Phonetic drift in a first language dominant environment

Wendy J. Herd, Robin L. Walden, Whitney L. Knight and Savana N. Alexander
Department of English, Mississippi State University, Mississippi State, MS; wherd@english.msstate.edu, rlf121@msstate.edu, whitneyleighknight@gmail.com, sna62@msstate.edu

Phonetic drift, changes in the first language (L1) sound system as a result of acquiring a second language (L2), has been documented in learners immersed in L2-dominant environments. Less attention has been given to phonetic drift in speakers learning an L2 in an L1-dominant environment. This study presents data from a cross-sectional analysis of English learners of Spanish in the United States at beginning (N=12), intermediate (N=12), advanced (N=10), and near native (N=6) proficiency levels. Participants were recorded reading a pseudo-randomized list of words including poll, bowl, toll, dole, coal, goal to measure drift in oral stops and heed, hayed, who'd, hoed to measure drift in vowels. Significant differences in the VOT of voiced stops and in vowel quality were found. Advanced and near native learners of Spanish produced voiced stops with more negative VOTs than beginning learners. Similarly, intermediate, advanced, and near native learners produced vowels in more peripheral positions of the vowel space than beginning learners. All of these effects were strongest in near native learners. These results suggest that phonetic drift occurs not only when learners are immersed in L2-dominant environments but also as a result of language instruction in L1-dominant environments.
1. Introduction

When language learners acquire a second language (L2), their first language (L1) sound system affects the acquisition of the L2 sound system, often resulting in a perceived “foreign accent” (e.g., Munro and Derwing, 1995; Derwing and Munro, 1997; see Moyer, 2013 for a thorough treatment of accent). Research also indicates that acquiring an L2 alters the L1 sound system, with the resultant alteration in the L1 sound system called phonetic drift or gestural drift. Phonetic drift has been well-attested in the speech of learners with experience in L2-dominant environments.

In a study of English learners of French living in Paris and French learners of English living in Chicago, Flege (1987) found that advanced English learners of French pronounced English /t/ with shorter voice onset times (VOTs) than English monolinguals and that advanced French learners of English pronounced French /t/ with longer VOTs than French monolinguals. In other words, acquiring advanced proficiency in French resulted in native English speakers pronouncing a more French-like /t/ in their native English. Similarly, advanced French learners of English pronounced a more English-like /t/ when speaking in their native French. In a related longitudinal study, Sancier and Fowler (1997) compared the VOTs of Portuguese /p/ produced by a Portuguese advanced learner of English after several months in Brazil and after several months in the United States. They reported that the VOTs of /p/ in the speaker’s native Portuguese were significantly longer, or more English-like, after a stay in the United States than after a stay in Brazil. More recently, Chang (2012, 2013), who studied native English speakers enrolled in a six-week intensive Korean course in South Korea, reported significant changes in the VOT and fundamental frequency (F0) of participants’ English voiced stops and in spectral qualities (F1 and F2) of the participants’ English vowels. This drift occurred within the first five weeks of language immersion, and phonetic drift among novice learners was stronger than that found in heritage speakers of Korean, suggesting that phonetic drift occurs early in language instruction but that it might stabilize during later stages of acquisition.

In the above studies, language learners were recorded in, or shortly after returning from, an environment where the L2 was the dominant language (i.e., L2-dominant environment). Thus, phonetic drift due to exposure to an L2-dominant environment has been documented in the speech of novice language learners and in the speech of highly advanced language learners. The purpose of this study is to determine if phonetic drift also occurs when language learners acquire an L2 in an L1-dominant environment. In particular, this study uses a cross-sectional design to investigate whether English speakers from Mississippi experience phonetic drift as a result of acquiring Spanish in an English-dominant environment.

