

# Measuring Foreign Accent in Spanish: How Much Does VOT Really Matter?

Elena Schoonmaker-Gates

## 1. Introduction

Previous studies have cited voice-onset time (VOT) as one of various cues that contribute to the perception of foreign accent in English and Spanish (González-Bueno, 1997; Riney and Takagi, 1999; Major, 1987; Magen, 1998). In fact, many textbooks and anecdotes on foreign accent cite VOT duration as one of the most problematic aspects of Spanish for language learners whose first language is English (Díaz-Campos, 2004; Flege, 1991; Flege and Eefting, 1988; González-Bueno, 1997; Lord, 2005; Schmidt and Flege, 1996; Zampini, 1998a, 1998b). This being said, little is known about how much VOT duration actually accounts for global foreign accent. The present study uses digital speech manipulation to explore the relationship between VOT and foreign accent in Spanish, investigating the perception of both native and nonnative listeners. Previous research on accent perception has primarily investigated the talker-specific and linguistic characteristics that correlate with native listeners' accent ratings. Probing the foreign accent perception of nonnative listeners has a different goal: to examine what second language (L2) learners perceive as foreign and how they assess certain cues in speech. This endeavor has implications for pedagogy and second language acquisition because it can inform teachers of what learners hear as foreign-accented and can shed light on what they perceive in the L2 speech signal.

Foreign accent is defined by Munro (1998, p. 139) as “nonpathological speech produced by L2 learners that differs in partially systematic ways from the speech characteristics of native speakers.” The present study examines listeners' ratings of native and nonnative speech on a scale of degree of foreign accent. It specifically explores listeners' ratings of utterances that differ solely in the duration of voice-onset time (VOT). Voice-onset time is the period of time between the release of a stop consonant and the following vowel. In English the VOT of voiceless stops /p t k/ is considerably longer than in Spanish. Lisker and Abramson (1964) measured the mean VOT durations of word-initial voiceless stops in English and Spanish and found noteworthy differences. In the read speech of 2 native Spanish and 4 native English speakers, Spanish VOT durations ranged from 4 to 29 milliseconds while English VOTs were between 58 and 80 milliseconds. The investigators also observed differences based on place of stop articulation, with VOTs shortest in bilabial stops and longest in velar stops, with dental stops in between in both languages. Numerous other sources have highlighted similar cross-linguistic phonetic differences between English and Spanish /p t k/ as well (Castañeda, 1986; Cho and Ladefoged, 1999; Lewis, 2001; Lisker and Abramson, 1964; Zampini and Green, 2001).

This paper provides a review of the relevant previous literature on foreign accent perception and voice-onset time, the methodology and research questions, the results and discussion of findings and a conclusion.

## 2. Previous literature

Research on nonnative English has shown that speech with longer, more native-like VOT durations is rated as less foreign-accented than speech with shorter VOTs (Major, 1987; Riney and

---

\* Elon University, egates2@elon.edu. I would like to express my gratitude to Erik Willis, Kimberly Geeslin, Manuel Díaz-Campos, Isabelle Darcy, and the anonymous reviewers for their comments on earlier versions of this paper. I would also like to thank the instructors who allowed me to collect data in their Spanish classrooms and the wonderful staff at the Indiana Statistical Counseling Center for their invaluable help. Any errors remain my own.

Takagi, 1999). These studies unfortunately did not control for other cues associated with foreign accent that may have also been present in speech with longer VOTs, so their conclusions are somewhat limited. The few studies that used manipulation to isolate the effects of VOT on foreign accent in L2 English and L2 Spanish (González-Bueno, 1997; Magen, 1998) used techniques that cut away potentially important pieces of the VOT. The present study improves on past attempts in that it employs a technique that maintains the VOT intact while still modifying duration.

Research on VOT acquisition in Spanish has investigated both native and nonnative perception and production of VOT (Díaz-Campos, 2004; Flege, 1991; Flege and Eefting, 1986, 1988; González-Bueno, 1997; Lord, 2005; Magloire and Green, 1999; Schmidt and Flege, 1996; Zampini, 1998a, 1998b). Past studies on the effects of VOT on foreign accent in English have shown that native listeners rate nonnative speech with longer (more native-like) VOTs as less foreign-accented (Flege and Eefting, 1986, 1987; Major, 1987; Riney and Takagi, 1999). In one of the earliest studies to examine VOT duration in terms of foreign accent perception, Major (1987) measured the VOTs of voiceless stops in nonnative English. The study reported on the assessments of foreign accent given by ten native speakers of English listening to short phrases produced by 53 native Portuguese speakers learning English and seven native speakers of English. Rather than using a categorical rating scale, a sliding lever was used to determine degree of foreign accent, and its continuum went from “no foreign accent at all” to “very heavy foreign accent.” VOT duration values for each speaker were calculated from spectrograms of four isolated words containing instances of the voiceless stops /p t k/ in word-initial position. Statistical analyses revealed a correlation between the mean global accent ratings given by native listeners and speakers’ VOT durations. Specifically, the listeners who were rated as more native-like also spoke with longer (more native-like) VOT durations. The results also showed considerable variation between VOTs that received similar accent ratings, which suggests that listeners were responding to other cues in speech as well.

