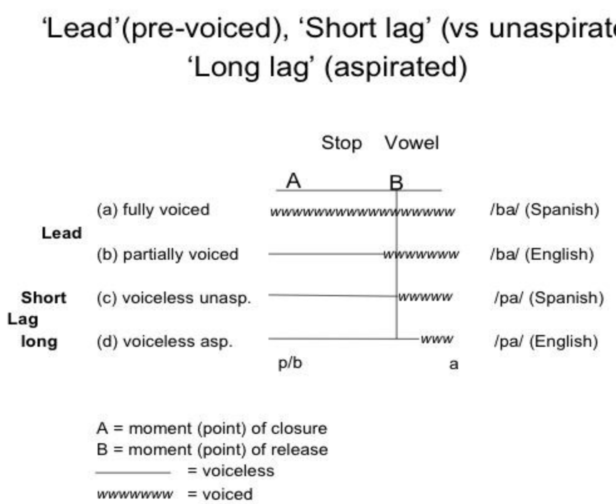
Factors Influencing the VOT of English Long Lag Stops and Interlanguage Phonology

Mehmet Yavaş

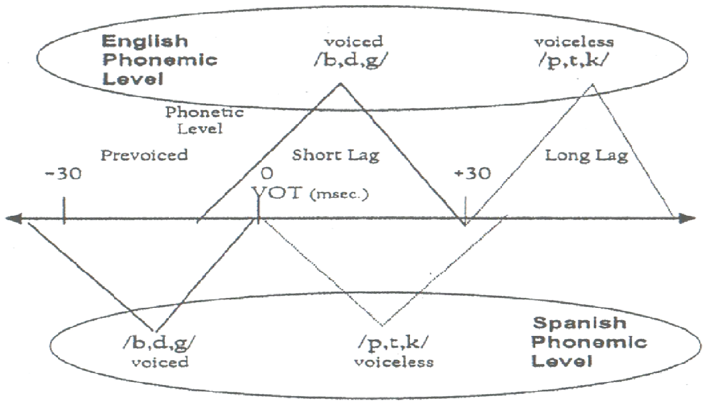
Florida International University, United States
yavasm@fiu.edu

1. Introduction

Stops /p, b, t, d, k, g/ are among the most common sounds in languages. However, their phonetic realizations vary from language to language. Both English and Spanish contain these six stops in three places of articulation which are distinguished by voicing. (English: 'peak' – 'beak', 'tin' – 'din', 'curl' – 'girl'; Spanish: 'peso' – 'beso', 'teja' – 'deja', 'callo' – 'gallo'). Despite this phonological similarity, the phonetic productions of the voiced and voiceless phonemes are rather different in the two languages. The best account of these differences is given via the acoustic measurements known as VOT (voice onset time). VOT refers the time that elapses between the release of the articulators for a stop and the onset of vocal cord vibration of the following segment. If the voicing starts before the release (i.e., during the closure phase) of the stop, then the result is described as 'voice lead' (or 'prevoiced) and is given a negative VOT value. If the voicing starts after the release of the stop, then the result is 'voice lag' and is described with a positive VOT value. The amount of lag is important to separate voiceless unaspirated ('short lag', with VOT values less than 30-35 milliseconds) from voiceless aspirated ('long lag', with VOT values greater than 35 milliseconds).



In Spanish, the voiceless stops /p, t, k/ are produced with a short lag, while the voiced stops /b, d, g/ are produced with a voice lag (vocal cord vibration preceding the stop release). English /p, t, k/ are produced with a long lag; /b, d, g/, on the other hand, may be produced with a short lead, no voicing, or may even have a short lag. The differences between the two languages are well described in the following figure adopted from Zampini & Green (2001).



Phonetic category classification of English and Spanish stop consonants

Zampini, M.L. & Green, K.P. (2001)

Several studies have shown that Spanish speakers experience difficulties in acquiring English long lag stops, because there is, unlike in the voiced stops, no overlap between the two languages (Avery & Ehrlich, 1992; Kenworthy, 1987; MacDonald, 1987; Nash, 1977). Typically, we observe the unaspirated [p, t, k] for the expected [p^h, t^h, k^h].

