2011-01-01

Resetting VOT in a Bilingual Region

Dan William Morgan

University of Texas at El Paso, dwmorgan@miners.utep.edu

Follow this and additional works at: https://digitalcommons.utep.edu/open_etd

Part of the Linguistics Commons

Recommended Citation

https://digitalcommons.utep.edu/open_etd/2349

This is brought to you for free and open access by DigitalCommons@UTEP. It has been accepted for inclusion in Open Access Theses & Dissertations by an authorized administrator of DigitalCommons@UTEP. For more information, please contact lweber@utep.edu.
RESETTING VOT IN A BILINGUAL REGION

DAN WILLIAM MORGAN
Department of Languages and Linguistics

APPROVED:

__________________________________________
Jon Amastae Ph.D., Chair

__________________________________________
Charles Elerick, Ph.D.

__________________________________________
Bess Sirmon Fjordbak, Ph.D.

__________________________________________
Benjamin C. Flores, Ph.D.
Acting Dean of the Graduate School
RESETTING VOT IN A BILINGUAL REGION

by

DAN WILLIAM MORGAN, B.A.

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF ARTS

Department of Languages and Linguistics

THE UNIVERSITY OF TEXAS AT EL PASO

August 2011
Acknowledgements

Though this thesis only has my name as the sole author, I would never have finished it without the guidance and efforts of a lot of people. There were times when I felt that there was no possible way that this project would ever be completed, and I will be forever indebted to these people.

I want to express my gratitude to my thesis committee for their support during the writing process. I am much appreciative of the time and effort that my director, Dr. Amastae put into helping me write my thesis. His suggestions and input kept me headed in the right direction during the entire process. Whenever I became unsure of a certain principle, he was always willing and able to discuss it with me.

I’d like to thank Dr. Elerick, who was important in helping me establish exactly what my research question was and how I wanted to research it. My thanks also go out to Dr. Fjordbak who was my outside reader. Her enthusiasm and kindness encouraged me to carry on. Also, her notes and suggestions were very helpful in finishing the project.

I want to thank Christine Susany. She was instrumental in helping me edit my thesis. She spent hours with me going over my writing. Her edits and suggestions helped transform my writing from a garbled mess at times into clear, logical statements. I cannot stress enough how much she helped and encouraged me.

A very special thank you goes out to all my family who has supported me unceasingly. My wife, Michelle, has always believed in me. She has supported me during the entire process, even when I got ‘snippy’ with her. Her love was always felt. To my mother and father, who were always willing to watch Virginia, my baby, so that I could work on my thesis I give thanks. I also want to thank my late grandmother who was always an inspiration in my life.
I would like to thank all of my teachers who have helped mentor me and guide me throughout my life in order to get me to where I am now. Finally I want to thank a good friend, Kaulana Shum, who introduced me to the world of linguists and encouraged me to take the introductory class.
# Table of Contents

Acknowledgements......................................................................................................................... ii

Table of Contents.................................................................................................................................. vi

List of Tables ........................................................................................................................................... ix

List of Figures .......................................................................................................................................... x

Chapter 1: Introduction to VOT .......................................................................................................... 1

Chapter 2: Characteristics of VOT within Language ......................................................................... 6

2.1 Average VOT in Spanish and English ......................................................................................... 6

Chapter 3: Bilingualism....................................................................................................................... 9

Chapter 4: Perception and Meaning of VOT .................................................................................... 12

4.1 Introduction ................................................................................................................................... 12

4.2 How Speech is Perceived ............................................................................................................. 13

4.3 Perception in Spanish and English .............................................................................................. 13

Chapter 5 History of Research on VOT ............................................................................................. 16

5.1 Introduction ................................................................................................................................... 16

5.2 Establishing Voiced and Voiceless Stops .................................................................................... 16

5.3 Compromised VOT values ............................................................................................................ 17

5.4 Search for Universals in VOT ....................................................................................................... 17

5.5 VOT with Japanese Subjects ......................................................................................................... 18
5.6 VOT Study in El Paso........................................................................21

Chapter 6 The Current Study......................................................................23