1.1 VOT differences between Spanish and Southern English stops

While Spanish and English include a similar inventory of stop consonants, the two languages differ in how they distinguish voiced and voiceless stops. In general, Spanish speakers produce word-initial voiced stops /b, d, ɡ/ with negative VOTs, or prevoicing, and voiceless /p, t, k/ as voiceless unaspirated stops with short positive VOTs, but English speakers tend to produce word-initial voiced stops /b, d, ɡ/ with short positive VOTs and voiceless stops /p, t, k/ with long positive VOTs, or aspiration (Lisker and Abramson, 1964; Magloire and Green, 1999).

It should also be noted that the participants in this study were native speakers of Southern English from Mississippi. Working with speakers from two different dialect regions, Jacewicz,
Fox and Lyle (2009) found that speakers of Southern English from North Carolina were more likely to produce voiced stops with negative VOTs than speakers from Wisconsin. While this means the Southern English speakers in our study might be more likely to produce voiced stops with negative VOTs than speakers of other varieties of English, the same is true of all of our participants, so cross-sectional comparisons based on Spanish experience will still be valid.

Based on these acoustic characteristics, we predicted that English-speaking learners of Spanish with more experience would produce both voiced and voiceless oral stops with shorter VOTs than learners with less experience. Specifically, we expected that learners with more experience would be more likely than their less experienced counterparts to produce voiced stops with negative VOTs and voiceless stops with short-lag VOTs.

1.2 Spectral differences between Spanish and Southern English vowels

In a cross-linguistic comparison of the vowels /i, e, u, o/ as produced by four native Spanish speakers from Madrid and four native English speakers from New York, Bradlow (2005) found that Spanish vowels were consistently produced with a lower F2 (i.e., further back in the mouth) than English vowels. With respect to F1, Spanish front vowels were produced with higher F1 values than English vowels (i.e., lower in the mouth), but Spanish back vowels were produced with lower F1 values than English vowels (i.e., higher in the mouth).

As with VOT, these F1 and F2 values must be considered within the context of dialect variation. In cross-dialectal studies, Fox and Jacewicz (2009) and Clopper, Pisoni and de Jong (2005) found that speakers of Southern English produce more centralized tense vowels than speakers from other dialects. Due to the Southern Vowel Shift, speakers of Southern English tend to produce both /i/ and /e/ with higher F1 and lower F2 values, in some cases even reversing position at the vowel nucleus with their lax counterparts /ɪ/ and /ɛ/ (Labov, 1998; Fox & Jacewicz, 2009; Clopper, Pisoni & de Jong, 2005). Southern English speakers also exhibit Back Vowel Fronting, which is also sometimes considered part of the Southern Vowel Shift, such that /u/ and /o/ are produced with higher F2 values than what is reported for speakers of other varieties of English (Labov, 1998; Clopper, Pisoni & de Jong 2005).

Thus, we predicted that Southern English speakers with more Spanish experience would pronounce the vowels /i, e, u, o/ in more acoustically peripheral positions than those with less Spanish experience. The prediction that speakers with more experience will produce /u/ and /o/ with lower F1 and F2 values is consistent with the findings from Bradlow (2005). While our prediction that more experience with Spanish will lead to producing /i/ and /e/ with lower F1 and higher F2 values seems to contradict Bradlow’s (2005) findings, this prediction is consistent with Fox and Jacewicz (2009) and Clopper, Pisoni and de Jong’s (2005) description of vowels in Southern English.