Riney and Takagi (1999) also investigated voice-onset time duration and its potential correlation with global foreign accent ratings. In an earlier study (Riney and Takagi, 1998) five native English listeners used a 9-point scale to rate five sentences spoken by native Japanese speakers learning English. Listeners rated speech that had been recorded during the speakers’ first and fourth years of college. They also analyzed speakers’ VOT durations from six words with word-initial voiceless stops, finding a positive correlation between the global foreign accent scores given and VOT durations just as Major (1987) had. However, the results showed that three nonnative speakers improved their global foreign accent ratings without improving their VOT, which indicates that further investigation of the effects of VOT on foreign accent perception is necessary.

Two additional studies used digital speech manipulation to tease VOT apart from other cues and to attempt to quantify its effects on foreign accent perception (González-Bueno, 1997; Magen, 1998). González-Bueno (1997) manipulated the VOT duration of the /k/ sound in the word ‘*casa*’ *house* produced by a single native speaker of English learning Spanish. The investigator created a total of 14 stimuli that ranged in VOT duration. She did this by removing 65 ms of the 75 ms VOT then inserting periods of silence in 5 ms increments until the original 75 ms duration was reached. Additionally, she created 14 ‘intact’ stimuli by removing 5 ms increments from the original VOT until only 10 ms remained. All 28 stimuli of the word ‘*casa*’ were played for 18 native Spanish listeners. Participants then rated each stimulus on a 7-point scale that ranged from ‘most native’ to ‘most foreign.’ The results showed that the VOTs between 15 and 35 ms were rated as most native-like regardless of the type of manipulation used. Because this study only examined one instance of velar /k/ spoken by one individual, further research is needed to determine whether these findings can be generalized to other speakers, contexts, and places of articulation.

Magen (1998) also used digital speech manipulation to modify the VOTs in the speech of two native Spanish speakers learning English in the United States. Two L2 learners read 96 English sentences that contained “difficult sounds” for Spanish speakers learning English. A native speaker also read the sentences, and the native VOTs (which contained longer, more native-like VOTs) were used to replace the nonnative VOTs. Ten native English listeners rated 200 natural and VOT-edited stimuli on a 7-point scale of closeness to native English. The results revealed that VOT was not a significant predictor of listeners’ accent ratings of either speaker. Contrary to the other previous findings, these results suggest that VOT may not be a prominent cue to foreign accent for native listeners of English. It is also possible that the manipulation method used to replace nonnative VOTs with native VOTs resulted in speech that did not sound natural or that lacked important transitions.

In sum, previous studies that measured VOT in natural speech found that VOT duration correlates with the perception of foreign accent in L2 English (Major, 1987; Riney and Takagi, 1999), though work that isolated and manipulated VOT did not always find it to be a significant predictor of native listener ratings of foreign accent. The previous work on the contributions of VOT to foreign accent perception in Spanish is also limited in that it only investigated the effects of VOT duration of the velar voiceless /k/ in a single word spoken by one native English speaker (González-Bueno, 1997). Further experimentation with VOT manipulation is needed to isolate it as a cue and to determine whether it contributes significantly to native and nonnative listener perception in Spanish.

The present study aims to add to previous findings by manipulating VOT durations of word-initial bilabial, dental, and velar voiceless stops in Spanish. It uses both native and nonnative speech manipulation, creating stimuli that represent a range of VOT durations that are both longer and shorter than the original speech. Furthermore, this study employs a manipulation process that, unlike past research, does not insert or remove pieces of the original VOT. Using the Pitch Synchronous Overlap-Add tool in Praat, the present study shortens and lengthens segments of speech while keeping the speaker's original VOT relatively intact and natural sounding. Finally, this investigation addresses the effects of VOT duration on nonnative listeners' perception of foreign accent in addition to native listener perception. This is meant to inform pedagogy, shedding light on what language learners hear as foreign-accented and how they assess VOT as a cue to foreign accent. This investigation, made up of two experiments, is guided by the following research questions:

- (a) Do native and nonnative listeners rely on VOT in assessing degree of foreign accent in Spanish?
- (b) How much does VOT contribute to listeners' perception of degree of foreign accent?
- (c) Do nonnative listener ratings of VOT-modified sentences vary according to place of articulation of the stop?