6.1 Introduction......................................................................................23

6.2 The Subjects....................................................................................25

6.3 Monolingual English Speakers.......................................................25

6.4 Monolingual Spanish Speakers.......................................................26

6.5 L1 English L2 Spanish Bilinguals....................................................27

6.6 L1 Spanish L2 English Bilinguals....................................................28

6.7 Tokens............................................................................................28

6.8 Procedures......................................................................................28

6.9 Study Limitations............................................................................31

Chapter 7 Results and Discussion.................................................................33

7.1 Introduction......................................................................................33

7.2 Monolingual Spanish Speakers.......................................................34

7.3 Monolingual English Speakers.......................................................36

7.4 L1 Spanish L2 English Bilinguals....................................................37

7.5 L1 English L2 Spanish Bilinguals....................................................38

Chapter 8 Discussion and Conclusion.........................................................39
References........................................................................................................................................47

Appendix 1 English questionnaire..................................................................................................49

Appendix 2 Spanish questionnaire..................................................................................................50

Appendix 3 Texts..................................................................................................................................51

Appendix 4 English informed consent form......................................................................................52

Appendix 5 Spanish informed consent form......................................................................................55

Vita .......................................................................................................................................................558
List of Tables

Table 1: .........................................................................................................................34

Table 2: ..........................................................................................................................35

Table 3: ..........................................................................................................................35

Table 4: ..........................................................................................................................36
List of Figures

Figure 1: ..................................................................................................................... 24
Figure 2: ..................................................................................................................... 30
Figure 3: ..................................................................................................................... 31
Figure 4: ..................................................................................................................... 40
Figure 5: ..................................................................................................................... 42
Chapter 1: Introduction to VOT

In articulatory phonetics, linguists categorize articulatory gestures using three distinct criteria: voicing, place of articulation, and manner of articulation. Voicing refers to the vibration of the vocal cords while producing speech. Place of articulation describes where multiple speech organs either touch or approach each other to produce a sound, while manner of articulation is the way a sound is articulated. The majority of articulated sounds are classified as stops, fricatives, affricates, and nasals. In this paper, the sounds of most interest are /p t k/ and on a lesser scale, /b d g/. The bars, / /, denote phonemes, which are the smallest units in a language used to convey a distinction in meaning. Minimal pairs, meaning pairs of words that differ by only one sound in the same position, are a common way to establish phonemes. For example, *dip* /dIp/ and *tip* /tIp/ are minimal pairs in English, and the the only difference between the two words is the initial consonant. By changing /t/ to /d/, the meaning of the word can be altered. Conversely, [tʰip] and [tip] do not have separate meanings, so in this example a minimal pair is not formed.

Both English and Spanish contain the phonemes /p t k/, as can be seen in the following examples from both language:

<table>
<thead>
<tr>
<th>English</th>
<th>Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>park</td>
<td>para</td>
</tr>
<tr>
<td>tea</td>
<td>taco</td>
</tr>
<tr>
<td>kin</td>
<td>castillo¹</td>
</tr>
</tbody>
</table>

¹ While castillo has no k in its orthography, or spelling, the initial phoneme is nonetheless /k/.
The phonemes /p t k/ are universally the unmarked stop consonants. Nevertheless, their articulation varies across languages. One major difference is the variation in voice onset time. Voice onset time, or VOT, is the period of time between the release of a stop and the start of the voicing of the next sound. When voice onset time is greater than 50 milliseconds, it is recognized as having a long lag, which is often called aspiration. Voice onset times that are shorter than 30 milliseconds are recognized as having a short lag (Kilpatrick 2003). Where /p t k/ begin to differ between English and Spanish is at the allophonic level.

Aspiration is not part of the lexical representation in English. Allophones are variants of a phoneme. They occur only in specific environments, or within surrounding sounds. This is to say that allophones are predictable and do not change the meanings of words. The /p/ in spill and pill might appear to be the same, but if examined closer, ill proves to be preceded by what seems to be a puff of air in pill but not in spill. Underspecification is a phenomenon wherein particular features are omitted in a representation. It states that features must only be underspecified if their values are predictable (Archangeli 1988). Aspiration is predictable in English, because it always happens in the same environment, and so /p t k/ become [pʰ tʰ kʰ] at the beginning of a stressed syllable.