2. Methods

2.1 Participants

In order to determine if experience with Spanish in an L1-dominant environment results in phonetic drift in L1 English, Southern English-speaking learners of Spanish in Mississippi were recruited from their university-level Spanish courses. Forty participants were recruited across four levels of proficiency: early learners of Spanish who were enrolled in Spanish 1
intermediate learners who were enrolled in Spanish 3 (N=12; 6 male and 6 female), advanced learners who were enrolled in advanced composition or conversation courses and who were majoring in Spanish (N=10; 4 male and 6 female), and graduate learners who were completing their MA in Spanish (N=6; 3 male and 3 female). All participants grew up in the South, with 38 participants from Mississippi, 1 graduate student from Pocahontas, TN, which is approximately 5 miles from Mississippi, and 1 graduate student from Tallahassee, FL. Considering the larger number of Spanish speakers living in Florida compared to Mississippi, it is important to note that the student from Florida did not report Spanish-speaking experiences different from the other graduate students. All participants were between the ages of 18 and 30. With the exception of the graduate students, none of the participants reported visiting a Spanish-speaking country for more than one week. Of the six graduate students, four reported visiting a Spanish-speaking country, with stays ranging from 1 week to 3 months. In short, none of the participants reported extensive or recent stays in Spanish-speaking countries. Participants were compensated $10 per hour for their participation in the study.

2.2 Procedure

After providing informed consent, participants completed a brief language background questionnaire in which they self-reported their basic demographic information (i.e., gender, age, hometown, and first language), knowledge of other languages, classroom experience with other languages, informal experience with other languages, and experience in L2-dominant environments. Participants were next recorded reading a word list three times while seated in a sound-attenuated booth (WhisperRoom 4848S) via an Electro-Voice N/D 767a microphone. Recordings were saved on a flash card using a Marantz Portable Solid State Recorder (PMD 661) at a sampling rate of 44.1 kHz.

2.3 Stimuli and Measurements

The experimental stimuli consisted of a minimal set of 6 items that contrasted oral stops of English, a minimal set of 11 items that contrasted vowels of English, and 75 additional filler items drawn from Egan’s (1948) PB-50 lists. The six oral stop tokens, poll, bowl, toll, dole, coal, goal, were included to determine if more advanced learners of Spanish produced English words with shorter (i.e., more Spanish-like) VOTs than less advanced learners. In the case of /p, t, k/ and short-lag /b, d, ɡ/, VOT was measured from the onset of the release burst to the first glottal pulse of the following vowel. When /b, d, ɡ/ were prevoiced, VOT was measured from the onset of voicing to the onset of the initial release burst. As noted in 2.2, participants produced three repetitions of each experimental stimulus. VOT values were averaged across repetitions for each consonant by speaker and submitted to statistical analyses.

In order to compare the production of vowels by speakers of different gender and size, all repetitions of all vowel items (i.e., heed, hid, hayed, head, had, hud, who’d, hood, hoed, hawed, hod) were used to normalize F1 and F2 values by speaker using a z-score normalization method (Lobanov, 1971). The Lobanov z-score method is noted for minimizing nonsystematic individual differences while maintaining systematic sociolinguistic variation (Adank, Smits, & van Hout, 2004). Vowel normalization was completed using the vowel package in R (Kendall & Thomas, 2009). Using Praat, F1 and F2 values were measured in a 20 ms window at vowel midpoint (Boersma & Weenink, 2013). The normalized values of the four vowels that overlap in English...
and Spanish (i.e., heed, hayed, who’d, hoed) were used for statistical analyses. As noted in 2.2, participants produced three repetitions of each experimental stimulus. After vowel normalization, F1 and F2 values were averaged across repetitions for each vowel by speaker.

2.4 Predictions

As mentioned in 1.1 and 1.2, we predicted that learners of Spanish with more L2 experience would produce shorter oral stop VOTs and more peripheral vowels in English than less experienced learners of Spanish. Specifically, we predicted that more Spanish experience would lead to the production of 1) voiceless oral stops with shorter positive VOTs, 2) voiced oral stops with more negative VOTs, 3) /i/ and /e/ with lower F1 and higher F2 values, and 4) /u/ and /o/ with lower F1 and F2 values.