It has been suggested that two variables – the place of articulation of the target voiceless stop, and the height of the vowel following the target stop – can have significant effects in the amount of long lag (aspiration) of the stop. In relation to the former, it has been stated that the lag is longer as we move the place of articulation from front to back (i.e., from bilabial to alveolar and then to velar). Evidence for this in English can be found in Klatt (1975), Laeufer (1996), Lisker and Abramson (1967), Macken and Barton (1979), Port and Rotunno (1979), Thornburg and Ryalls (1998), Volaitis and Miller (1992), Yavaş (2002), Yavaş and Wildermuth (2006), and Zlatin (1974). The rationale cited for this comes from the degree of abrubtness of the pressure drop upon the release of the stop. The more sudden (abrupt) the pressure drop is, the sooner the voicing of the next segment starts. Consequently, this results in less aspiration (i.e., shorter lag). When we look at stops at the three places of articulation, we see that the tongue dorsum separates more slowly (less abrupt) from the velum for /k/ than the tip from the alveolar ridge for /t/, or from the lips for /p/.

In addition to the place of articulation of the stop, some studies suggest the effect of the height of the following vowel. Specifically, greater lag was observed when stops were followed by high vowels (more open articulations) than when they are followed by low vowels (narrower opening) (Klatt, 1975; Thurnburg & Ryalls, 1998; Yavaş, 2002; Yavaş & Wildermuth, 2006). The rationale for this effect again is related to the abruptness of the pressure drop. High vowels have a more obstructed cavity than low vowels. Since the high tongue position that is assumed during the stop closure in anticipation of a subsequent high vowel would result in a less abrupt pressure drop, a stop produced as such will have longer lag than the one produced before a low vowel.

2. The study

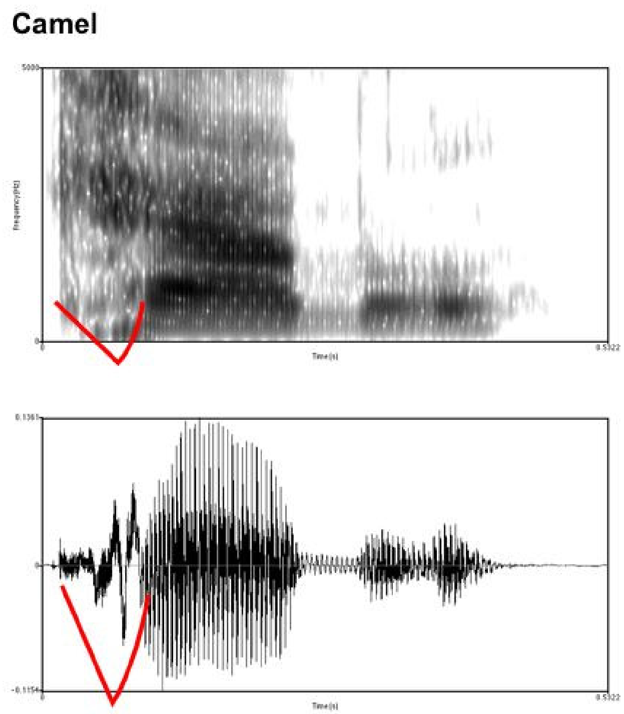
The purpose of this study was to revisit the effects of these two variables in the amount of aspiration. Specifically, it was intended for the validation of Yavaş and Wildermuth (2006) with further data. The two hypotheses that are tested can be stated in the following:
1) VOT will increase as we move the place of articulation from front to back. Thus, should have the less difficulty with the bilabial /p/, then the alveolar /t/, and lastly with the velar /k/.
2) VOT will increase when the stop target occurs before high vowels than before low vowels.

2.1 Subjects

The subjects were 16 adult students, 9 male and 7 female, of the English Language Institute, a private English Language School on the campus of Florida International University in Miami, Florida. All students were native speakers of Spanish with the following countries of origin and numbers: Cuba (3), Colombia (5), Venezuela (4), Costa Rica (2), Nicaragua (1), and Peru (1). None were married to or living with native speakers of English. The ages of the subjects were between 19 and 37; each subject has been in the U.S. five years or less.

2.2 Method

A list of 12 disyllabic target words was placed in carrier sentences. Each of the words contained a voiceless stop in initial position of a stressed syllable. The targets were preceded by a semantically (but not phonetically) close word in the sentence (e.g., "Don't say *cushion,* say *pillow* instead") whereby 'pillow' was the target and 'cushion' was the semantically close word. The 12 targets revealed 6 combinations, with 2 words for each. The combinations were comprised of each of the three places of articulation with two vowel heights: 'bilabial stop + high V' (*picture, pillow*), 'bilabial stop + low V' (*puppy*, *package*) 'alveolar stop + high V' (*timber, ticket*), 'alveolar stop + low V' (*tanker, tablet*), 'velar stop + high V' (*kitten, kicker*), and 'velar stop + low V' (*caller, camel)*. The subjects were asked to read the sentences which were randomly ordered. 12 target words from 16 subjects resulted in 192 tokens. The utterances were acoustically analyzed via PRAAT. The VOT values of the target stops were obtained from the waveform and verified with the spectrogram. The beginning of the lag (positive VOT) was identified by a sharp spike where the waveform changes from quiescent to transient; the end point (onset of vocal cord vibration) is determined from where the waveform becomes periodic. As for the spectrographic readings, VOT intervals from the beginning of the release burst to the onset of voicing were analyzed. The energy burst represents the release of the articulatory constriction, and the first of the regularly spaced vertical striations represents the vocal cord vibration.



