

The effects of stress/accent on VOT depend on language (English, Spanish), consonant (/d/, /t/) and linguistic experience (monolinguals, bilinguals)

Miquel Simonet, Joseph V. Casillas, Yamile Díaz

Department of Spanish & Portuguese
University of Arizona, Tucson, Arizona U.S.A.

{simonet, jvcasill, ydiaz44}@email.arizona.edu

Abstract

This study examines Voice Onset Times of coronal stops in utterance-initial position in two languages. Crucially, the effects of lexical stress (stressed, unstressed syllable) on VOT are analyzed. The study investigates aspirated stops (English /t/), short-lag voiceless stops (English /d/, Spanish /t/) and prevoiced stops (Spanish /d/). Three groups of speakers provide data: English monolinguals, Spanish monolinguals, and proficient Spanish-English bilinguals. The study finds that lexical stress lengthens aspiration (English /t/) and prevoicing (Spanish /d/) but it does not alter significantly short-lag stops (Spanish /t/, English /d/). Monolinguals and bilinguals differ slightly in their phonetic behavior. Implications for gestural coordination as well as for feature theory are discussed.

Index Terms: VOT, stress, Spanish, English, bilingualism

1. Introduction

The present study is concerned with the effects of lexical stress on Voice Onset Times (VOT) of /d/ and /t/ in two languages, Spanish and English. The study examines the productions of monolingual and bilingual speakers and assesses the potential impact of linguistic experience on lexical stress effects for Spanish and English /d/ and /t/.

While both Spanish and English possess a contrast between *fortis* (/p t k/) and *lenis* stops (/b d g/), the phonetic implementation of the contrast differs for the two languages. English presents a contrast between a set of aspirated, voiceless stops (/p t k/) and a set of unaspirated, voiceless stops (/b d g/). The Spanish contrast is that between a set of unaspirated, voiceless stops (/p t k/) and a set of voiced stops (/b d g/) [1]. (Note that this description applies to utterance-initial position because Spanish /b d g/ are regularly spirantized in other positions). According to one view, the phonological opposition in English depends on the feature [spread glottis], which is specified for the *fortis* stops but not for the *lenis* ones; on the other hand, the Spanish contrast depends on [voice], which is specified for the *lenis* stops but not for the *fortis* ones [2].

A large body of research has demonstrated that prosodic structure affects fine-phonetic detail in segments. Experimental studies have shown that consonants in prominent syllables (stressed, accented) are hyperarticulated relative to those in non-prominent syllables (unstressed, unaccented) in a number of languages ([3–7, among others]). Hyperarticulation is captured by articulatory and acoustic measurements. One of the acoustic correlates that manifest hyperarticulation is VOT. In a seminal study, [7] found that VOTs in English aspirated stops (/p t k/) were longer when these consonants were found at the onset of lexically stressed syllables than when they were found at the on-

set of unstressed syllables. This effect was corroborated, for this language, in [3] and, for phrase-accented syllables in connected speech, in [5]. In sum, for consonants with long-lag VOT, such as English /p t k/, the effects of lexical stress are straightforward, and they seem to be of an additive nature—stress lengthens the aspiration period of aspirated consonants.

Stress effects also have been investigated for consonants with short-lag VOTs, although the results are less straightforward than those for aspirated consonants [8–10]. For instance, [7] reported a trend for English /b d g/ according to which these consonants showed a *shorter* VOT when stressed than when unstressed. On the other hand, [5]’s findings are not in line with those in [7] regarding English /b d g/. In [5]’s study, VOT is lengthened when the consonant bears prominence, and that applies to both /p t k/ (English aspirated stops) as well as to /b d g/ (English short-lag, unaspirated stops). The results of a study of Dutch *fortis* (/t/) and *lenis* (/d/) stops are also informative in this regard [4]. Dutch presents a *fortis-lenis* contrast of the “Spanish type”, rather than the “English type”; that is, Dutch /t/ is a voiceless, unaspirated stop—a stop with a short-lag VOT. [4] explores the effects of lexical stress and phrasal accent separately. In this study, VOT length of Dutch /t/ is longest when unstressed and unaccented and shortest when stressed and accented. Thus, the effects of stress on this consonant are subtractive (in terms of VOT), rather than additive, and the effects of lexical stress and phrasal accent are cumulative.