However, aspiration does not always have to be predictable as it is in English. For example, Thai, the official language of Thailand, also uses aspiration, but aspirated stops are phonemes as opposed to allophones. By looking at the following examples of Thai, it is evident that aspiration in /p/ does form minimal pairs, thus showing /pʰ/ to be a separate phoneme from /p/.
Aspiration is often referred to as a puff of air, since aspiration does refer to breathing. However, the term, aspiration, can be misleading when associated to long VOT. VOT is measured in terms of length and not air pressure. This could be due to the difficulty in measuring air. If the long VOT of pill is shortened considerably, the sound of the words becomes strange but the meaning remains unaltered. Long VOT occurs in certain environments in English, and we see that word initial /p/ followed by a stressed vowel results in aspiration. This is also holds true for /t/ and /k/. Spanish, on the other hand, contains no aspiration with regards to /p t k/ in any environment.

This paper examines three different categories of VOT. In the first category, the start of voicing significantly corresponds with a stop release. Sounds generated when the start of voicing delays moderately after a stop release fit the second category, and the third category is when the start of voicing delays greatly after a stop release. Voiced and voiceless stops in Spanish fit the first category. Conversely, voiceless stops in English belong in the third category, while voiced stops in English may be placed in the first category, even when preceding a vowel at the beginning of a word or stressed syllable. That the VOT of voiced and voiceless stops in Spanish and English belong in separate categories illustrates the problem that Cynthia Kilpatrick observed, that voiceless stops in Spanish and voiced stops in English share a similarity that could pose problems for learners.
Both English and Spanish learners face a challenge in learning their respective second languages. They must reset the phonetic rules of their first language in order to match the rules of their new target, or second, language. This is not an easy task. A study done by Kilpatrick in 2003 confirmed this to be true. Kilpatrick showed that while most of her participants understood that VOTs of voiceless stops in English and Spanish differed, the average VOT of the bilingual English and Spanish speakers varied significantly from that of English or Spanish monolinguals. Her results also showed that in order to get non-compromised VOT values, a bilingual must have learned both of his/her languages while of a young age. Age, according to Kilpatrick, among other linguists, was a significant determinant in successful language acquisition.

Spanish speakers are accustomed to /p t k/ having no aspirated allophones, but the English language phonemes /p t k/ can all be aspirated with relatively long VOTs. While word-initial stops with a short lag before following vowels are perceived as voiced in English, such is not the case in Spanish. This complicates matters for Spanish speakers learning English, who must be able to both perceive stops with long lags as voiceless and also lengthen the VOTs of these voiceless stops in their speech.

In contrast, the English speakers learning Spanish have the advantage of already being accustomed to short VOTs with voiced stops like /b d g/, that should have no VOT since they are voiced along with their following vowel. Regardless, these language students must still learn how to shorten the VOTs associated with /p t k/ in word-initial and medial stress position. Kilpatrick argued that it is easier for English speakers to acquire Spanish VOT values than the other way around, because
native English speakers acquire non-aspirated voiceless stops in non-stressed medial position as in the word, apple. Spanish speakers, instead, do not acquire any voiceless stop with an aspiration feature.

In order to test the claim by Kilpatrick that English speakers really do have an advantage over Spanish while learning their L2, a study of comparative and contrastive data from native English L2 Spanish speakers and native Spanish L2 English speakers was initiated. This study attempted to prove if in fact and to what extent such an advantage exists for either group of people.
Linguists Lisker and Abramson (1964) revealed that VOT is predictable. While varying among languages, and to a lesser degree, among speakers of the same language, VOT is predictable according to the environment in which it is found. So, VOT is affected by its environment. Voiced consonants are generally shorter than their corresponding voiceless consonant, e.g. /g/ and /k/, /g/ being the voiced consonant and /k/ the voiceless. Also, the place of articulation affects VOT. Bilabial stops, /b, p/, are produced with shorter VOTs than velar stops, /k, g/. This is because bilabial stops are articulated at the front of the mouth. The shape of the front of the mouth causes the VOT in stops to be shorter. Velar stops, on the other hand, are produced at the back of the mouth, and the VOT of velar stops are longer. This is caused by a small space behind the tongue and a large space in front. This triggers a delay in the vibration of the vocal chords (Kilpatrick 2003). Alveolar stops, /t, d/, at least in English, fit somewhere in the middle.