VOT measurements were rounded to the nearest ms. For inter-judge reliability, 10% of the tokens were analyzed using the same experimental protocol by two individuals. The average difference between the original measurements and those of the reliability was less than 10 ms.

3. Results and discussion

After measuring the VOT values from each token for the 16 subjects, we calculated the mean VOT per speaker for the six different combinations (3 places of articulation with 2 vowel heights). Table 1 gives the averages subject by subject.

		B	A	V
S.1	H.V.	28	45	54
	L.V.	20	46	53
S.2		34	48	60
		22	43	53
S.3		41	51	67
		24	41	57
S.4		20	39	49
		11	41	40
S.5		25	41	48
		19	42	40
S.6		24	40	59
		17	41	41
S.7		40	55	52
		16	52	39
S.8		32	46	54
		23	42	35
S.9		40	66	60
		24	51	55
S.10		28	49	70
		15	41	55
S.11		36	53	55
		19	40	49
S.12		27	39	71
		20	40	58
S.13		39	46	61
		15	47	50
S.14		28	42	56
		17	34	50
S.15		39	64	59
		17	48	47
S.16		37	46	56
		20	44	49

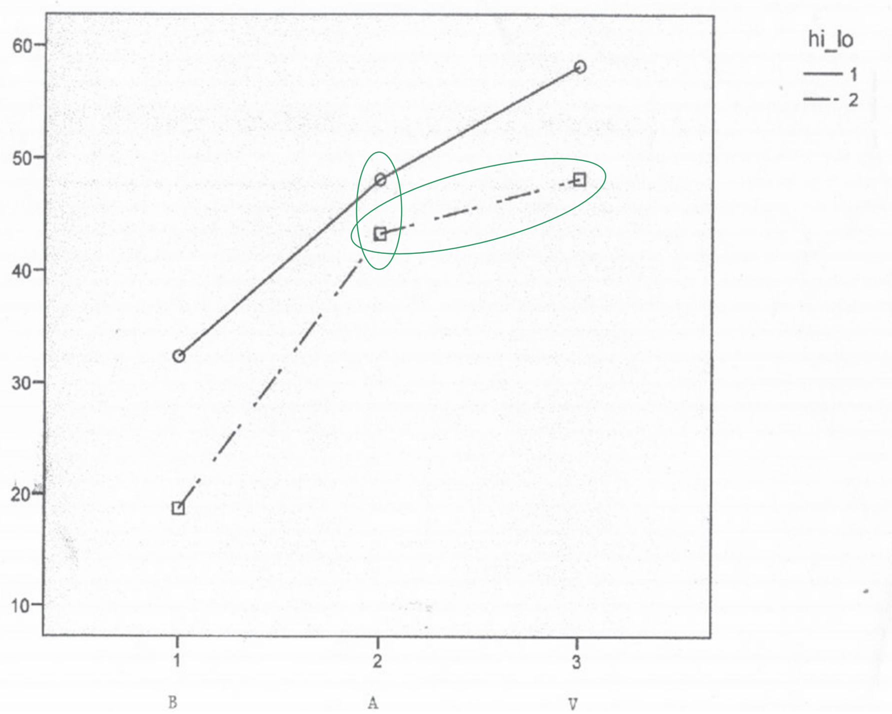
As we note here, the first hypothesis (greater lag as the place of articulation moves from front to back) was negated by 5 subjects in alveolar-velar dimension with low vowels (S. 4, 5, 6, 8, 15). The same hypothesis was negated by 3 subjects (S. 7, 9, 15) in alveolar-velar dimension with high vowels. No subject negated this hypothesis in bilabial-alveolar dimension. The second hypothesis (greater lag is expected when the target stop followed by a high vowel than when followed by a low vowel) was negated by 5 subjects (S. 4, 5, 6, 12, 13) only in alveolars. No subject negated this hypothesis in bilabials or in velars. 6 subjects (S. 2, 3, 10, 11, 14, 16) confirmed both hypotheses without exceptions.

The group results are given in the following display.

Means (S.D.)

	/p/	/t/	/k/
H.V.	32.3 (6.7)	48.1 (8.1)	58.2 (6.7)
L.V.	18.7 (3.6)	43.3 (6.5)	48.2 (7.1)

These are repeated in the graph below.



