

Voice Onset Timing in the English of Spanish Heritage Speakers

Tiempo de inicio de emisión de la voz en el inglés de hablantes de español como lengua de herencia

Molly Perara-Lunde*
University of New Mexico
Jessica Lindsey**
University of New Mexico
Richard J. File Muriel***
University of New Mexico

Recibido 14-10-15 / Aceptado 10-07-16

Abstract

The phonological system of bilinguals has been the subject of many recent studies. Several studies have specifically looked at Voice Onset Time (VOT) production in bilinguals in order to answer questions about the phonological system or systems of different types of bilinguals. For example, prior studies have looked at the effects of bilingual phonological production based on age of acquisition of a second language, cross-linguistic influence, and childhood experience with the language. This study will specifically examine the VOT production of a group of bilingual English-Spanish students who are enrolled in a Spanish as a Heritage Language Program (henceforth, SHL) at the University of New Mexico in Albuquerque, New Mexico. This population is of particular interest as Spanish and English have been in long-standing contact in this region of the country. This study will measure the VOT of this diverse group of bilinguals in order to assess whether and to what extent exposure to and/or proficiency in Spanish has an effect on the production of English VOTs in this region of the country.

Keywords: *Spanish as a Heritage Language, Bilingualism, Voice Onset Timing (VOT), Spanish in the U.S.*

* Profesora Universidad de Nuevo México. mlunde@unm.edu

** Profesora adjunta de español como lengua extranjera de Strayer University y Universidad de Nuevo México. jessica.a.lindsey@gmail.com

*** Director de estudios graduados de la Universidad de Nuevo México. richfile@unm.edu

Resumen

El sistema fonológico de los bilingües ha sido el tema de muchas investigaciones recientes. Específicamente, algunos estudios se han enfocado en el tiempo de inicio de la emisión de la voz –VOT, por sus iniciales en inglés– en hablantes bilingües para responder preguntas sobre los sistemas fonológicos de varias clasificaciones de bilingües. Por ejemplo, estudios anteriores han investigado temas como 1) el efecto de la edad de adquisición de la segunda lengua en la producción fonológica; 2) los efectos translingüísticos, y 3) el grado de efecto de las experiencias lingüísticas durante la niñez sobre la adquisición de una segunda lengua en la edad adulta. Este estudio específicamente examinará la producción VOT de un grupo de estudiantes bilingües español-inglés, matriculados en un programa de Español como Lengua de Herencia (SHL) en la Universidad de Nuevo México ubicada en Albuquerque, Nuevo México. Esta población es de interés particular ya que el español y el inglés tienen una larga historia de contacto en la región sudoeste de los Estados Unidos. Este estudio mide la producción VOT de este diverso grupo de bilingües a fin de evaluar si la exposición al español y/o la competencia en español durante la niñez tendrá algún efecto fonológico en la producción del VOT en esta región del país.

Palabras clave: *Español como lengua de herencia, bilingüismo, tiempo de inicio de la emisión de la voz (VOT), español hablado en los Estados Unidos.*

The phonological system of bilinguals has been the subject of many recent studies. Several studies have specifically looked at Voice Onset Time (VOT) production in bilinguals in order to answer questions about the phonological systems of different types of bilinguals. For example, prior studies have looked at the effects of bilingual phonological production based on age of acquisition of a second language, cross-linguistic influence, and childhood experience with the language. This study will add to that body of research by examining the VOT of a group of bilingual English-Spanish students enrolled in a Spanish for Heritage Speakers course at the University of New Mexico in Albuquerque, New Mexico. The purpose of this study is to explore the effect that exposure to Spanish may have on the English production of stops in this region of the country. It is well known that in the southwest, Spanish and English have been in contact for hundreds of years; this long-standing contact between the two languages could result in “phonological transfer” (Silva-Corvalán, 2001).