Another important characteristic of VOT is that it is affected by the speed as well as the context in which words are produced. Major (1992) showed average VOT values to be shorter in conversational speech than formal speech. He stated that this is due to the speed in which people talk. Formal speech is slower and more careful, while conversational speech is generally quicker. With this in mind, VOT values should be even longer when read from a passage.

Unfortunately, linguists have always had to struggle with the observer’s paradox (Labov 1972). They want to observe natural, casual speech, but if natural
speech is being observed, then the speech is no longer completely natural. Natural speech involves someone speaking while someone else listens and responds. There are few, if any, natural speech environments in normal life where somebody speaks, while somebody else listens to that speech and analyzes it. The closest thing that researchers can hope for is referred to as naturalistic data. Naturalistic data is elicited by researchers as they record data in a natural setting instead of in a lab. There is a varying degree of styles of speech, including minimal pairs, world lists, reading passages, and interview styles. In the interview styles, the researcher attempts to elicit casual speech from subjects by asking them questions that bring out stories tied to strong emotional memories. The idea is that the more the subject is concentrating on an emotional story, the more casual the subjects’ speech becomes. On the other hand, minimal pairs, word lists, and reading passages elicit more formal speech. Formal speech is not bad and should not be ignored. For one thing, formal speech allows researchers to obtain a frame of reference for how their subjects think and understand language (Labov 1972).

2.1 Average VOT in English and Spanish
In order to comprehend the value of VOT in bilinguals, it is necessary to look at the values of VOT in both the spoken English and spoken Spanish of monolinguals. Many studies have been completed to determine the average VOT for speakers of English. These times fluctuate from as low as 58 ms (Lisker and Abramson 1964) and as high as 93 ms (Major 1992). Contrastingly, Spanish shows much lower VOTs, which range from 4 ms to 29 ms. (Lisker and Abramson 1964). These VOT values were produced by monolingual speakers inside a lab and not within a natural setting.
A bilingual, native English speaker who learned Spanish after childhood could have a reduced VOT due to transfer from Spanish. The reverse would also happen to a native Spanish speaker who acquired English after childhood. Meanwhile, native Spanish speaker’s VOT might increase because of transfer from English.
Chapter 3: Bilingualism

As a matter of convenience, a person’s first language will be referred as L1, and a second learned language will be L2. A person who speaks English as a native language, and can speak Spanish as a second language will be referred to as a L1 English L2 Spanish bilingual.

What is required for a person to be regarded as a bilingual? If one were to rely on the words of John Edwards (Edwards 2006), then everyone would be bilingual, because everyone knows at least a word or two words from another language, be it adios or sayonara – even my spell checker knows these two words. Edwards clarified his ideas about bilingualism according to degrees. Thus, there are many levels attached to bilingualism, namely the level of fluency in reading, writing, speaking, and listening. Can someone be bilingual if they are able to listen and speak fluently in two languages but can only read and write fluently in one language? Where do we draw the line at in defining what is and isn’t a bilingual?

Other linguists have been vague in their descriptions of bilingualism. (Weinreich 1953) described bilingualism as the varying use of two languages. Haugen (1953) addressed this issue when he stated that a bilingual is one who has the ability to converse meaningfully in another language. Neither of these definitions gives a clear meaning to what a bilingual is and leaves much up to interpretation. Other linguists, such as Bloomfield (1933), are more definite in defining a bilingual.
He classifies bilinguals as people with native-like ability in two languages. To define the term bilingual is very complicated, as it differs greatly among people.

Without listing out the vast array of the different types of bilinguals, it is necessary to at least have a general understanding of all the diverse types of bilinguals. Although there is not a uniformed method of classifying bilinguals among today’s linguists, there are some terms that are commonly used to classify bilinguals. Bilinguals can be classified by age or by skill. In terms of age, a bilingual can be either an early bilingual or a late bilingual. An early bilingual is someone who started learning two languages at an early age, although one language could have been acquired first. A late bilingual is someone who learns a second language after their childhood. These types of bilinguals usually are not able to communicate in their L2 as well as their L1. In terms of skill, a bilingual can be classified as receptive, dominant, or balanced. A receptive bilingual is someone who is able to understand an L2 in written form and/or spoken form but might not be able to speak or write it. A dominant bilingual is one who is more proficient in one language than in another. A balanced bilingual is someone who speaks two languages with the same proficiency.