No study seems to have straightforwardly addressed the effects of lexical stress on negative VOT or prevoiced stops [8, 9, cf.]. [4] analyzed the duration of voicing during closure in Dutch /d/. In the data in this study, /d/ always appeared in intervocalic, rather than utterance-initial, position. Although this measurement might differ from lead VOT (since voicing during closure in intervocalic /d/ could be due partially to carryover laryngeal coarticulation), the findings in [4] with respect to /d/ are perhaps relevant—lexical stress was found to affect voicing length. In particular, prominent syllables displayed longer voicing periods during closure than unstressed syllables. The fact remains that, to our knowledge, the effects of stress on prevoiced stops remains to be investigated in depth.

The present study analyzes lexical stress effects on long-lag stops (English /t/), short-lag stops (English /d/, Spanish /t/) and prevoiced stops (Spanish /d/). Thus, we explore a language with a [spread glottis] contrast (English), one with a [voice] contrast (Spanish), and, importantly, we study speakers possessing both features and thus all three stop types (Spanish-English bilinguals). The goal of the present study is to gain a broader understanding of stress effects on different VOT categories in order to begin to grasp the impact of prosody on segments more generally.

2. Methods

In order to collect the production data, we used a delayed shadowing task, widely used in the literature on second-language speech learning (e.g., [11]). In this task, speakers listen to and then repeat out loud target phrases.

2.1. Materials

2.1.1. Target phrases

The materials were a list of words beginning with /d/ or /t/. Half of these words were stressed on the word-initial syllable, and the other half were stressed on the second syllable. Therefore, stops could appear in a stressed syllable or unstressed syllable. There were 40 words total, 20 per language. The words were balanced for Consonant (/d/, /t/) and Stress (stressed, unstressed). There were five word items per design cell: 2 (languages) \times 2 (consonants) \times 2 (stress configurations) = 8 design cells \times 5 items = 40 words. The target words were interspersed among many fillers or distractors ($n = 36$).

2.1.2. Auditory stimuli

The 40 words were printed out on two lists as a function of language—20 English and 20 Spanish words. Three male native speakers of each language (six ‘talkers’ in total) were asked to read their corresponding word list out loud. Each ‘talker’ received a different randomization of the list. Words were produced in utterance-initial position in the carrier sentences ‘... is the word’ and ‘... es la palabra.’ In this way, three auditory models were recorded for each word item, one per ‘talker.’ This amounted to 120 (20 items \times 2 languages \times 3 auditory models or ‘talkers’) different auditory stimuli to be used in the delayed shadowing task.

2.2. Speakers

A total of 47 volunteers participated in this experiment. All speakers were female, and they were all between 18 and 23 years of age. There were three groups of speakers: (i) a group of Spanish-English bilinguals ($n = 19$), (ii) a group of monolingual Spanish speakers ($n = 22$), and (iii) a group of monolingual English speakers ($n = 7$). The monolingual Spanish speakers were recruited from among the student body of the Universitat de les Illes Balears on Majorca, Spain. These participants are bilingual in Catalan. (Note that there are no reported differences between Catalan and Spanish stop consonants with respect to their VOT [12].) The monolingual English speakers, as well as the Spanish-English bilinguals, are/were undergraduate students at the University of Arizona in Tucson, Arizona. The English speakers consider English to be their native language, and they were exposed to English exclusively in their family circle while growing up. The bilinguals, Mexican-Americans born and raised in Southern Arizona, were brought up by Spanish-speaking families and were schooled mostly in English. They use both languages daily both in the classroom as well as with their friends and relatives. A bilingual profile questionnaire was used to establish the groups [13]. In this report, data from the three groups are analyzed separately.