3. Results

3.1 Oral Stops

In order to determine if English oral stop VOTs varied with Spanish proficiency, one-way ANOVAs were run on the VOT of each oral stop with group (i.e., Spanish proficiency) as a between-subjects factor. As clearly seen in Figure 1, Spanish learners of all proficiencies produced voiceless oral stops /p, t, k/ with long positive VOTs; no group differences were found (all p values > 0.05). However, the groups differed significantly in their production of /b/ [F (3, 39) = 9.338, p < 0.001], /d/ [F (3, 39) = 3.583, p = 0.023], and /ɡ/ [F (3, 39) = 3.132, p = 0.037]. Pairwise comparisons (Fisher’s LSD) determined that graduate students produced all voiced oral stops with significantly more negative VOTs than beginning (/b/: p < 0.001, /d/: p = 0.017, /ɡ/: p = 0.005), intermediate (/b/: p = 0.001, /d/: p = 0.012, /ɡ/: p = 0.019), and advanced (/b/: p = 0.001, /d/: p = 0.003, /ɡ/: p = 0.038) students.

![FIG. 1. Mean VOT of English (L1) oral stops by Spanish (L2) proficiency level.](image_url)

While voiceless oral stops were consistently produced with long-lag VOTs by all participants from all groups, there was more variation in the production of voiced oral stops such that some individuals produced primarily short-lag positive VOTs, some produced primarily
negative VOTs, and others produced a combination of the two. As such, the values used in the analyses above and shown in Figure 1 sometimes represented the average of positive and negative VOTs, particularly in the case of beginning, intermediate, and advanced learners. Thus, a one-way ANOVA was conducted on the percentage of oral stops produced with negative VOTs with group as a between-subjects factor to see if advanced learners prevoiced stops more often than less advanced learners. The percentage of prevoiced stops produced by group is presented in Figure 2. There was a significant difference in the percentage of voiced stops produced with negative VOTs such that advanced and graduate students produced more prevoiced oral stops than other participants \[F(3, 39) = 4.642, p = 0.008\]. Pairwise comparisons (Fischer’s LSD) revealed that graduate students produced significantly more prevoiced stops than beginning \(p = 0.002\) and intermediate \(p = 0.006\) students and that advanced students produced significantly more than beginning students \(p = 0.034\).

**FIG. 2.** Percentage of English (L1) voiced oral stops with negative VOTs by Spanish (L2) proficiency level.

### 3.2 Vowels

In order to determine if F1 and F2 values of English vowels varied with Spanish proficiency, separate one-way ANOVAs were run on the Lobanov-normalized F1 and F2 values of /i, e, u, o/ with group as a between-subjects factor. Normalized values for these vowels as produced by proficiency group are shown in Figure 3. For the front vowels /i, e/, groups differed significantly in the F1 values of /i/ \[F(3, 39) = 3.069, p = 0.040\], the F1 values of /e/ \[F(3, 39) = 4.017, p = 0.015\] and the F2 values of /e/ \[F(3, 39) = 6.806, p = 0.001\]. While graduate and intermediate students produced /i/ with numerically lower F1 values than advanced and beginning students, these differences did not reach significance in pairwise comparisons (all \(p\) values > 0.05). With respect to /e/, graduate students produced significantly lower F1 values than beginning \(p = 0.003\) and intermediate \(p = 0.045\) students, and advanced students produced significantly lower F1 values than beginning students \(p = 0.023\). Likewise, graduate students produced significantly higher F2 values than beginning \(p < 0.001\) and advanced \(p = 0.004\) students, but the difference between graduate and intermediate student F2 values only neared...
As predicted, graduate students and advanced students produced a higher /e/ than less experienced learners, and graduate students’ production of /e/ was more fronted than less experienced learners.

Groups also differed significantly in F1 values for /u/ \([F (3, 39) = 3.244, p = 0.033]\) and /o/ \([F (3, 39) = 7.133, p = 0.001]\). According to pairwise comparisons, both graduate (\(p = 0.049\)) and intermediate (\(p = 0.005\)) students produced /u/ with significantly lower F1 values than beginning students. While advanced students also produced /u/ with numerically lower F1 values than beginning students, the difference did not reach significance. For /o/, however, graduate (\(p = 0.001\)), advanced (\(p = 0.002\)), and intermediate (\(p = 0.001\)) students produced significantly lower F1 values than beginning students. Generally, students with more Spanish experience produced higher back vowels than beginning Spanish students. Differences in vowel fronting (i.e., F2 value) did not reach significance for /u/ or /o/.

