

Zampini, M. L. (1993). *Spanish and English voiced stop phonemes and spirantization: A study in second language acquisition.* (Unpublished Doctoral Dissertation). Georgetown University, Washington, DC.

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.