### 3. Experiment 1: Method

Experiment 1 was done to address the first two research questions, regarding native and nonnative listeners' reliance on VOT for assessing foreign accent and the amount that VOT contributes to their perception of degree of foreign accent.

#### 3.1. Listeners

Both native and nonnative Spanish speakers participated in the present study as listeners. The nonnative participants were 160 native English speakers enrolled in undergraduate Spanish courses at a large Midwestern university. Nonnative listeners were recruited from intact third-semester and fifth-semester courses, and from third and fourth year Spanish courses. The ratings of the first 40 from each instructional level were included in the data analysis. Twenty-six native Spanish listeners also participated. These were monolingual Spanish speakers living in central Spain, recruited in public libraries and universities. Most had completed at least some college courses at the time of participation.

#### 3.2. Talkers and Stimuli

Stimuli used in the experiment consisted of sentences read by three native and 12 nonnative speakers of Spanish. The three native Spanish speakers were two females and one male from central Spain who reported varying degrees of fluency in English. All had resided in the United States for at least 3 years prior to the study. The nonnative speakers were 12 L1 English speakers enrolled in undergraduate and graduate-level Spanish courses. These speakers were chosen to represent varying degrees of foreign accent and so were from various instructional levels.

Although stimuli from 15 different speakers were played for listeners, all but four were distracters. The speech of two native Spanish and two nonnative speakers was chosen for digital manipulation and

rating analysis. The native Spanish speakers were chosen because they were females from the same area of central Spain. The nonnative/native English speakers, also both female, were chosen because they had moderate to heavy foreign accents in Spanish. This was determined by a small pilot study conducted with 4 native and 20 nonnative listeners who rated sentences produced by all speakers on a 9-point scale. Although in most cases the 2 nonnative speakers' VOT durations differed by less than 10 milliseconds, in some sentences there was considerable variability between the two speakers. In these sentences the difference was generally between 20 and 30 milliseconds. It is important to note that neither nonnative speaker had uniformly longer or shorter VOT durations than the other, and so their speech was deemed comparable for the purposes of the present study.

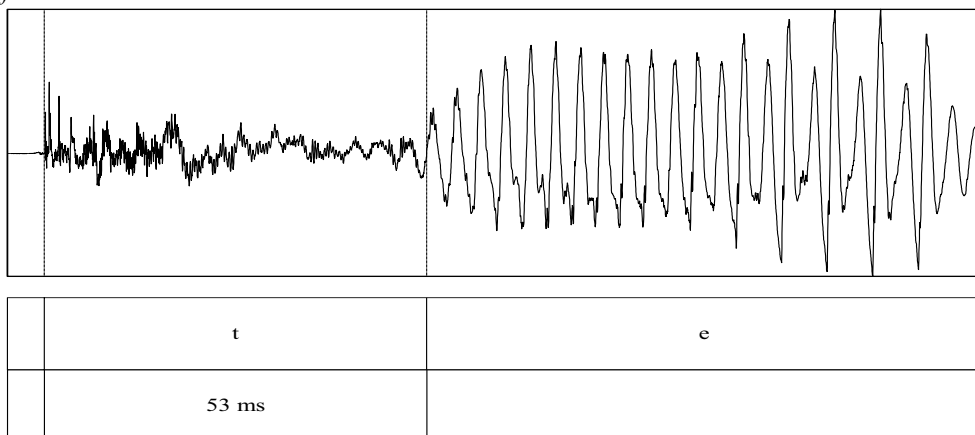
All talkers read a total of ten Spanish sentences that ranged from seven to eleven syllables. Five warm-up sentences were used in a short practice activity and five experimental sentences were used in the perception activity. Each sentence contained three instances of word-initial voiceless stops. Across both sets of five sentences there was one instance of each of the following consonant-vowel combinations: /pi, pe, pa, po, pu, ti, te, ta, to, tu, ki, ke, ka, ko, ku/.

The five experimental sentences were:

1. Él come pavo en su piso. (*He eats turkey in his apartment.*)
2. A ti te doy el queso. (*I give you the cheese.*)
3. Él puso una flor en tu pelo. (*He put a flower in your hair.*)
4. Yo tomé un taxi en Cuba. (*I took a taxi in Cuba.*)
5. Él se quiso casar en el polo sur. (*He wanted to get married in the south pole.*)

In each sentence the three VOTs were measured following Cho and Ladefoged (1999), starting at the onset of the release of the stop closure and ending at the zero crossing of the onset of the first complete vocal fold vibration. An example of a nonnative speaker's pre-manipulation VOT measurement is shown in Figure 1.