As can be seen from the above, with the exception of the circled pair-wise comparisons (i.e., 'alveolar + high V' vs. 'alveolar + low V, and 'alveolar + low V vs. 'velar + low V') all other comparisons resulted in significant differences. That is, changing the vowel height made a significant difference in bilabials and velars, and changing the place of articulation resulted in significantly different VOT in all combinations, except for between alveolars and velars when the following vowel was low.

We also wanted to compare the strength of the two variables – place of articulation of the target stop and the height of the following vowel – since the results showed the influential role of both variables in determining the VOT. To make a comparison of strength between these two variables, we compared two pairs of contexts in which one member in each pair has the combination of a favorable place of articulation with an unfavorable vowel height. More specifically, we compared the targets with a) 'bilabial stop + high V' with 'alveolar stop + low V', and b) 'alveolar stop + high V' with 'velar stop + low V'. The results of these comparisons revealed that in comparison a) targets with 'alveolar + low V' had significantly greater mean VOT (m=43.31 ms) than 'bilabial stop + high V' (m=32.38 ms). This suggests the strength of the place of articulation over vowel height. The results of the other comparison (i.e., 'alveolar stop + high V' versus 'velar stop + low V'), however, do not show any significant difference (m= 48.13 ms , and m= 48.19 ms, respectively).

The results of this experimental study confirm partially the influences of the two variables in changing the lag values of /p, t, k/ and thus validate the results of Yavaş and Wildermuth (2006). Hypothesis 1 (i.e., changing the place of articulation) was confirmed in all pairs except between alveolars and velars when preceding low vowels. Hypothesis 2 (i.e., changing the vowel height) was confirmed in bilabials and velars, but was not confirmed in alveolars. As for the relative strength of the variables, our results suggest the strength of the place of articulation over the height of the following vowel from bilabial to alveolar; the same cannot be said from alveolar to velar.

This study demonstrates the fact that in dealing with interlanguage phonology, we need to go beyond simple contrastive phonological information and examine the relative markedness of variables that are speech aerodynamics. Acquiring competency in the sound system of a second/foreign language is a highly structured process. Identifying the linguistic contexts that are influential in promoting or inhibiting the longer lag of target stops provides valuable insights to professionals who deal with teaching/remediation of sound patterns.

References

Avery, P., & Ehrlich, S. (1992). *Teaching American English pronunciation*. Oxford: Oxford University Press.

Kenworthy, J. (1987). *Teaching English pronunciation*. London: Longman.

Klatt, D. H. (1975). Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research, 18*, 686-706.

Laeufer, C. (1996). The acquisition of complex phonological contrast: Voice timing patterns of English initial stops by native French speakers. *Phoneticam, 53*, 86-110.

Lisker, L., & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. *Language and Speech, 10*, 1-28.

MacDonald, M. G. (1987). The influence of Spanish phonology on the English spoken by United States Hispanics. In P. Bjarkman, & R. Hammond (Eds.), *American Spanish pronunciation: Theoretical and applied perspectives* (pp. 215-236). Washington, D.C.: Georgetown University Press.

Macken, M., & Barton, D. (1979). The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language, 7*, 41-74.

Nash, R. (1977). *Comparing English and Spanish*. New York: Regents.

Port, R. F., & Rotunno, R. (1979). Relation between voice-onset time and vowel duration. *Journal of the Acoustical Society of America, 66*, 654-662.

Thornburg, D., & Ryalls, J. H. (1998). Voice onset time in Spanish-English bilinguals: Early versus late learners of English. *Journal of Communication Disorders, 31*, 215-229.

Volaitis, L. E., & Miller, J. L. (1992). Phonetic prototypes: Influence of place of Articulation and speaking rate on the internal structure of voicing categories. *Journal of the Acoustical Society of America, 92*, 723-735.

Yavaş, M. (2002). Voice onset time patterns in bilingual phonological development. In F. Windsor, M.L. Kelly, & N. Hewlett (Eds.), *Investigations in clinical phonetics and linguistics* (pp. 341-350). Mahway, NJ: Lawrence Erlbaum Associates.

Yavaş, M., & Wildermuth, R. (2006). The effects of place of articulation and vowel height in the acquisition of English aspirated stops by Spanish speakers. *IRAL, 44*, 251-263.

Zampini, M. L., & Green, K. P. (2001). The voicing contrast in English and Spanish: The relationship between perception and production. In J. L. Nicol (Ed.), *One mind, two languages* (pp. 23-48). Boston: Blackwell.

Zlatin, M. A. (1974). Voicing contrast: Perceptual and productive voice onset time characteristics of adults. *Journal of the Acoustical Society of America*, *56*, 981-994.