Essentially, people learn languages at different times of their lives, at different paces, and in different areas (classroom or natural environment). This makes it nearly impossible to find people who share the exact same level of bilingualism.

For this study it was necessary to find bilinguals that fit into the dominant bilingual category. This study needed bilinguals who clearly favored their L1 over their L2. Early bilinguals who are also proficient bilinguals would be similarly skilled in English as they are in Spanish, thus, using them as participants would remove the
variable from the study given that the purpose was to determine whether either native English speakers or native Spanish speakers can acquire the VOT norms of their L2.
Chapter 4: Perception and Meaning of VOT

4.1 INTRODUCTION

VOT can determine the perception of voicing and aspiration contrast in stops, e.g. /k/ vs. /g/ or [k] vs. [kʰ]. Stimuli are not typically recognized categorically (Kess 1992). For example, we do not see a color spectrum from red to violet as either pure red or pure violet with nothing in between. A color can be blue-green or forest green, or even “sort of” green. However, a stop cannot be “sort of” /p/ or “kind of” /b/; it is either a /p/ or a /b/, and it must be that way in order for speakers to understand each other. People perceive speech\(^2\), or phonemes, categorically.

However, this idea of category is not to be confused with people’s perception being skewed by orthography. When one hears the word, cotton, spoken in fast speech, he or she may claim to hear a /t/ even if only a glottal stop was uttered in place of that /t/. Orthographical confusion can also occur between languages. The word pizza is pronounced /pɪtsa/ in English and /bɪtsa/ in Arabic. If someone speaking Arabic uttered /bɪtsa/ and native English speakers not familiar with Arabic heard the utterance, they would probably state that the initial consonant sounded like a /b/ but would not identify it categorically as a /b/, even though Arabic only contains a /b/ and no /p/ phoneme. This confusion occurs because pizza is spelled with a ‘p’ in English.

\(^2\)This is not to say the perception is easy. Speakers trying to perceive sounds in a language different from their L1 could find perception to be a difficult task. For example, the word “Pakistan”, when spoken by a native Urdu speaker, sounds like /bakistan/ to an American. Hindi-Urdu distinguishes between /p/, [pʰ], and /b/, so its speakers should have no problems physically perceiving as well as producing aspirated stops, but they do.
4.2 How Speech is Perceived

When studying a language, people usually think of the production of speech, but what is often forgotten is the perception. For example, speech is most often defined in terms of articulation—/b/ as a voiced bilabial fricative. A common illustration of perception gone askew is seen with the “Mcgurk Effect”, where /ba/ is uttered, but /ga/ is articulated. This causes the hearer to most likely perceive either /ga/ or /da/. So perception is different from articulation.

According to Thomas (2011), our auditory systems do not hear the same things that we first hear with our ears. The cochlea, an organ inside our inner ear, converts sound waves into neural signals and sends them to the brain. When sections of the cochlea vibrate, tiny hairs excite, affecting nearby nerves. The sensitivity of the cochlea with respect to frequency and amplitude is complicated, and Thomas (2011) refers to it as logarithmic. The cochlea perceives amplitude differences better with certain frequencies than with others, and the way that these frequencies are perceived is known as pitch. The way that people perceive vowels and consonants is dependent upon these differences of pitch and what Thomas calls raw frequency (Thomas 2011).

4.3 Perception in English and Spanish

As stated earlier, the release of a voiceless stop in Spanish occurs right before the onset of voicing with regards to the following vowel. The same is true with respect to voiced stops in English. For example, /d/ in English and /t/ in Spanish will have

---

3 The McGurk Effect was first explained by Harry McGurk and John MacDonald in 1976. It works by placing a dubbed recording of one phoneme (/ga/) and matching it to a video of another phoneme (/ba/) being produced. What results is usually a 3rd perceived phoneme (/da/) (McGurk 1976)
approximately the same VOT value. Both /d/ and /t/ have the same place of articulation, on the alveolar ridge, and the same manner of articulation, a stop. The only difference between these two phonemes is that /d/ is voiced while /t/ is voiceless. With the release and onset of voicing happening at about the same time, English voiced stops and English voiceless stops are pronounced almost indistinguishably (Kilpatrick 2003).