Figure 1. *VOT Duration Measurement in Nonnative Sentence 'A ti te doy el queso' Before Modification*



Using the Pitch-Synchronous Overlap-Add (PSOLA) function in Praat, each VOT was made 200% (two times) and 50% (half) the duration of the original. In addition, the native English speakers' VOTs were made 25% (one quarter) the original and the native Spanish speakers' VOTs were made 300% (three times) the original. This was done to create an overlapping category in which native Spanish speech had nonnative VOT durations and nonnative Spanish speech had native VOT durations. The modification procedure itself involved creating a new duration tier with six duration points to isolate the three VOTs. This way only the VOT segments were affected by the duration modifications in the resynthesized stimulus. The PSOLA function was used in order to modify the duration without removing potentially important cues from the VOT. It modifies duration either by removing or reiterating small overlapping portions of the speech signal without affecting pitch or

formant values. The original and modified VOT durations for both native and nonnative speakers are shown in Table 1 below. The overlapping categories are shaded.

Table 1. *Native and Nonnative Speaker VOT Durations With and Without Modification*

#### Nonnative speakers

<b>Voiceless stop</b>	<b>25% VOT</b>	<b>50% VOT</b>	<b>Unmodified</b>	<b>200% VOT</b>
/p/	16 ms	28ms	56ms	109ms
/t/	19 ms	32ms	62ms	121ms
/k/	19 ms	37ms	70ms	136ms

#### Native speakers

<b>Voiceless stop</b>	<b>50% VOT</b>	<b>Unmodified</b>	<b>200% VOT</b>	<b>300% VOT</b>
/p/	10ms	18ms	36ms	53 ms
/t/	12ms	21ms	41ms	62 ms
/k/	17ms	34ms	64ms	96 ms

### 3.3. Procedure

The nonnative listeners completed the experiment on individual computers in a computer lab. Native Spanish listeners completed the experiment individually on a personal laptop computer in a quiet place. All participants used over-the-ear headphones to complete the warm-up and the perception activity.

Listeners completed a 23-item warm-up to become familiarized with the task and the range of accents they would hear. Sentences were presented in pseudorandom order using Praat (Boersma and Weenink, 2012) and listeners had the option to replay each sentence once. They used a 9-point rating scale to rate the degree of foreign accent they perceived in each sentence. The question they were asked was ‘How close to native Spanish is the pronunciation of the sentence’ and the rating scale ranged from 1 to 9, 1 being *closer to native* and 9 being *less close to native*. The rating scale was displayed in Praat and after hearing each stimulus listeners were asked to select a number on the screen and finalize their decision by clicking an OK button that would appear. A screen shot of the rating scale as it appeared with each stimulus is shown in Figure 2 below.

Figure 2. *Screen Shot of Rating Scale for Nonnative Listener*

How close to native Spanish is the pronunciation of the sentence?

1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---

1 = closer to native Spanish      9 = less close to native Spanish

Click to replay

After listeners finalized their rating decision the program automatically played the next stimulus and the process began again. The perception activity consisted of 80 target items (4 Speakers x 5 Sentences x 4 VOT Durations) and 130 distracter sentences for a total of 210 items. Some of the distracter sentences had been modified for speech rate as part of another study. Participants were encouraged to work at their own pace and spent approximately 30 minutes on the task.

In addition to the warm-up and the perception activity, listeners completed a background questionnaire to gather information about their linguistic experience and a grammatical proficiency test. The information gathered was used for statistical analysis.

## 4. Experiment 1: Results

### 4.1. Difference ratings for VOT manipulations

In order to address the contributions of VOT to listener ratings of foreign accent, listeners heard four versions of each sentence spoken by the same speaker: unmodified, 200% VOT, 50% VOT, and an additional overlapping category. If the speaker was native then their VOT was modified to be 300% the original duration and if they were nonnative it was made 25% the original. Rather than analyzing the raw scores given to speech by listeners, the difference between the unmodified version and each of the manipulated versions was calculated instead. This was done to normalize listeners' ratings and to determine the amount of change in foreign accent that listeners heard as a result of VOT duration.

A mixed model ANOVA analyzed native Spanish and nonnative/native English listeners' difference ratings with VOT manipulation as a fixed factor. The analysis revealed VOT to be a significant predictor of both native and nonnative listeners' difference ratings  $F(4, 12636)=198.81$ ,  $p<0.001$ , as shown in Table 2. The negative difference ratings in Table 2 indicate that listeners rated VOT modified speech lower, as less accented, on the foreign accent scale. The positive ratings indicate that modified speech was rated higher, as more accented, on the rating scale.

