The Pronunciation of the Aspirated Consonants *P, T,* and *K* in English by Native Speakers of Spanish and French

Philip Suarez
Florida International University, USA

Abstract: This study examines whether explicit phonetic instruction helps Spanish and French speakers more closely approximate native pronunciation of the aspirated variants of the English consonants *p, t,* and *k.* The study results indicate that phonetic instruction clearly benefited the subjects, and the findings warrant further investigation.

Native speakers of English consistently use aspirated variants of the consonants *p, t* and *k* when they occur at the beginning of words (i.e., these three consonants are pronounced with a small puff of air when they occur in initial position). These aspirated variants are *phonemes,* basic units of sound that native speakers can distinguish; phonemes are indicated by a slash mark on either side. The use of the aspirated phonemes helps the listener to distinguish minimal pairs in English, which are pairs of words, such as *pill/bill,* that differ only by a single sound. English consonants that are aspirated often present difficulties for speakers of languages, such as Spanish or French, which lack these sounds. Does explicit phonetic instruction that targets the description and practice of these aspirated consonants offer any benefit to native speakers of Spanish and French?

This study specifically focuses on the production of the aspirated variants of *voiceless stop consonants* in English by native speakers of Spanish or French. English, like Spanish and French, has two sets of *stop consonants,* consonants produced by stopping the airflow completely. The phonemes /b/, /d/ and /g/ are *voiced stops* because the vocal chords vibrate; /p/, /t/ and /k/ are *voiceless stops* because the vocal chords do not vibrate. This underscores the function and importance of aspirating initial voiceless *stops* in English because the single small difference between the initial consonants in pairs like *pill/bill* is that one is voiced and the other is voiceless. Spanish and French speakers may simply fail to aspirate initial voiceless *stops* in English as a result of familiarity with the written forms of Spanish, French, and English. They may not realize that the letters *p, t,* and *k* do not necessarily represent the same phonemes in English as they do in Spanish or French. Further, they may not understand that aspirating initial voiceless *stops* is an important phonological cue for the listener and that communication may be unsuccessful when the expected aspiration is absent.

This paper reports on a study of pronunciation patterns for English language learners whose native language, which is hereafter referred to as L1, is Spanish or French. The students were tested, both before and after phonetic instruction, for the pronunciation of /p/, /t/, and /k/ in initial position as they occur in English, the second language that the student is learning. The student's second language is hereafter referred to as L2. The data and results from the studies reviewed below rather convincingly demonstrate the benefits of explicit phonetic instruction in developing phonological L2 competence. The bulk of existing studies concentrate on L1 English speakers who are learning Spanish or French. Consequently, it seems justified to develop a methodology that measures the effect of

Suarez, P. (2008). The pronunciation of the aspirated consonants *p, t,* and *k* in English by native speakers of Spanish and French. In M. S. Plakhotnik & S. M. Nielsen (Eds.), *Proceedings of the Seventh Annual College of Education Research Conference: Urban and International Education Section* (pp. 144-150). Miami: Florida International University.
http://coeweb.fiu.edu/research_conference/

focused phonetic instruction on L2 English phonological competency by L1 Spanish and French speakers because this has received relatively scant attention in the literature.

Review of the Literature

Much of the current literature (e.g., Brecht, Davidson, & Ginsberg, 1995; Diaz-Campos, 2004; Lapkin, Hart, & Swain, 1995) concerning phonetic instruction that addresses the differences between English phonology and Spanish or French phonology has focused on L1 English speakers learning Spanish or French, most often in study abroad programs. The relative dearth of information regarding the phonological problems of L1 Spanish or French speakers learning English reveals a need that remains to be addressed adequately. This is particularly true in light of the large immigrant population of Spanish speakers in the United States and the high volume of commerce between the United States, Mexico, and Canada. Any research that evaluates methods of English phonetic instruction for Spanish or French speakers offers a much-needed contribution to pedagogy.

The most basic question to address is to what extent L2 learners acquire new phonetic categories in the process of learning English. Flege and Eefting (1988) gathered data from a group of L1 English speakers, a group of monolingual L1 Spanish speakers and a group of bilingual L1 Spanish speakers. They created 16 test sounds on a continuum ranging from an unaspirated variant of the phoneme /d/ to an aspirated variant of the phoneme /t/. Their findings revealed that the bilingual L1 Spanish speakers were able to imitate and distinguish three categories, from non-aspirated to aspirated variants, defined by voice onset time (VOT; i.e., the length of time between the beginning of the consonant to the beginning of the following vowel). They concluded that the bilingual L1 Spanish speakers had acquired a new phonetic category for the aspirated variant of /t/.