In a study that examined adult bilingual speakers of English and Italian in Canada, Flege, Schirru, and MacKay (2003) found that participants varied their production of /e/ depending on their length

of residency in Canada and the frequency with which they reported using Italian. Participants who had more recently arrived and used Italian regularly would “overshoot” their production of /e/, an act that the researchers interpreted as an attempt to distinguish Italian and English /e/ as much as possible. In contrast, those participants who had lived longer in Canada and reported using English regularly produced what many have referred to in the literature as a “compromised” /e/ with an acoustic value somewhere in between Italian and English /e/, suggesting that these bilinguals merged characteristics of their two languages to create a novel vowel category.

In the case of this study, we are particularly interested in phonological transfer of VOT, which is typically much longer in English than Spanish. While many young Hispanic people living in this region may have very limited or no fluency in Spanish, they are undoubtedly exposed to Spanish to some extent either through family members or other social networks. Our primary question is whether heritage speakers will exhibit a shorter, Spanish-like VOT due to either 1) their proficiency in Spanish or 2) close contact with family members and friends who are proficient in Spanish.

Review of Literature

Childhood Language Experiences

Several studies have sought to explore the effects that childhood language experiences can have on learning a second language in adulthood. Flege (1991) found that the VOT production of /t/ in Spanish-English bilinguals who learned Spanish in childhood was similar to the monolingual Spanish production of /t/ as opposed to those bilinguals who learned Spanish as an adult. Other studies have found similar results, drawing attention to the fact that childhood language experiences do have long-term effects insofar as the gestural timing (MacLeod & Stoel Gammon, 2005; 2010; Olson, 2013; Thornburgh & Ryalls, 1998). Many of these studies conclude that the reason early bilinguals tend to produce values that approximate those of monolinguals is because children start to form phonological categories very early on (Flege, 1991), and adults are not as proficient at establishing these categories. Other studies have pointed to the fact that these two separate phonological systems developing in children interact and do possibly affect one another (Kehoe, Lleó & Rakow, 2004).

Of particular note for this study are the findings from a group of researchers who have looked at the long-term effects that hearing a language during childhood can have on the learning of that language as an adult (Au, Knightly, Jun, & Oh, 2002). The researchers found that there are phonological advantages to hearing a language during childhood, as participants in their study exhibited more “native-like” accents than L2 learners who did not have any exposure to the language growing up: “The over-hearers were virtually native-like by this measure for the word-initial position; their VOT was reliably shorter than the typical late L2 learner” (p. 240). In a follow up study, this same group of researchers examined the perceptive skills

of childhood overhearers in a group of heritage Korean speakers. Similar to their earlier findings, childhood overhearers were able to perceive Korean /t/ more accurately than adult second language learners (Oh, Jun, Knightly, & Au, 2003). What is missing from these studies is a more realistic examination of bilingual communities: “los miembros de una comunidad comparten tanto normas lingüísticas como normas sociales que regulan las diferentes formas de hablar” (Silva-Corvalán, 2001, p. 10), signifying that in bilingual communities, it may be common practice to use both languages interchangeably and at the same time (Sánchez, 1982). It is precisely this type of language environment, which arises in multilingualism, that we are interested in examining.

Spanish-English bilingualism in the Southwest

Amengual (2012) analyzes the VOT of several groups of bilingual English-Spanish speakers in order to determine whether there is any cross linguistic influence at varying levels of proficiency in the phonological production for bilinguals. The study also investigates whether there are phonological effects on the production of Spanish/English cognates. The researcher examined the production of the voiceless stop /t/ in an unstressed position and found that Spanish/English bilinguals produce longer VOT for /t/ for Spanish/English cognates. This finding suggests that cognate status influences phonological production in the bilingual lexicon. The researchers interpret this as strong evidence for an exemplar model of representation, given that this model is well suited for this type of data in which words and chunks may be linked due to their similarity in meaning and/or form.

In addition, Amengual (2012) finds that Spanish heritage speakers produce shorter VOT values in English: “However, for the early bilingual groups, on average the Spanish Heritage group residing in

the United States produced VOT's of 17.1 ms and those of the English heritage group residing in Spain averaged 19 ms, a difference in an unexpected direction, that is, the Spanish heritage group's VOTs were lower than the English heritage group, who in addition to growing up hearing Spanish since birth also lived in a Spanish-speaking country" (p. 523). Furthermore, Balukas and Koops (2014) found that bilingual Spanish-English speakers in the state of New Mexico exhibited shorter VOT durations in English when approaching a code switch point. VOT values in Spanish, however, were left unaffected. While neither of these authors assume phonological convergence, it is possible to posit that contact with Spanish is affecting the VOT value of English.