Table 2. *Native and Nonnative Listener Difference Ratings by VOT manipulation*

	Native listeners	sd	Nonnative listeners	sd
50% VOT	-0.15	1.3	-0.17*	1.4
200% VOT	0.43*	1.3	0.31*	1.3
25% VOT (Nonnative speaker only)	-0.52*	1.5	-0.51*	1.6
300% VOT (Native speaker only)	0.65*	1.4	0.50*	1.4

\* $p<0.001$

Table 2 shows that both native and nonnative listeners rated speech with VOTs that were longer as significantly higher or more foreign-accented on the scale than unmodified speech. Listeners in both groups also rated speech with VOTs that were shorter (25% the original duration) as significantly lower or less accented on the scale than unmodified speech. Additionally, nonnative listeners rated speech with VOTs that were 50% the original duration as significantly less accented than unmodified speech.

### 4.2. Raw ratings of nonnative speech with native-like VOT durations

The raw scores given by listeners to nonnative speech with native-like VOTs were used to examine how much VOT duration accounts for foreign accent in nonnative speech. In order to do this, listeners' ratings of unmodified speech and speech in the overlap category (25% the original VOT duration) were compared. Mean ratings of both native and nonnative Spanish listeners are shown in Table 3.

Table 3. *Native and Nonnative Listener Mean Ratings of Nonnative Speech in Overlap Category (Unmodified and Modified)*

	Native listeners	Difference	Nonnative listeners	Difference
25% VOT in nonnative speech	6.25		6.96	
		+ 0.51		+ 0.51
Unmodified nonnative speech	6.76		7.47	

Table 3 shows that native and nonnative listeners rated nonnative speech with native-like VOTs as less accented than the original speech, a difference of half a point on the overall rating scale. A mixed model ANOVA found speech modification to be a significant predictor of listeners' foreign accent ratings of native listeners  $F(1, 493) = 21.78$ ,  $p<0.001$  and nonnative listeners  $F(1, 3039) = 99.98$ ,  $p<0.001$ . Put otherwise, the same nonnative speech was rated as half a point less foreign-accented when VOTs were native-like in duration, a difference that was significant.

#### 4.3. Raw ratings of native speech with nonnative-like VOT durations

Listeners' ratings were also used to examine how much foreign accent ratings differed when native speech was produced with nonnative-like VOTs. For this analysis, listeners' ratings of unmodified native Spanish and speech in the overlap category (300% the original VOT duration) were compared. Mean ratings of both native and nonnative listeners are shown in Table 4.

Table 4. *Native and Nonnative Listener Mean Ratings of Native Speech in Overlap Category (Unmodified and Modified)*

	Native listeners	Difference	Nonnative listeners	Difference
Unmodified native speech	1.62		1.83	
		+ 0.67		+ 0.50
300% VOT in native speech	2.29		2.33	

Table 4 shows that native Spanish listeners rated native speech that had nonnative VOT durations as 0.67 points more accented than original speech on the rating scale. Nonnative/native English listeners also rated native Spanish with nonnative VOTs as significantly more accented than original speech, a difference of 0.5 points on the rating scale. A mixed model ANOVA revealed a significant effect of these two speech modifications for native listener  $F(1, 493) = 36.68, p < 0.001$  and nonnative listener  $F(1, 3039) = 145.84, p < 0.001$ . These findings indicate that when native Spanish had nonnative VOTs, listeners rated it between 0.5 and 0.67 points more foreign-accented than the same speech with native VOTs.

#### 4.4. Comparison of unmodified speech and speech in the VOT overlap category

Considering rating scale differences between unmodified speech and speech in the VOT overlap category renders a more complete picture of VOT duration as it contributes to foreign accent in Spanish. Table 5 presents the mean ratings given by native and nonnative listeners to unmodified utterances and versions of those same sentences with overlapping VOT durations. Again, the overlapping category specifically refers to the nonnative speech with 25% VOT durations and the native speech in Spanish with 300% VOT durations.

Table 5. *Native and Nonnative Listener Mean Ratings of VOT-Overlap Speech*

	Native listeners	Difference	Nonnative listeners	Difference
Unmodified native speech	1.62		1.83	
		+ 4.63		+ 5.13
25% VOT in nonnative speech	6.25		6.96	
Unmodified nonnative speech	6.76		7.47	
		- 4.47		- 5.14
300% VOT in native speech	2.29		2.33	

Table 5 shows that the nonnative speech with native-like VOTs was still rated by native listeners as 4.63 points higher or more accented than native speech. Similar trends were found for nonnative listeners, who rated the modified nonnative speech as 5.13 points higher or more accented on the 9-point scale. The values in Table 5 also indicate that native listeners rated native speech that had nonnative VOT durations as 4.47 points lower or less accented than unmodified nonnative speech. Nonnative listener means show similar trends, with modified native speech rated as 5.13 points lower than the unmodified nonnative speech. These findings suggest that despite the modifications to VOT duration, nonnative speech continued to sound nonnative and native Spanish continued to sound relatively native.