Although the above study provides the most relevant type of data, much more attention has been devoted to L1 English speakers learning Spanish, comparing the effects of both phonetic instruction and study abroad. These studies also create a better understanding of the role that English and Spanish phonology play in L2 acquisition. However, this kind of data can only be extrapolated in order to gain insight into the acquisition of L2 English. González-Bueno (1997) studied the use of aspirated variants of voiceless *stops,* such as /k/, in the perception of foreign accent in the speech of L1 English speakers learning Spanish. She created 28 unaspirated and aspirated stimuli. Eighteen L1 Spanish speakers from Seville, Spain, were asked to judge the speech samples. The variants having a VOT ranging from 15 ms to 35 ms were perceived as native. Since the perception of native speakers varied according to how VOT was manipulated, González-Bueno suggested that phonetic instruction be used to shorten the VOT of voiceless *stops* in order to produce more native-like pronunciation in L2 Spanish.

Even more striking was the data that Lord (2000) obtained in a study of the effects of phonetic instruction versus study abroad. Her focus was on the production of voiced *stops* (*b, d,* and *g*) when they occur between two vowels, which in Spanish become fricatives (i.e., the airflow is constricted but does not stop). Eight students were enrolled in an 8-week study-abroad program in Mexico. The experimental group of four students had taken a Spanish phonetics course one to two semesters before the study-abroad program, the other four had not. None of the students had any problem producing voiced *stops* (*b, d,* and *g*) in the pretest and post-test. Lord suggested that the *stop* variants were the default value as a result of transferring the voiced *stop* values from L1 English. For the intervocalic fricative variants, the average accuracy for the control group was 3.3%, compared with 8.6% for the

experimental group. The post-test showed 5.8% accuracy for the control group and 28.7% accuracy for the experimental group. The effect of phonetic instruction here, especially in combination with study abroad, is impossible to overlook.

The transfer of voiced *stops* from L1 English that Lord (2000) noticed indicates that they are unproblematic for English speakers. In the same way, because the variants of English voiceless *stops* are unaspirated when they are not in initial position, they will be easy to produce for L1 Spanish or French speakers, compared with the aspirated variants. The use of generative phonology by Flege and Eefting (1988) to explain the acquisition of a new phonetic category for the aspirated variant of /t/ bears directly on studies of L1 Spanish or French speakers acquiring the aspirated variants of initial voiceless *stops* in English.

Three ESL primary school teachers in Murcia, Spain, conducted another significant study. Blanco, Gayoso, and Carrillo (2001) sought to measure the effectiveness of an initial teaching program of English that concentrated on L2 phonology. They worked with experimental and control groups of second-graders (with an average age of 7) and tested for oral and reading skills. Blanco et al. (2001) postulate the process of "equivalent classification" of L2 sounds as L1 sounds, blocking the formation of new L2 categories. The results of their research indicate that the sounds with only slight differences from the L1 categories are the most difficult to acquire, whereas the sounds with no close L1 equivalent are easiest to learn. Kuhl (1993), another researcher investigating the difficulty of acquiring L2 phonology, suggests that beginning at the age of one, L1 phonological prototypes act as perceptive magnets that attract similar L2 sounds.

Blanco et al. (2001) intended to overcome this phonological deficit by teaching explicit and analytical phonology to 2nd grade students. Their research suggests that the capacity to perceive and form new phonological categories is not lost and that training in the perception and articulation of L2 sounds can help learners to form new phonological categories. The authors then go on to discuss the Motor Theory of Speech Perception (Liberman & Mattingly, 1985), which claims a very close connection between the perception and articulation of speech. Proponents of this theory have tried to demonstrate that the basic unit of speech perception is speech articulation, not the acoustic signal (Bowman & Goldstein, 1989).

Blanco et al. (2001) used the aforementioned research by Liberman and Mattingly (1985) and Bowman and Goldstein (1989) to justify designing the program around complimentary components of perceptive and articulatory training. They concentrated on activities that developed the ability to discern rime units of varying complexity (VC, VCC, VCCC) that form parts of the words being analyzed, with the students encouraged to categorize and label the different phonemic units. Articulation was approached as a cumulative process, using the vowel nucleus as the starting point and gradually adding consonants until students could pronounce the entire rime unit correctly. The post-test results of their study confirmed the effectiveness of the experimental training in the phonological skills and also in the areas of reading, word repetition, and spelling. Particularly useful is the premise of Blanco et al. that articulation rather than acoustic signals generates the process of distinguishing and producing new phonetic categories. To a certain extent this mirrors intuitive understanding, as we can often observe everyday examples of L2 users who have the benefit of years or decades of L2 acoustic input yet do not acquire certain L2 phonetic categories.