The Spanish-English bilinguals were recorded in both of their languages. The English monolinguals were recorded in English. The Spanish monolinguals were recorded in Spanish.

2.3. Procedure

The 47 female speakers were asked to listen to and then repeat out loud the auditory stimuli. The English monolinguals heard and produced only the English materials (60 tokens). The Spanish monolinguals heard and produced only the Spanish materials (60 tokens). The Spanish-English bilinguals, however, were recorded in both of their languages—they listened to and repeated the English as well as the Spanish materials (120 tokens). Each speaker heard a different randomization of the stimuli.

Importantly, the bilinguals heard and produced the English and Spanish materials in a single block, in random order. In other words, the Spanish and English auditory stimuli were randomized in one single production session so that the English items were interspersed amongst the Spanish items. This hypothetically maximizes the degree of interlingual transfer of phonetic characteristics for these bilinguals [14], and it thus explores this population in a situation diametrically opposed to the one researched in [15]. Speakers were asked to listen to the entire utterance first (‘... is the word’ or/and ‘... es la palabra’) and then to repeat the entire sequence aloud. In this way, shadowing of the target consonant was slightly delayed relative to the time in which it was heard.

Recordings were made through a head-mounted dynamic microphone (Shure SM10A), a pre-amp (Sound Devices MM-1) and a digital voice recorder (Marantz PMD660). Digitization was at 44.1 kHz, 16-bit. The Spanish monolinguals were recorded in a quiet laboratory at the Universitat de les Illes Balears. The English monolinguals and the Spanish-English bilinguals were recorded in a sound-attenuated booth at the Applied Phonetics Laboratory of the University of Arizona. Speakers were recorded one at a time.

2.4. Analysis

A total of 4,020 word tokens were collected for the present study. The English speakers provided 420 tokens, the Spanish speakers provided 1,320 tokens, and the Spanish-English bilinguals provided 2,280 tokens, 1,140 in Spanish and 1,140 in English. The study considered two within-speaker factors: (i) Consonant: /d/, /t/; and (ii) Stress: first syllable in word is either stressed or unstressed. For one of the speaker groups (Spanish-English bilinguals) a third within-speaker factor was added—Language spoken: Spanish, English.

The target /d/ and /t/ tokens were measured for VOT [16]. VOT is an acoustic metric that captures, for stops, the time lag between the burst (i.e., the release of the articulators after the stop closure) and the initiation of periodicity (i.e., the onset of modal voicing). VOT is positive if the burst occurs before the onset of periodicity (since burst = 0 ms) and negative if it occurs after it. In this study, bursts and voicing onsets were hand-located by exploring spectrographic and sound wave displays. The onset of voicing was marked at a zero-crossing on a sound wave display. Voicing onset in lag VOTs was marked by taking into account the initiation of F_2 in order to make sure that we labeled the initiation of modal voicing rather than voicing *per se*.

3. Results

3.1. Spanish speakers

The Spanish materials produced by the Spanish speakers were analyzed through a repeated-measures ANOVA with VOT (ms) as the dependent variable and Consonant (/t/, /d/) and Stress

	/d/	/t/
stressed	-69.2 (3.52)	15.3 (1.04)
unstressed	-50.4 (3.51)	16.3 (1.38)

Table 1: Mean (+ SE) VOT (ms) as a function of Consonant (/t/, /d/) and Stress (stressed, unstressed) in the Spanish data.

	/d/	/t/
stressed	21.9 (3.66)	78.4 (6.72)
unstressed	26.3 (4.81)	71.4 (5.89)

Table 2: Mean (+ SE) VOT (ms) as a function of Consonant (/t/, /d/) and Stress (stressed, unstressed) in the English data.