Kilpatrick (2003) uses the example of *tia* and *dia* to explain that for a monolingual English speaker, both words may sound the same because the /t/ in *tia* has basically the same VOT value as /d/ in English. Kilpatrick (2003) calls this a problem with perception. On the other hand, Spanish speakers have what Kilpatrick (2003) calls a problem with production. Looking at the minimal pairs, *tip* and *dip*, a Spanish speaker will produce something that to a monolingual speaker sounds like *dip* and *dip*. This can cause confusion when two bilingual speakers of differing languages converse.

How much aspiration does a Spanish speaker need to produce when uttering voiceless stops in order to make themselves clear? How much aspiration is needed for an English speaker to understand a Spanish speaker with respect to voiceless stops? Kilpatrick (2003) refers to a term known as crossover. Crossover is the juncture where voiced stops become perceived as voiceless stops, and vice versa. This crossover point is at about 25 ms for native English speakers, while the crossover point for native Spanish speakers is significantly lower. Stops with VOT values above 25 ms. are recognized as voiceless, and stops below that crossover point are recognized as voiced for English speakers. This crossover is where bilinguals can become confused in terms of VOT of stops.
Since speech is perceived categorically, it is quite difficult for speakers to differentiate between all speech sounds. Generally speaking, speakers are only able to differentiate speech sounds of their own native language. When a language contains speech sounds different from one’s own native language, one must train the ear to even be able to perceive it. Thus, when Spanish speakers hear an aspirated consonant, they will not perceive any difference from a non-aspirated consonant, unless they have had enough exposure to English.
Chapter 5: History of Research on VOT

5.1 INTRODUCTION

Lisker and Abramson (1964) were the first researchers to study VOT. They defined VOT to be the measure of the time between the release of a stop and the start of the next phoneme. Lisker and Abramson studied eleven different languages with the purpose of establishing universals among languages with respect to VOT. They determined that three categories, 1) voiceless unaspirated stops (0–25 ms), called short lag; 2) voiceless aspirated stops (50–100 ms), called long lag; and 3) voiced stops, when the start of vibration begins before the release of the stop, were universal to all the languages they were studying. However, they also determined that all languages do not share the same boundaries to distinguish categories of stops. For example, English voiceless stops, which fit into the sixty to one hundred milliseconds range, are longer than Spanish voiceless stops.

5.2 ESTABLISHING VOICED AND VOICELESS STOPS

Lisker, et al. (1977) were the first linguists to perform a study of VOT as a basis for categorizing stops. They were looking into what they called, VTD, voiced transition duration. Prior to this, Stevens and Klatt (1970) had hypothesized that initial voiced and voiceless stops were perceived as such due to the voiced transition into the following vowel, as opposed to the preceding voiceless interval. Lisker, et al. determined stops in English might not be perceived as voiced or voiceless due entirely to the voiceless interval of VOT, but they also concluded that VTD had much less to do with determining voiced or voiceless stops.
5.3 COMPROMISED VOT VALUES

Fledge and Hillenbrand (1984) performed a study of bilingual speakers of both English and French and English and Spanish. Fledge and Hillenbrand (1984) along with Kilpatrick (2003) found that native Spanish speakers who were late learners of English produced compromised VOT values. These values were somewhere in the middle of what normal values of English and Spanish should be. Fledge and Hillenbrand (1984) found that native English speakers studying French produced initial /t/ with VOT values somewhere between French and English norms. Likewise, Kilpatrick found that bilinguals who started acquiring a new language as early as the age of six still produced compromised VOT values. It seems as though languages have a great impact on each other in regards to VOT. This could be due to language transfer, which occurs when the knowledge of a person’s L1 interferes with their L2. This can be either a conscious or unconscious choice, but any transfer of VOT would seem to be on the unconscious level.

5.4 SEARCH FOR UNIVERSALS IN VOT

In a later study by done by Cho and Ladefoged (1999), eighteen languages were studied and analyzed in an effort to establish universals among them. As stated earlier in this paper, fifty milliseconds is the length of time where aspiration and nonaspiration are divided. Cho and Ladefoged states that their data did not show this to be true. In fact, they reported that their data did not support any kind of “statistical clumping” (Cho and Ladefoged pg. 223).

However, Cho and Ladefoged (1999) did affirm that there were only three contrasting values of VOT: voiced, voiceless “unaspirated,” and “aspirated.” While
there are languages, like Hindi, that have four contrasting homorganic stops, most languages contain only the above mentioned three contrasting values.