Method

This section describes the study subjects, elicitation material, pretest procedures, test procedure, and post-test procedures.

Subjects

The eight subjects were native speakers of Spanish or French, languages that lack the aspirated variants of voiceless *stops* that are found in initial position in English.

Elicitation Material

The researcher gave the subjects sentences containing the monosyllabic English words, *pill, tip,* and *cod,* examples of all three types of aspirated voiceless *stops* in initial position. The tokens with initial voiceless *stops* were contained in complete sentences in order to generate natural speech patterns. The tokens, *pill* and *tip,* contain [ɪ], a high vowel, and the token, *cod,* contains [a], a low vowel. Both the vowel height and the tenseness or laxness of the vowel that follows the initial voiceless *stop* may very well affect aspiration. The phoneme [ɪ] is a lax rather than tense vowel, which minimizes obstruction to the airflow; and because it is a high vowel, it is pronounced with the lips relatively close together, which concentrates and maximizes the puff of air produced during aspiration. It might be expected that the low vowel [a] does not maximize aspiration as much as the high vowel [ɪ]. This factor needs to be taken into consideration when analyzing the samples produced by the subjects. The sentences containing the tokens, *pill, tip,* and *cod,* were combined into a list of seven sentences and randomized as follows, with sentence 1 repeated a second time, just before sentence 6:

1. She gets a *pill* everyday. 2. He wants a good *dip*. 3. He's always talking about *cod*.
4. She gets a *bill* everyday. 5. He wants a good *tip*. 6. He's always talking about *God*.

Pretest Procedures

In Pretest Procedure 1, the researcher gave the subjects the list of sentences from the elicitation material. The subjects were asked about any unfamiliar words. The researcher gave the subjects a definition and also told them when they were pronouncing the word correctly. The researcher did not model the words for the subjects but only verified that the subjects understood the vocabulary.

In Pretest Procedure 2, the researcher asked the subjects to read the sentences at a natural rate of speed. The examples were recorded using an Olympus digital voice recorder. The tokens from the pretest recordings were measured for voice onset time (VOT) using spectrograms produced with sound analysis software.

Test Procedure

The researcher conducted a brief lesson that explained the aspiration of initial voiceless *stops.* This lesson served as a means for evaluating the efficacy of phonetic instruction as regards the acquisition of the aspirated variants of voiceless *stops* in English.

Post-test Procedures

The subjects read the elicitation material to generate a second round of recorded samples. The tokens from the post-test recordings were measured for VOT using spectrograms produced with sound analysis software.

Results

The recorded sample from Subject 1 produced a typical set of pretest (see Figure 1) and post-test (see Figure 2) spectrograms for the variant of /p/ which occurs in initial

position. The time sequence is from left to right. Going from left to right, the large dark area starts to cover the upper half of the image about halfway across the spectrogram. This indicates where the sound of the consonant *p* begins. The thick line with graph points begins to the right of this. This indicates where voicing begins, that is to say, where the vowel sound begins. Therefore, the time frame between where the dark area begins and where the thick graph line begins indicates the length of time from the beginning of the voiceless *stop* /p/ to the beginning of the following vowel, or the VOT. This is how the degree of aspiration for initial voiceless *stops* is defined.

Figure 1. Subject 1—Pretest spectrograph

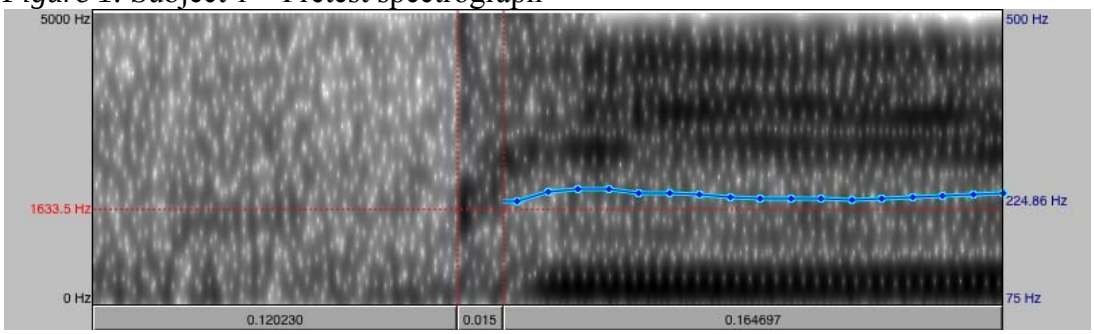
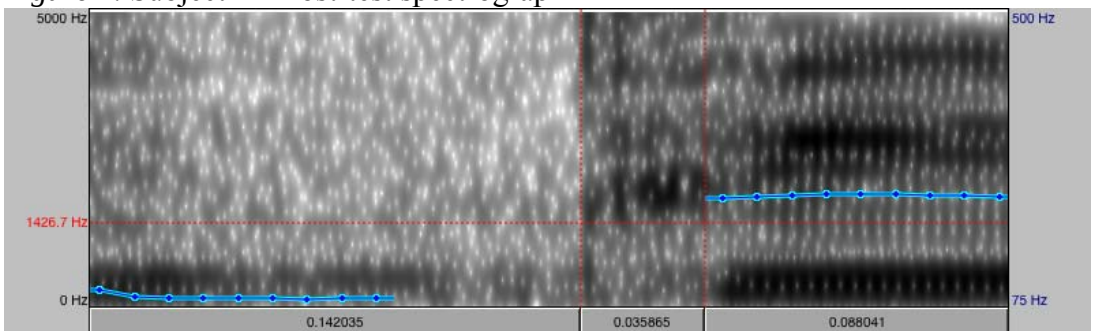