(stressed, unstressed) as within-subjects factors. Individual speaker ($n = 22$) was the error term. The descriptive statistics of these data are shown in Table 1. The statistical test revealed significant effects of Consonant ($F(1, 21) = 511.7, p < 0.001$) and Stress ($F(1, 21) = 123, p < 0.001$), as well as a two-way significant interaction ($F(1, 21) = 135.7, p < 0.001$). For these Spanish speakers, /d/ has a lead, negative VOT with a mean of -59.8 ms and /t/ has a short-lag, positive VOT with a mean of 15.8 ms.

In order to explore the interaction, the main effects of Stress were investigated for the two stop consonants separately; this was done by means of paired t-tests. The alpha level was adjusted accordingly ($0.05/2 = 0.025$). On the one hand, /d/ was shown to present longer prevoicing when stressed than when unstressed by a factor of 1.373 ($t(21) = -12.02, p < 0.001$). On the other hand, /t/ was found not to be affected by Stress ($t(21) = -1.85, p > 0.05$).

3.2. English speakers

The English VOT data were explored by means of a repeated-measures ANOVA with Consonant (/t/, /d/) and Stress (stressed, unstressed) as main factors. Subject ($n = 7$) was the Error term. The ANOVA yielded significant Consonant effects ($F(1, 6) = 244.9, p < 0.001$), but no significant effects of Stress ($F(1, 6) = 1.27, p > 0.05$). The model revealed a significant two-way interaction ($F(1, 6) = 19.22, p < 0.01$). Descriptive statistics are provided in Table 2.

The interaction was examined by means of two separate paired t-tests, and the alpha level was adjusted accordingly. The interaction was due to the fact that Stress effects were robust for /t/ ($t(6) = 3.8, p < 0.01$), but not for /d/ ($t(6) = -2.7, p = 0.03$). At most, one could claim that a marginal trend exists for /d/ as well, although in the opposite direction. The effects for /t/ were due to the fact that stressed /t/ displays a longer period of aspiration than unstressed /t/ by a factor of 1.098.

3.3. Spanish-English bilinguals

The bilingual VOT data were submitted to a repeated-measures ANOVA with Language (Spanish, English), Consonant (/t/, /d/) and Lexical Stress (stressed, unstressed) as fixed factors. All these were within-subjects factors since bilinguals were recorded in both languages. Subject ($n = 19$) was the error term. The statistical model yielded significant effects of all three main effects: Language ($F(1, 18) = 249.5, p < 0.001$), Consonant ($F(1, 18) = 152.4, p < 0.001$), and Stress ($F(1, 18) = 7.58, p = 0.01$). More relevant was the fact that the model revealed three significant two-way interactions (Con-

Figure 1: Effects of lexical stress on VOT in English /t/ and /d/ in the bilingual speakers' productions.

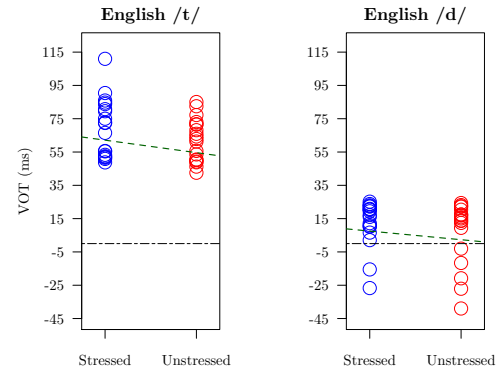
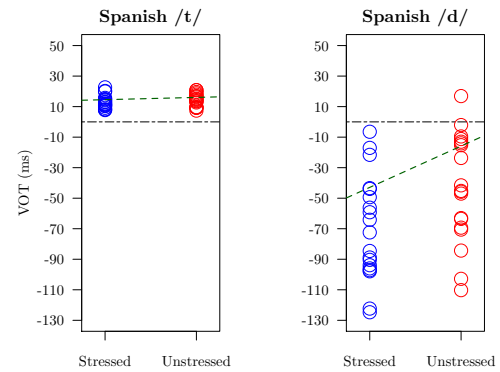


Figure 2: Effects of lexical stress on VOT in Spanish /t/ and /d/ in the bilingual speakers' productions.



sonant by Language, $F(1, 18) = 5.3, p = 0.03$; Language by Stress, $F(1, 18) = 54.5, p < 0.001$; Consonant by Stress, $F(1, 18) = 16.95, p < 0.001$ and a significant three-way interaction ($F(1, 18) = 31.6, p < 0.001$). In order to explore the interactions, the dataset was divided into subsets.