Evidence from the studies reviewed here lead us to believe that VOT is an appropriate place to begin looking at possible points of influence in the speech of bilingual communities in this region of the country.

Data Set

This study uses data from semi-structured interviews to analyze the English VOT of 8 students between the ages of 19 and 26 enrolled in the Spanish as a Heritage Language (SHL) program at the University of New Mexico (UNM). The term "heritage learner" has been a contentious issue and many educators and researchers have questioned who exactly qualifies as a heritage learner. This program holds that it is important to take into account the sociopolitical history of the region when considering the term "heritage." In fact, researchers have documented instances of physical and psychological abuse inflicted on Spanish speaking students in the recent past for using Spanish at school (MacGregor-Mendoza, 2000). This has direct implications for the maintenance of Spanish among Hispanic communities in the region.

In response to the social, political, and linguistic history of the state of New Mexico, the SHL pro-

gram at UNM has adopted a very broad definition of a heritage learner in order to include those individuals who may feel that the Spanish language is an important part of their identity even if they are not proficient in the language. Students enrolled in the SHL program at UNM therefore exhibit a vast range of proficiency in the Spanish language, ranging from receptive skills to very high levels of oral proficiency. What this group does have in common is a cultural relationship with the Spanish language (Wilson & Martinez, 2011). All participants in this study were exposed to Spanish in some way during their childhood. Below is a brief profile of the study participants. Taking into account the broad definition of heritage learner that is used in this program, the participants in this study exhibit a great variation of experience with the Spanish language and thus are likely to exhibit variation in their VOT durations.

Participants

Participants in this study were enrolled in a fourth semester Spanish class. All participants are from New Mexico and all but one is from the city of Albuquerque (participant #2 is originally from Santa Fe, a city about 60 miles north of Albuquerque). The participants' age range was 19-26, with an average age of 20. When asked what language they spoke growing up, only two participants said Spanish. The majority of participants said mostly English with some Spanish, and one participant said only English. When asked who in their family spoke Spanish, there were various answers ranging from "everyone" to "only my dad's parents," suggesting that the participants in this study experienced different levels of exposure to Spanish during their childhood. Figure 1 below reports background information of each participant.

Since ethnic identity is likely to influence speech patterns, the participants were also asked what ethnic labels they use to describe themselves. Figure 2 below represents their responses to this question.

Figure 1: Background information for each participant

Participant	Age	Hometown	Language spoken at home during childhood	Who speaks Spanish in your family?
1	19	Albuquerque, Rio Rancho	Spanish	Everyone
2	19	Santa Fe	Mostly English, but some Spanish	Grandparents and parents
3	19	Albuquerque	Mostly English, some Spanish	Mom and her family
4	19	Albuquerque, Rio Rancho	English	Grandparents on mom's side
5	20	Albuquerque	Mostly English, some Spanish	Grandparents on mom's side
6	19	Albuquerque	Mostly English, some Spanish	Dad's parents, dad a little
7	26	Albuquerque, South Valley	Mostly English, some Spanish	Grandparents
8	19	Albuquerque	Spanish	Everyone

Figure 2. Labels to describe own ethnic identity

Participant	What labels do you use to describe yourself?
1	Hispana, Latina, Mexicana
2	Hispanic
3	Hispanic, White
4	Hispanic, Mexican
5	Hispanic
6	Hispanic generally and Mexican
7	Spanish, American, Chicana
8	Hispanic

This again demonstrates that, while these participants all find themselves in the same Spanish class, they identify themselves in distinct ways, often using more than one label to describe their background. Heritage learners are therefore a diverse group in that they have various levels of exposure to Spanish and use different labels to describe themselves ethnically.