Figure 2. Subject 1—Post-test spectrograph



The number under the small time section in the middle of the spectrograph, then, is the VOT, as measured in milliseconds. The pretest VOT is 15 ms. The post-test VOT is 35 ms. Subject 1 more than doubled the degree of aspiration for the initial *p* variant. Native English speakers in North America exhibit a VOT for initial voiceless *stops* ranging between 20 ms and 60 ms. Listed in the Table 1 are the results from all eight subjects for the

Table 1

Voice Onset Time for Initial P Allophone

	Pretest	Post-test
Subject 1	15	36
Subject 2	0	10
Subject 3	0	10
Subject 4	14	25
Subject 5	15	32
Subject 6	37	54
Subject 7	0	15
Subject 8	10	35

initial *p* variant. The subjects exhibited similar improvement for the initial *t* and initial *k* variants.

Significance

Average voice onset time for speakers in North America is in the range of 20 ms to 60 ms. If we allow that speakers in England might possibly have a voice onset time of up to 100 (twice the North American average), at least for British Received Pronunciation, then we have a maximum scale of no more than 100 for the entire English-speaking world. On a scale of 100, a standard deviation of 10 is used for a small sample such as this. A z-test yields:

$z = 15.75/(10/\text{sqrt}(8)) \approx 4.455$

In a one-way analysis of variance, if the probability of observing a value greater than 4.455 for the standard normal distribution is less than .05, then the results can be considered significant. A rough estimate is that the *p* is less than .01, which would make it very unlikely that the above results would be obtained under the null hypothesis. In this case, then, the null hypothesis can be rejected, although no further conclusions can be made.

Conclusion

Typological markedness is a concept from Universal Grammar, a theory of linguistics postulating principles of grammar shared by all languages. This concept predicts that a less natural form, such as the aspirated *stops* that are found in relatively few languages other than English and are not found in Spanish or French, will be more difficult for L2 learners to acquire. This is especially true because this is a phonetic category that is close to and easily replaced by an already existing L1 category in Spanish and French (unaspirated voiceless *stops).* One measure of markedness is the fact that out of over 200 languages in Europe, only English, German, and Icelandic have aspirated *stops,* although they occur sporadically in languages outside Europe.

The data from the small sample in this study indicates that brief, focused phonetic instruction may be beneficial to L1 Spanish or French speakers in acquiring the aspirated English variants of voiceless *stops.* Study on a larger scale would yield more definitive results. Beyond that, there is potential for the use of explicit instruction in dealing with other problems in phonology which impact L2 learning.

References

Blanco, M., Gayoso, E., & Carrillo, M. (2001). *Analytical training for the teaching of EFL in primary school.* Paper presented at the Phonetics Teaching & Learning Conference 2001, London.

Bowman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Haskins Laboratories Status Report on Speech Research, SR-99-100,* 69-101.

Brecht, R., Davidson, D., & Ginsberg, R. (1995). Predictors of foreign language gain during study abroad. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 37-66). Amsterdam: Benjamins.

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

Flege, J., & Eefting, W. (1988). Imitation of a VOT continuum by native speakers of English and Spanish. *Journal of the Acoustical Society of America, 83,* 729-740.

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

Kuhl, P. K. (1993). Infant speech perception: A window on psycholinguistic development. *International Journal of Psycholinguistics, 9,* 3-56.

Lapkin, S., Hart, D., & Swain, M. (1995). A Canadian interprovincial exchange: Evaluating the linguistic impact of a three-month stay in Quebec. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 67-94). Amsterdam: Benjamins.

Liberman, A. L., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21,* 1-36.

Lord, G. (2000). *The combined benefits of instruction and immersion on L2 pronunciation.* Unpublished manuscript, University of Florida, Gainesville.