Firstly, the effects of Consonant and Stress were examined for the two languages separately. The first ANOVA investigated the effects of these two fixed, within-subjects factors on VOT for the English data only. These data are plotted in Figure 1. This model yielded significant effects of Consonant ($F(1, 18) = 140, p < 0.001$), as well as Stress ($F(1, 18) = 17.3, p < 0.001$). Importantly, there was no interaction between the two factors ($F(1, 18) < 1$). The effects of Stress were due to the fact that stressed consonants had a longer VOT (a VOT further away from zero) than unstressed consonants by a factor of 1.18. Note that both consonants (/t/, /d/) were similarly affected by the effects of stress in this data subset.

The second follow-up ANOVA explored the effects of Consonant and Stress for the Spanish materials only. (See Figure 2.) In this analysis, Consonant and Stress were within-subjects factors and speaker was the error term. The model revealed significant effects of Consonant ($F(1, 18) = 99.25, p < 0.001$) and Stress ($F(1, 18) = 35.38, p < 0.001$), as well as a significant two-way interaction ($F(1, 18) = 28.42, p < 0.001$). The interaction was due to the fact that /d/ was affected by Stress ($t(18) = -5.73, p < 0.001$) while /t/ was not ($t(18) = -1.6, p > 0.1$). In these data, stressed /d/ had longer prevoicing than unstressed /d/ by a factor of 1.63.

4. Discussion and conclusions

The present study has reported on the results of a production experiment in which the effects of lexical stress on VOT length are examined. VOTs were measured for /d/ and /t/ in both Spanish and English. Three groups of speakers were recorded: Spanish and English monolinguals, and Spanish-English bilinguals. The three groups were investigated separately.

Regarding English /d/ and /t/, the study found that the *fortis*, aspirated stop (/t/) was affected by stress while the *lenis* stop (/d/) was not (or marginally so). The effects of stress on English /t/ were additive—lexical stress lengthened aspiration in this consonant. Therefore, it could be said that stress enhances the acoustic difference between the two members of the /d/-/t/ contrast by affecting one of the two.

The situation for Spanish was different from that for English. In Spanish, it was the *lenis* stop (/d/), rather than the *fortis* one (/t/), the one to be impacted by lexical stress. Spanish /t/ was not modulated by stress. The effects of stress on Spanish /d/ were straightforward—stress lengthens prevoicing in this consonant. Once again, therefore, it could be claimed that stress enhances the acoustic opposition between /d/ and /t/ by displacing one of the two members of the contrast.

The findings of the present study regarding the effects of lexical stress on three VOT categories of two different languages are reminiscent of the literatures on the effects of (i) clear speech and (ii) of speech rate on these sound categories. For instance, [17] found that, in clear speech, aspiration is lengthened in English *fortis* stops and prevoicing is lengthened in Croatian *lenis* stops—short-lag VOT categories are largely unaffected by clear speech.