## 5. Experiment 2: Method

### 5.1. Participants and Stimuli

The second experiment addressed the third research question, regarding whether place of articulation affected nonnative listeners' ratings of VOT-modified speech. Listeners in the second experiment were 74 L1 English speakers enrolled in accelerated first-year Spanish classes designed for students who had taken at least one year of high-school Spanish prior to their enrollment in the class. Most reported at least two years of high-school Spanish and none had participated in Experiment 1. Native Spanish listeners were not included in Experiment 2 because its primary focus was on the effects of place of articulation on language learners' perception.

The same 15 talkers and 10 sentences from the first experiment were used in the second experiment, and the speech of the same four native and nonnative Spanish speakers was manipulated. However, the manipulations differed in the second experiment so that in each utterance only one of the three word-initial VOTs was manipulated. This was done to control for place of the stop articulation. In the second experiment, VOTs for each place of articulation were made 50% (half) and 200% (two times) the original duration. The overlapping (25% and 300%) category from the first experiment was omitted due to time constraints. This was necessary in order to keep the task length reasonable for participants and to minimize their fatigue. Listeners were presented with 6 modified versions of each utterance spoken by the target speakers (3 VOTs x 2 VOT Durations), plus the original unmodified utterance. In total, listeners heard and rated a total of 140 utterances spoken by the 2 native and 2 nonnative speakers (7 Versions x 5 Sentences x 4 Speakers). With 70 distracter sentences added, the perception activity consisted of 210 items presented in pseudorandom order.

### 5.2. Procedure

The procedure of the second experiment was very similar to the first. Listeners performed a warm-up activity consisting of 23 practice items to familiarize them with the task. In this warm-up they were exposed to speech manipulations similar to those they would hear in the actual experiment. Again, the perception activity took listeners approximately 30 minutes to complete. Listeners used the same 9-point rating scale to rate the degree of foreign accent they perceived as was used in the first experiment.

## 6. Experiment 2: Results

The nonnative/native English listeners' mean difference ratings between the unmodified utterances and both modifications were calculated for each place of articulation: bilabial /p/, dental /t/, and velar /k/. Table 6 shows the listeners' mean difference ratings.

Table 6. *Nonnative Listener Difference Ratings of 50% and 200% VOT Durations*

place of articulation	50% VOT	sd	200% VOT	sd
bilabial	-0.13	1.4	+0.07	1.4
dental	-0.22	1.2	+0.07	1.2
velar	-0.16	1.3	+0.05	1.2

Two mixed model ANOVAs tested the effects of place of articulation (segment, in this case) on listeners' difference ratings for 50% VOTs and 200% VOTs separately. The ANOVAs revealed no significant effect of place of articulation in predicting difference ratings for shorter VOTs  $F(2,4364) = 1.71$ ,  $p=0.18$  or for longer VOTs  $F(2,4364) = 0.19$ ,  $p=0.84$ . However, as was found in the first experiment, type of VOT modification (50% and 200%) was a significant predictor of listener ratings for all stops. This indicates that as in the first experiment, these listeners still differentiated between shorter and longer VOTs when they assessed degree of foreign accent.



## 7. Discussion

The first research question asked whether native and nonnative listeners relied on VOT in assessing degree of foreign accent in Spanish. In order to answer this question, the amount that listeners' ratings of each VOT modification differed from the original sentence ratings was calculated and compared. The results revealed that regardless of whether listeners heard native or nonnative speech, both listener groups rated speech with longer VOTs as more foreign-accented and speech with shorter VOTs as less foreign-accented. The statistical analysis showed that VOT duration was a significant predictor of the ratings given by native and nonnative listeners. This suggests that listeners were able to hear and attend to differences in VOT duration when they made their foreign accent assessments. This was true of both native listeners and language learners from a variety of enrollment levels.

The finding that native Spanish listeners used VOT to assess foreign accent in Spanish extends what González-Bueno (1997) found for the effects of VOT duration of velar /k/ on foreign accent perception to other utterances, places of articulation, speakers, and listeners. Most previous research in this area did not isolate cues like VOT by way of speech manipulation, making it difficult to determine beyond correlational data what listeners were attending to in accented speech. However, the present findings strongly suggest that both native and nonnative listeners attend to VOT duration when they assess foreign accent. Because in the present study VOT duration was the only cue that differed between manipulated and original utterances, the rating differences observed indicate that both listener groups used VOT in making their foreign accent judgments. The fact that this was true of both native and nonnative speech further strengthens these claims, since native speech had no other indicators to accent beyond the VOT differences.