Any other contrast is made by some movement of the larynx. Cho and Ladefoged suggested that VOT is definable in terms of not only the difference in time between the initial burst of a word initial consonant until the start of the proceeding vowel, but also the instigation of the laryngeal gesture which is responsible for all vocal fold vibrations. Measurements of any kind in relation to laryngeal gestures will have to be left for a study at a later date.

5.5 VOT with Japanese Subjects
A study by Tetsuo Harada (2007) examined the acoustical side of language acquisition; specifically, with L1 English speaking children in a Japanese speaking environment. The study analyzed not only the production of VOT for /p t k/ of English but also the Japanese of English-speaking children within a Japanese immersion program, but Harada was really more focused on age of acquisition. Harada’s study looked at children around the age of puberty, because studies have shown that people who begin learning a second language after puberty, what is referenced to as a late bilingual, fail to acquire native-like VOT values of that second language. Other studies like (Kilpatrick 2003) have gone even further to explain that even people who start learning an L2 before puberty often fall short of native-like VOT values.

Harada’s study found that the earlier the English speaking children were exposed to Japanese, the better able they were to distinguish VOT values for English and Japanese. This supports what Kilpatrick found in her study. According to
Kilpatrick’s data, speakers who learned both English and Spanish from infancy were successful in maintaining an average Spanish VOT of less than 30 ms. and an average English VOT of 54 ms. Alternatively, speakers who learned Spanish or English as an L2 after the age of six produced either longer VOT averages in Spanish or shorter VOT averages in English, respectively.

Harada’s study also found that the children’s VOT values varied from not only normal VOT values established from monolingual Japanese children but also normal VOT values established from monolingual English children. This suggests some type of interference occurring between Japanese and English.

While Japanese has voiced and voiceless stops just like English, Japanese /p, t, k/ are not aspirated. However there is some debate about whether this is true (Vance 1987). This is because the Japanese VOT in voiceless stops is shorter than voiceless stops in English but longer than VOT in other languages like English. A study generated by Homma (1980) measured the VOT of six Japanese disyllabic words. He found that the average VOT of his Japanese subjects in word initial /t/ to be about 25 ms. This was a much lower value when compared to the average VOT value of 70 ms that Lisker and Abramson (1964) found for monolingual English speakers.

Harada studied English speaking children living in a Japanese immersion school. The immersion school required these students to speak Japanese almost the entire time while in the school, although the children did not speak much Japanese outside of the school. The children did not intermingle with any monolingual Japanese speakers. Their teachers were bilingual L1 Japanese L2 English speakers.
The volunteers who participated in Harada’s study consisted of fifteen L1 English L2 Japanese children, five L1 Japanese L2 English bilingual adults, ten monolingual Japanese children, five monolingual Japanese adults, and five monolingual English adults. The fifteen children from the immersion program were between the ages of six and ten. The monolingual children were approximately the same age.

Harada recorded each of the participants with an audio tape recorder during twenty minute sessions. They were shown pictures designed to elicit voiceless stop tokens. The tokens consisted of word initial voiceless stops. 945 Japanese tokens along with 675 English token were elicited.

The L1 English immersion children uttered voiceless stops with longer VOT values than their immersion teachers or the monolingual Japanese children. This showed that the immersion children were not able to acquire a native-like production of VOT, but the children produced Japanese voiceless stops with significantly shorter VOT values than those of their English VOT values. This could mean that these students are making a cognitive distinction in regards to VOT between spoken English and Japanese. These changing VOT values are still not as long as the average VOT of a native Japanese child, but this could be due to interference from their L1, English. This could also be due to their immersion teachers. The data showed that their teachers produced VOT values between 35 and 50 ms, while the monolingual Japanese children produced VOT values of 20 to 30 ms. The immersion children’s input from the teachers was already significantly longer than monolingual Japanese speakers in terms of VOT. This could have had a great effect on the immersion children’s acquisition of Japanese.
5.6 VOT Study in El Paso

Cynthia Kilpatrick’s study examined VOT with the expectation of determining norms for monolingual English and Spanish speakers in El Paso. Earlier studies have presented results to conclude that monolinguals residing within a bilingual speech community produce speech differently from the speech of monolinguals living in a monolingual speech community. She wanted to determine if and to what extent these norms differed between English and Spanish in El Paso, and she wanted to know if those norms differed significantly from bilinguals. She also sought to discover whether or not VOT values become stabilized for bilinguals.