With respect to speech rate effects, studies typically find that as speech rate decreases (i.e., speech becomes slower) VOTs lengthen, but not for all stop consonants [18, 19]. In parallel to the studies on clear speech (and stress effects), the literature on speech rate has shown that aspiration is lengthened in English *fortis* stops while VOT is not impacted (or only minimally so) in stops characterized by short-lag VOTs [18]. In a fundamental cross-linguistic study, [19] researched the effects of speech rate on Thai, French and English stops. French has a contrast of the “Spanish-type”—*fortis* stops are voiceless, unaspirated and *lenis* stops are prevoiced. Thai has a three-way contrast—it possesses sets of (i) voiceless, aspirated stops, (ii) voiceless, unaspirated stops, and (iii) prevoiced stops. The study found that slow speech lengthened aspiration in English *fortis* stops and French *lenis* stops while short-lag VOT (English *lenis*, French *fortis*) stops were not affected. Regarding the case of Thai, the study found that short-lag VOT categories were not modulated by speech rate while the other two VOT categories were—VOTs were lengthened in the predictable directions.

A review of the literature suggests that the effects of speech rate, clear speech and lexical stress on VOT categories are very similar. What these studies have in common is that they find that, in two-way contrasts, short-lag VOT categories are unaffected while the other category is affected. In three-way contrasts, short-lag VOT categories are also the ones that remain largely unmodified. This could be summarized with the statement that short-lag VOT stop categories act as anchors in situations that lead to contrast enhancement, such as slow and/or clear speech and prosodic prominence.

The “contrast-enhancement” perspective, however, has been challenged by a study that investigated the effects of speech rate on Central Standard Swedish [2]. This language has a two-way contrast. It opposes a set of aspirated, voiceless

stops with a set of prevoiced stops—a short-lag VOT category is not found in this language. In Swedish, both stop consonants are affected by slow speech. In sum, the same categories that are affected in French, English and Thai are also affected in Swedish—the only difference is that Swedish does not have an “anchor”. The reasoning in [2] is that the Swedish contrast is already acoustically large in faster speech and does not need to be enhanced. [2] propose that speech rate, together with clear speech (and lexical stress), affects VOT categories that are *specified for a phonological feature*. In this view, aspirated stops are specified for [spread glottis], prevoiced stops are specified for [voice] and, importantly, short-lag VOT stops are left unspecified. Contrast enhancement is viewed as an artifact of feature modulation.

An alternative view that does not exploit feature theory is one that depends on articulatory gesture coordination. The VOT metric measures the time lag between two acoustic events that stand for two articulatory gestures, a laryngeal and a supralaryngeal gesture. As such, the VOT metric is a measure of gesture coordination. It is possible that speech rate, clear speech and prosodic prominence do not affect gestural coordination patterns for sounds whose two gestures are robustly synchronized (short-lag VOT stops), but do for those that are loosely synchronized (prevoiced and aspirated stops).

An analysis of the bilingual data found the following: (i) Spanish /d/ is robustly affected by lexical stress while Spanish /t/ is not, and (ii) both English consonants are slightly affected by lexical stress. Therefore, for these bilingual subjects, consonants with lead VOT (Spanish /d/) and long-lag VOT (English /t/) are impacted by lexical stress, although the former more so than the latter. This is the same result that was found for the native English- and Spanish-speaking groups. From the two short-lag VOT consonant (Spanish /t/, English /d/), only English /t/ is affected by lexical stress. While English monolinguals did not display robust stress effects on their English /d/, the bilinguals did.

A possible interpretation of the bilingual findings is suggested by the fact that a subgroup of the Spanish-English bilinguals produced English /d/s with some prevoicing—see Figure 1. This could be a transfer effect from Spanish /d/ for this subgroup. Since negative VOT in prevoiced stops is indeed affected by lexical stress (Spanish /d/, for instance) it is perhaps not surprising that English /d/, as produced by this particular group, is somewhat modulated by stress. Alternatively, English, but not Spanish, *lenis* stops, in general, could be impacted by lexical stress in this population. A strong interpretation of the “featural specification hypothesis” of [2], according to which only consonants with specified features are affected by prosodic modulations, is therefore not possible for the bilingual group. It could be concluded that the Spanish-English bilinguals in this study possess a phonological system consisting of [voice] (Spanish /d/), \emptyset (Spanish /d/), and [spread glottis] (English /t/). The nature of the English *lenis* stop would deserve more attention. If [voice] were to be chosen as the featural specification for English /d/, the question arises as to why the phonetic facts for this consonant differ from those for Spanish /d/.