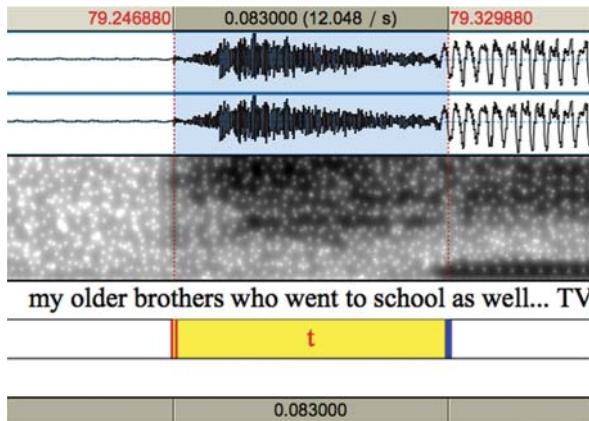
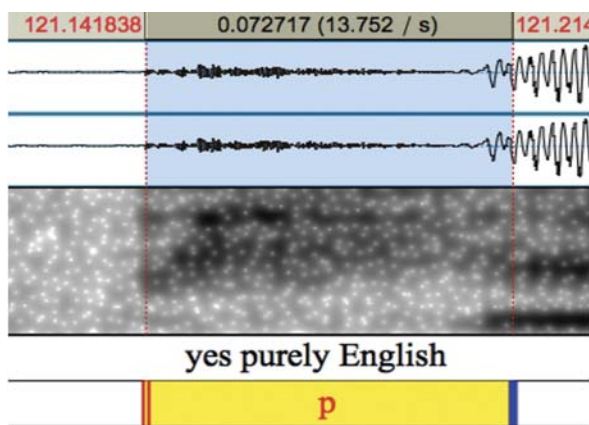
Hypothesis

In this paper we have three hypotheses. First, we predict that heritage Spanish speakers will evidence different VOT durations from previously used norms for American English (30-90 ms). Furthermore, we predict that this group will evidence shorter VOTs than American English VOT

norms. Our third hypothesis is that we expect our findings will show variation along a continuum in that VOTs of learners with the most exposure to Spanish will be the shortest, demonstrating a greater phonological influence from Spanish.

Methodology

This project analyzed the English Voice Onset Timing of voiceless stops /p, t, k/ in the speech of the 8 participants described above. Data was analyzed using Praat. We measured VOT from the initial burst and marked the end of the stop at the onset of the following vowel (see Examples 1 & 2 below).

Example 1**Example 2**

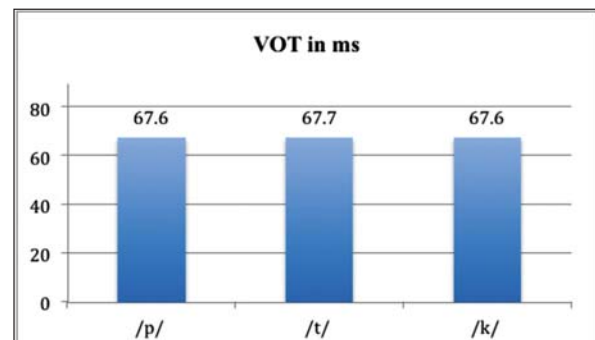
In cases where the beginning of periodic waveforms was difficult to assess, the rise of the second formant was also used to mark the end of the stop.

The VOT durations for 407 tokens were delimited using Praat v.5.4.21 (Boersma & Weenink, 2015) and entered into an excel spreadsheet, which included 101 instances of /p/, 139 /t/, and 167 /k/. Below we report the average durations of VOT for each voiceless stop. The linguistic variable considered was the duration of VOT and the social variable considered was self-reported childhood language use. We used Excel to find the average duration of each of the three stops as well as the average duration of each stop by participant. A

one-way ANOVA was run to determine if there were any patterns according to childhood language experience.

Results

This section will report the findings of this study, beginning with descriptive data characterizing the average VOT by token place of articulation /p, t, k/ and then, by participant. Below is a table that reports the average VOT by place of articulation across all speakers.