The fact that nonnative listeners associated VOT duration with foreign accent also suggests that language learners know that longer Spanish VOTs are less native-like or more accented than shorter VOTs. How they come upon this knowledge, what form it takes, and whether or not they are conscious of this difference was not explored in the present study and so can only be speculated. It may be that nonnative listeners learn about the differences between English and Spanish VOT through explicit or implicit L2 instruction, through L2 input, through their knowledge of English, or through some combination of these factors. Future research of naïve listeners might shed light on this topic.

The second research question asked how much VOT duration contributes to listeners' perception of degree of foreign accent. This was addressed by examining listeners' ratings of native Spanish with nonnative VOT durations (300% the original) and nonnative Spanish with native VOT durations (25% the original). The results showed that both listener groups rated nonnative speech as 0.51 points less foreign-accented on the 9-point scale when VOTs were native-like. Native speech was rated 0.51 and 0.67 points more foreign-accented on the scale when its VOT durations were nonnative-like.

This finding provides a quantitative measure to how much VOT contributes to foreign accent perception in the Spanish of native English speakers. This means that we can tell our students that learning to produce native-like VOTs can positively impact their foreign accent in Spanish, and that it may be beneficial to overtly teach language learners about cross-linguistic differences in VOT. These results also suggest that VOT has a substantial effect on foreign accent. Although the results show that VOT duration alone is not enough for nonnative speech ratings to align with native speech ratings, the differences in ratings between modified and unmodified speech were significant. What the findings from the first two research questions indicate is that VOT duration contributes significantly to foreign accent perception in Spanish, results that go beyond previous findings that relied solely on correlational data.

The third research question asked whether nonnative/native English listener ratings of VOT-modified sentences varied according to the place of articulation of the stop. This was to determine whether certain stops contributed more to the perception of foreign accent observed or not. New versions of each sentence were created that controlled for place of articulation, and these stimuli were rated for degree of foreign accent by nonnative listeners. The findings revealed that listeners did not rate VOT-modified speech differently as a result of the place of articulation of the modified stop. In other words, listener ratings of foreign accent did not vary based on whether the modified stop was bilabial, dental, or velar.

One would have hypothesized that exceptionally long bilabial VOTs would sound more accented than exceptionally long velar stops, since bilabial stops have considerably shorter VOTs than dental or velar stops. However, this was not the case, which was a surprising finding. The fact that place of articulation was not found to influence foreign accent perception may have resulted from the methodology employed, which involved lengthening and shortening VOTs in relation to their initial duration. Even nonnative speakers generally produced their shortest VOTs in bilabials and their longest VOTs in velars, perhaps because this is mirrored in English. As previously mentioned, Lisker and Abramson (1964) showed that although Spanish VOTs ranged from 4-29 ms and English VOTs ranged from 58-80 ms, the shortest in both languages were the bilabials and the longest were the velars. Because this relationship is similar in English and in Spanish, language learners may find differentiating place of articulation to be easier than learning to shorten or lengthen voiceless stop VOTs in general. This finding can provide instructors with valuable information on what to stress in the language classroom in teaching about foreign accent and Spanish pronunciation.

The findings from the three research questions show that although language learners rate longer VOTs in Spanish as less native-like than shorter VOTs, they do so regardless of the place of articulation of the stop. One stop does not account for most or all of the VOT-related foreign accent perceived by these listeners. In general, both native and nonnative listeners heard and attended to VOT duration differences in drawing conclusions about foreign accent in Spanish.

## 8. Conclusion

To conclude, voiceless stop VOT duration plays a role in the foreign accent perception of both native Spanish and nonnative/native English listeners. This shows that listeners use differences in VOT duration when they assess foreign accent, which indicates that in language classrooms we are right to tell students that producing native-like VOTs can reduce the degree of foreign accent they have. Although there are many different cues that contribute to the perception of a foreign accent in Spanish, the present study provides a means of quantifying the degree to which VOT duration of voiceless stops influences foreign accent. Specifically, this investigation presents data to show that native speech with nonnative VOT durations sounds significantly more accented and nonnative speech with native VOT durations sounds significantly less accented. This trend was observed in both the native and the nonnative listener groups. Furthermore, this effect is not created by only one stop, as shown by the place of articulation data that confirm that nonnative listeners cue in to VOT duration regardless of the place of articulation of the stop.