4. Discussion

The purpose of this study was to determine if English learners of Spanish who are acquiring Spanish in an English-dominant environment exhibit phonetic drift in the same way as learners acquiring an additional language in an L2-dominant environment. Group differences were found for both stops and vowels; however, graduate students exhibited the most robust evidence of phonetic drift.

Significant group differences were found in the VOT of voiced, but not voiceless, stops. Participants across all proficiency levels produced voiceless stops with the long lag VOTs typical of English, but the average VOT for voiced stops was negative for all groups, with numerically more negative values as proficiency level increased. First, it is notable that such a large number of stops were prevoiced, with even the beginning learners of Spanish prevoicing 44%. While this large amount of prevoicing may seem atypical for English speakers, it is consistent with
Jacewicz, Fox and Lyle (2009), who found that Southern English speakers from North Carolina produced more prevoiced tokens than speakers from Wisconsin. Second, the percentage of prevoicing increased with more advanced proficiency, with graduate students producing significantly more negative VOTs than less experienced participants and with both advanced and graduate students producing significantly larger percentages of prevoiced tokens than less experienced learners. These results suggest phonetic drift, as measured by changes in VOT of oral stops, is possible in L1-dominant environments.

Similarly, significant group differences were found in the production of vowels. In particular /e, u, o/ exhibited the most drift. The vowel /e/, which is typically centralized (i.e., higher F1 and lower F2 values) by speakers of Southern English, was produced in a more peripheral position by advanced and graduate students than less experienced students. Likewise, more advanced learners produced /u/ and /o/ higher in the mouth (i.e., with lower F1 values) than beginning students. The F1 results are consistent with our general prediction that advanced learners would produce more peripheralized vowels than beginning learners, but advanced learners did not produce /u, o/ in a less fronted position than beginning learners, particularly not for /u/. In fact, graduate, advanced, and intermediate learners fronted /u/ more than beginning students as seen in Figure 3, but the difference did not reach significance. There are several possible explanations for the behavior of /u/. First, the results could be due to the pervasiveness of Back Vowel Fronting, particularly in Southern English, but, if this is the case, there is no reason to predict beginning students would perform differently than learners with more Spanish experience. Another possibility to consider is that participants were well-matched for gender across groups but that no attempt was made to match participants for ethnicity or race. If the beginning student group was more ethnically diverse than the other three groups, this could account for their lack of /u/ fronting. Finally, and most optimistically, it is possible that the more advanced students formed separate categories for English and Spanish /u/, differentiating the more fronted English /u/ from a Spanish /u/ with lower F2 values. Unfortunately, the data collected for the current study does not address these possibilities. To assess the latter two, the participants would also need to self-report their ethnicity or race and would need to be recorded reading a Spanish word list. Additionally, the cross-sectional design of this study may miss the type of early phonetic drift Chang (2012) and Chang (2013) reported in English learners of Korean. A longitudinal study that follows English learners of Spanish might be more sensitive to early rapid changes as well as to individual differences in the amount of prevoicing and as to the presence of the Southern Vowel Shift and Back Vowel Fronting prior to Spanish acquisition.

As expected, graduate students exhibited the most phonetic drift, but one could argue that the graduate students, most of whom had studied abroad if only briefly, represent a different population than the less experienced learners. It is thus particularly important to note that advanced students, learners who had never studied in an L2-dominant environment, produced significantly more prevoiced tokens than beginning students. Similarly, intermediate and advanced students with no study abroad experience produced vowels differently from beginning students. Taken together, these results suggest that phonetic drift in an L1-dominant environment is possible but that it is likely a slower process than that reported in L2-dominant environments.
References