Figure 3

These results indicate that on average, this group of Spanish heritage speakers is producing a VOT within the traditional range of American dialects of English. These findings could support studies that have suggested that bilingual children are able to form phonological categories early on that contrast VOT in Spanish and English (Flege, 1991), especially observing that the two participants who reported speaking only Spanish growing up exhibited 'traditional' English VOTs. However, it's difficult to speculate about this as data was only gathered in English; to truly explore this theory, data would be needed in Spanish as well.

Researchers have found that

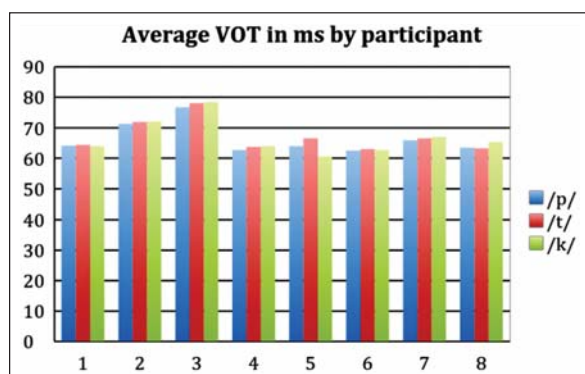
the most well-studied factor in VOT variation is place of articulation. It has been

confirmed in various studies that VOT increases when the point of constriction moves from the lips to the velum, both in isolated word reading and read speech (Zue, 1976; Crystal & House, 1988; Byrd, 1993; among others), and this pattern is not limited to the English language (Cho & Ladefoged, 1999) (Yao, 2009, p. 29-30).

However, this data set reveals an extremely homogeneous VOT value across all three places of articulation.

Figure 4 below gives the average duration of VOT of /p, t, k/ across all eight participants:

Figure 4



The data show a relatively homogeneous production of VOT across all three stops and across all eight participants, ranging between 60 and 79 milliseconds in duration. While the overall average values for /p, t, k/ indicate that VOT durations fall within the expected range of other monolingual American English varieties, it is interesting to note how similar the production of VOT is across participants.

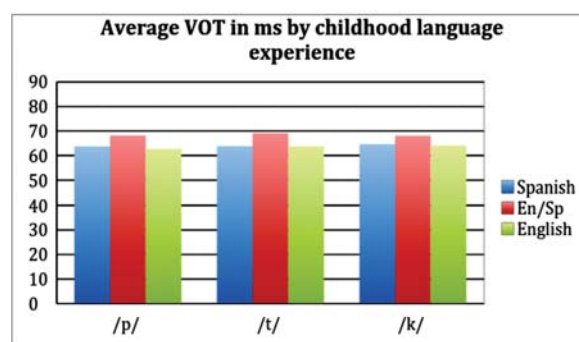
While VOT production is very similar across participants, there is some variation across individuals that could possibly correlate to ethnic identity labels. Participant #3, the only participant in this group that reported identifying as “White,”

produced the longest VOT values, and hence, the least Spanish-like. Future research should investigate this in more detail, by adding more participants that self-identify as “White” we could determine if ethnic identity is a significant factor in explaining this variation.

The primary interest of this study is to assess whether 1) Spanish is having any phonological effect on the English of bilingual communities in New Mexico and 2) whether there is a correlation between participants’ childhood language and production of VOT in English. For this reason, participants were put into three groups according to their self-reported childhood language (see Figure 1). Two participants (Participants 1 and 8) were labeled as “Spanish” for childhood language; one participant (Participant 4) was labeled as “English”, and the five remaining participants reported that they spoke predominantly English but some Spanish during their childhood and were therefore put in a category labeled “English/some Spanish”.

Figure 5 below reports the average VOT of each stop for each of the three language groups:

Figure 5



Of immediate interest is the fact that those participants who reported using mostly English and some Spanish during childhood show slightly longer VOT values across all three stops, contrary to the hypothesis that the “English childhood

language” group would have the longest VOT’s. ANOVAs were run to assess whether this difference was statistically significant. A one way between subjects ANOVA was run to compare the duration of VOT in all three stops with the variable childhood language experience. No statistical significance was found for any other of the three ANOVA’s run (for /p/, [F(2, 98) = 1.5, p = .22], for /t/, [F(2, 136) = 1.1, p = .32], and for /k/, [F(2, 164) = 1.0, p = .36]

Conclusions

The hypothesis that Spanish is having a phonological effect on English is not supported by this set of data. This group of Spanish heritage speakers produced relatively uniform VOTs that fall into the prescribed range of American English VOT. Furthermore, there seems to be no correlation between childhood language and duration of VOT in English. While the two participants who spoke Spanish during childhood did exhibit shorter VOT’s than the participants who spoke primarily English and some Spanish, the one participant who spoke English at home during childhood exhibited the shortest VOTs, showing no correlation between the amount of exposure to Spanish during childhood and length of VOT in English.