In general, these findings confirm the role of VOT duration as it contributes to foreign accent in Spanish, providing more compelling evidence than previous research provided. They also show that VOT duration is associated with foreign accent in language learners' concept of what constitutes native or foreign-accented Spanish, results that can inform language instructors of what learners hear and attend to in the L2 speech signal.

## References

- Castañeda, María L. (1986). El VOT de las oclusivas sordas y sonoras españolas. *Estudios de Fonética Experimental, II*, 91-110.
- Cho, Taehong & Ladefoged, Peter (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics, 27*, 207-229.
- Díaz-Campos, Manuel. (2004). Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition, 26*, 249-273.
- Elliott, A. Raymond (1995). Field independence/dependence, hemispheric specialization, and attitude in relation to pronunciation accuracy in Spanish as a foreign language. *The Modern Language Journal, 79*(3), 356-371.
- Flege, James E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America, 84*, 70-79.
- Flege, James E. & Eefting, Wieke (1988). Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *Journal of the Acoustical Society of America, 83*(2), 729-740.
- Flege, James E., & Fletcher, Kathryn L. (1992). Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America, 91*, 370-389.

- González-Bueno, Manuela (1997). Voice-onset-time in the perception of foreign accent by native listeners of Spanish. *IRAL: International Review of Applied Linguistics in Language Teaching*, 35(4), 251.
- Lewis, Anthony M. (2001). *Weakening of intervocalic /p, t, k/ in two Spanish dialects: Toward the quantification of lenition processes*. Unpublished Doctoral dissertation, University of Illinois at Urbana-Champaign.
- Lisker, Leigh and Abramson, Arthur S. (1964). A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *Word*, 20, 384-422.
- Lord, Gillian. (2005). (How) Can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania*, 88(3), 557-567.
- Magen, Harriet S. (1998). The perception of foreign-accented speech. *Journal of Phonetics*, 26(4), 381-400.
- Magloire, Joël & Green, Kerry P. (1999). A cross-language comparison of speaking rate effects on the production of voice onset time in English and Spanish. *Phonetica*, 56, 158-185.
- Major, Roy C. (1987). A model for Interlanguage pronunciation. In G. Ioup & S. Weinberger (Eds.), *Interlanguage Phonology: The acquisition of a second language sound system*. Cambridge: Newbury House.
- Munro, Murray J. (1998). The effects of noise on the intelligibility of foreign-accented speech. *Studies in Second Language Acquisition*, 20, 139-154.
- Riney, Timothy J. & Takagi, Naoyuki (1999). Global foreign accent and voice onset time among Japanese EFL speakers. *Language Learning*, 49(2), 275-302.
- Schmidt, Anna M. & Flege, James E. (1996). Speaking rate effects on stops produced by Spanish and English monolinguals and Spanish/English bilinguals. *Phonetica*, 53, 162-179.
- Zampini, Mary L. (1998a). The relationship between the production and perception of L2 Spanish stops. *Texas Papers in Foreign Language Education*, 3, 85-100.
- Zampini Mary L. (1998b). Strategies for the production of Spanish stop consonants by native speakers of English. Paper presented at the AAAL Conference, Seattle, Washington.
- Zampini, Mary L. & Green, Kerry P. (2001). The voicing contrast in English and Spanish: The relationship between perception and production. In J. L. Nicol (Ed.), *One mind, two languages*. Boston: Blackwell.

# Selected Proceedings of the 6th Conference on Laboratory Approaches to Romance Phonology

edited by Erik W. Willis,  
Pedro Martín Butragueño,  
and Esther Herrera Zendejas

Cascadilla Proceedings Project Somerville, MA 2015

## Copyright information

Selected Proceedings of the 6th Conference on Laboratory Approaches to Romance Phonology  
© 2015 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-467-6 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.  
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

## Ordering information

Orders for the library binding edition are handled by Cascadilla Press.  
To place an order, go to [www.lingref.com](http://www.lingref.com) or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA  
phone: 1-617-776-2370, fax: 1-617-776-2271, [sales@cascadilla.com](mailto:sales@cascadilla.com)

## Web access and citation information

This entire proceedings can also be viewed on the web at [www.lingref.com](http://www.lingref.com). Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Schoonmaker-Gates, Elena. 2015. Measuring Foreign Accent in Spanish: How Much Does VOT Really Matter?  
In *Selected Proceedings of the 6th Conference on Laboratory Approaches to Romance Phonology*, ed. Erik W. Willis et al., 95-105. Somerville, MA: Cascadilla Proceedings Project. [www.lingref.com](http://www.lingref.com), document #3195.