5. Acknowledgments

The authors wish to express their gratitude to the following: Mark Amengual, Melinda Porta, Olivia Obeso, Miquel Llompart, and the 47 volunteers who participated in the study. The authors are also grateful to three anonymous reviewers, and to the SP-7 audience, for constructive criticism.

6. References

- [1] Rosner, B and López-Bascuas, L and García-Albea, J and Fahey, R, "Voice-onset times for Castilian Spanish initial stops," *Journal of Phonetics*, vol. 28, pp. 217–224, 2000.
- [2] Beckman, J and Helgason, P and McMurray, B and Ringen, C, "Rate effects on Swedish VOT: Evidence for phonological overspecification," *Journal of Phonetics*, vol. 39, pp. 39–49, 2011.
- [3] Cho, T and Keating, P, "Effects of initial position versus prominence in English," *Journal of Phonetics*, vol. 37, pp. 466–485, 2009.
- [4] Cho, T and McQueen, J, "Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress," *Journal of Phonetics*, vol. 33, pp. 121–157, 2005.
- [5] Cole, J and Choi, H and Kim, H and Hasegawa-Johnson, M, "The effect of accent on the acoustic cues to stop voicing in Radio News Speech," in *Proc. of the International Congress of Phonetic Sciences*, 2003.
- [6] de Jong, K, "The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation," *The Journal of the Acoustical Society of America*, vol. 97, pp. 491–504, 1995.
- [7] Lisker, L and Abramson, A, "Some effects of context on Voice Onset Time in English stops," *Language and Speech*, vol. 10, pp. 1–28, 1967.
- [8] M. Castañeda, "El VOT de las oclusivas sordas y sonoras españolas," *Estudios de Fonética Experimental*, vol. 2, pp. 91–110, 1986.
- [9] D. Poch, "Caractérisation acoustique des occlusives de l'espagnol: Le problème du VOT," *Revue de Phonétique Appliquée*, vol. 77, pp. 477–490, 1985.
- [10] M. Troya, "El VOT de las oclusivas sordas en la norma culta de Las Palmas de Gran Canaria," *Boletín de Lingüística*, vol. 17, pp. 31–38, 2005.
- [11] Guion, S, "The vowel systems of Quichua-Spanish bilinguals," *Phonetica*, vol. 60, pp. 98–128, 2003.
- [12] Amengual, M, "Interlingual influence in bilingual speech: Cognate status effect in a continuum of bilingualism," *Bilingualism: Language and Cognition*, vol. 15, pp. 517–530, 2012.
- [13] Birdsong, D and Gertken, L and Amengual, M, "Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism," 2012.
- [14] Olson, D, "Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production," *Journal of Phonetics*, vol. 41, pp. 407–420, 2013.
- [15] Magloire, J and Green, K, "A cross-language comparison of speaking rate effects on the production of voice onset time in English and Spanish," *Phonetica*, vol. 56, pp. 158–185, 1999.
- [16] Lisker, L and Abramson, A, "Crosslanguage study of voicing in initial stops," *Word*, vol. 20, pp. 384–422, 1963.
- [17] R. Smiljanic and A. Bradlow, "Stability of temporal constraints across speaking styles in English and Croatian," *Journal of Phonetics*, vol. 36, pp. 91–113, 2008.
- [18] J. Allen and J. Miller, "Effects of syllable-initial voicing and speaking rate on the temporal characteristics of monosyllabic words," *The Journal of the Acoustical Society of America*, vol. 106, pp. 2031–2039, 1999.
- [19] R. Kessinger and S. Blumstein, "Effects of speaking rate on voice-onset time in Thai, French and English," *Journal of Phonetics*, vol. 25, pp. 143–168, 1997.