Other studies have pointed to factors that may be contributing to phonological convergence between Spanish and English in the southwest region of the United States. Balukas and Koops (2014) found that English VOT was shortened when approaching a code switching point and Amengaul (2012) found that Spanish heritage speakers from the Southwest produced a short, Spanish-like VOT in English. The data gathered in this study does not align with these previous findings and there are several factors that could be at work. The graduate students who conducted these interviews were not members of the same social networks that the participants were a part of, possibly contributing to the

participants using a more standardized form of English during the interview. The interviews were conducted on campus in an academic setting, possibly also contributing to the use of a standard variety of English. Furthermore, while there was no information gathered on the socioeconomic status of the participants, all of them were enrolled in college courses. It’s possible to infer that socioeconomic status may be a contributing factor in the speech of this group of college students. It would be ideal to reproduce this study using interviews conducted by local community members in order to compare the results. Investigating other linguistic variables such as speech rate could also yield interesting results, as both researchers perceptually noted this specific variable while collecting data.

In order to get a more complete picture of the VOT of this group of bilinguals, it would be necessary to also examine their production of VOT in Spanish, as well as the VOT of monolingual English speakers from this same community. Another unexplored possibility is that this group is producing longer VOTs in Spanish, and that phonological convergence is working in the opposite direction that was expected in this study. Unfortunately, there is no available access to this group’s production of Spanish. While there were some very limited pieces of data that could qualify as code switching, it was not substantial enough to incorporate into this data.

Limitations and Future Research

There are several limitations to this study. Due to the nature of the interviews, the researchers coded a high number of repeated words. Assuming an exemplar-based phonological representation, it’s very likely that individual words develop their own unique phonological and gestural patterns. Thus, not including “word” as a predictor might miss any patterns that could be lexically unique.

These interviews were conducted by two graduate students from the University of New Mexico, neither of them were natives of New Mexico and therefore did not speak a dialect of Spanish or English from that state. It's possible that study participants chose to use a more standard form of English during the interviews due to the place and setting of the interviews. In fact, one interviewer was a native English speaker and the other was a native Spanish speaker. Whether the interviewees accommodated their language in some way according to the native language of the interviewer was not explored in this study.

Because only one participant in the study was male, gender was not a variable considered. Future research should seek to include a higher number of male participants to see whether male and female participants differ in VOT production. As previously mentioned, it would be interesting to gather more data from participants who identify as "White" to further explore whether ethnic identity plays a role in VOT production in this region of the country. It's also possible that socioeconomic status played a part in the findings of this study; while there was no information gathered about the socioeconomic status of the participants, they were all enrolled in college courses, possibly suggesting that they come from a middle or upper class background. Other linguistic variables should have been considered, including speech rate, as this was a salient factor observed by both researchers in the duration of VOT.

While some statistical analysis was conducted, it is important to note that the three childhood language groups were not evenly comprised of participants, nor were there comparable numbers of tokens gathered across groups. A higher token count for some individual participants, as well as groups comprised of equal numbers of participants, may yield different results.

In the future, it would be optimal to gather data in both English and Spanish in order to gain a fuller picture of the speech of bilingual speakers. It was difficult to find intervocalic, word initial voiceless stops because many words in English end in consonants. It could be interesting to approach voiceless stops in English from a different perspective and include the nasalized and affricated tokens that were not examined in this study.

References

- Amengual, M. (2012). Interlingual influence in bilingual speech: Cognate status effect in a continuum of bilingualism. *Bilingualism: Language and Cognition*, 15(03), 517-530.
- Au, T. K. F., Knightly, L. M., Jun, S. A., & Oh, J. S. (2002). Overhearing a language during childhood. *Psychological science*, 13(3), 238-243.
- Balukas, C., & Koops, C. (2014). Spanish-English bilingual voice onset time in spontaneous code-switching. *International Journal of Bilingualism*. doi1367006913516035.
- Boersma, P. & Weenink, D. (2015). *Praat: doing phonetics by computer* [Computer program]. Version 5.4.21, retrieved from <http://www.praat.org/>
- Byrd, D. (1993). 54,000 American stops. *UCLA Working Papers in Phonetics*, 83, 97-116.
- Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of phonetics*, 27(2), 207-229.
- Crystal, T. H., & House, A. S. (1988). Segmental durations in connected-speech signals: Current results. *The journal of the acoustical society of America*, 83(4), 1553-1573.
- Flege, J. E., Schirru, C., & MacKay, I. R. (2003). Interaction between the native and second language phonetic subsystems. *Speech communication*, 40(4), 467-491.
- Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *The Journal of the Acoustical Society of America*, 89(1), 395-411.
- Kehoe, M. M., Lleó, C., & Rakow, M. (2004). Voice onset time in bilingual German-Spanish children. *Bilingualism: Language and Cognition*, 7(01), 71-88.

- Knightly, L. M., Jun, S. A., Oh, J. S., & Au, T. K. F. (2003). Production benefits of childhood over-hearing. *The Journal of the Acoustical Society of America*, 114(1), 465-474.
- MacGregor-Mendoza, P. (2000). Aquí no se habla Español: Stories of linguistic repression in southwest schools. *Bilingual Research Journal*, 24(4), 355-367.
- MacLeod, A., & Stoel-Gammon, C. (2005). Are bilinguals different? What VOT tells us about simultaneous bilinguals. *Journal of Multilingual Communication Disorders*, 3(2), 118-127.
- MacLeod, A. A., & Stoel-Gammon, C. (2010). What is the impact of age of second language acquisition on the production of consonants and vowels among childhood bilinguals? *International Journal of Bilingualism*, 14, 400-421.
- Magloire, J. E. L., & Green, K. P. (1999). A cross-language comparison of speaking rate effects on the production of voice onset time in English and Spanish. *Phonetica*, 56(3-4), 158-185.
- Oh, J. S., Jun, S., Knightly, L. M., & Au, T. K. (2003). "Holding on to childhood language memory," *Cognition* 86, B53-B64.
- Olson, D. J. (2013). Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics*, 41(6), 407-420.
- Sánchez, R. (1982). Our linguistic and social context. In J. Amastae, L. Elías-Olivares. *Spanish in the United States: Sociolinguistic aspects*. (pp. 9-46). Cambridge: Cambridge University Press.
- Silva-Corvalán, C., & Enrique-Arias, A. (2001). *Lenguas en contacto y bilingüismo. Sociolingüística y pragmática del español*. 269-332. Washington: Georgetown University Press.
- Thomas, E. (2010). *Sociophonetics: an introduction*. New York: Palgrave Macmillan.
- Thornburgh, D. F., & Ryalls, J. H. (1998). Voice onset time in Spanish-English bilinguals: Early versus late learners of English. *Journal of communication disorders*, 31(3), 215-229.
- Wilson, D. V., & Martinez, R. (2011). Diversity in Definition: Integrating History and Student Attitudes in Understanding Heritage Learners of Spanish in New Mexico. *Heritage Language Journal*, 8(2), 115-133.
- Yao, Y. (2009). "An exemplar-based approach to automatic burst detection in spontaneous speech," in UC Berkeley Phonology Lab 2009 Annual Report (UC Berkeley Phonology Laboratory, Berkeley, CA), pp. 13-28.
- Zue, V. W. (1976). Acoustic characteristics of stop consonants: A controlled study (Ph. D. Dissertation, Massachusetts Institute of Technology). Retrieved from https://dspace.mit.edu/bitstream/handle/1721.1/29469/Zue_Victor_ScD_1976.pdf?sequence=1