She discerned that acquiring native-like VOT for bilinguals is an enormous challenge. Moreover, she discovered that even her bilingual participants were not able to express VOT within the range of monolingual speakers, be it Spanish or English. Kilpatrick (2003) recognized that bilinguals in El Paso developed separate phonetic categories for stops of English and Spanish. This was determined because, for bilinguals, most voiceless stops in English were above 50 ms while most stops in Spanish were around 25 ms.

Kilpatrick suggested that age of acquisition played a significant role in determining how well bilinguals were able to produce native like VOT values in their L2. Her cut off age was six. Participants who learned both languages before they were six successfully maintained VOT values over 54 ms in English stops and under 30 ms in Spanish stops. Participants who learned one language after six years of age had either longer VOT values in Spanish or shorter VOT values in English. They were not able to consistently produce native-like VOT values, but they were able to
present established separate categories for stops for each language, like the participants who acquired their L2 before the age of six.

Kilpatrick’s data indicate that VOT is not at all stable and can be affected by exposure to another language. Even if someone has spoken Spanish since birth, his/her VOT values can become compromised by exposure to English through speaking or even listening. She mentions that although linguists, like Major (1992), have stated that this is due to interference; Kilpatrick (2003) suggests that this actually might be due to the mere nature of VOT. VOT is different from categorical features. It is not a question of VOT being present or not but, rather, how long the VOT is. VOT varies significantly, and attempting to duplicate one’s own VOT is a difficult task. Duplicating the VOT of another individual within the same language is next to impossible.

On the other hand, Kilpatrick states that the VOT values of her monolingual English subjects were not affected by Spanish. This seems to be difficult to understand. It seems only natural that the VOT values be affected in the same way for both sets of monolinguals, especially with the English monolinguals living in a bilingual area. She states that this could be because the most of the English monolinguals were not enrolled in any Spanish classes, unlike the Spanish monolinguals who were all enrolled in an English class, suggesting that actively learning a new language results in an added unstable situation.
Chapter 6: The Current Study

6.1 INTRODUCTION

The purpose of this linguistics research project is to determine whether or not it is easier for English speakers to acquire Spanish VOT values than it is for Spanish speakers to acquire English VOT values. For Spanish speakers /p t k/ are short lag voiceless stops and are perceived as voiceless. If a /p/, /t/, or /k/ were to be pronounced with a long lag in Spanish by accident or for any other reason, the phonemes would still be perceived as voiceless.

However for English speakers, word initial stops with long lag are perceived as voiceless and short lag stops are perceived as voiced (Kilpatrick 2003). Native English speakers learning Spanish need only “manipulate the allophonic and phonemic status of the stops” in order to produce short lag word initial voiceless stops. (Kilpatrick pg. 76).

As demonstrated in the following word pairs, short lag stops /b d g/ contrast with [pʰ tʰ kʰ] in initial position.

\[
\begin{align*}
\text{pig} & \hspace{1cm} \text{teal} & \hspace{1cm} \text{cap} \\
\text{big} & \hspace{1cm} \text{deal} & \hspace{1cm} \text{gap}
\end{align*}
\]

According to Kirkpatrick (2003) voiceless stops in medial position in an unstressed syllable have short VOT. However, native English speakers can still distinguish flappy from flabby, where the former contains a voiceless, unaspirated stop and the latter contains a voiced stop. Kirkpatrick (2003) suggests that English speakers may have already established the perceptual boundaries between an

\[3\] All statements about perception in this paper are anecdotes and conjectures, as neither Kilpatrick nor I have done any direct testing of perception. Speakers’ perceptions can only be inferred by using production data. This type of inference has validity in that speakers/acquirers produce what they perceive.
unaspirated voiceless stop and a voiced stop, and therefore need only to transfer this perception to word initial position.

Meanwhile, a Spanish speaker who hears a word initial stop followed by a short lag could perceive the stop to be voiceless, as it would be in Spanish. That Spanish speaker must learn to not only perceive a new set of long VOT stops, but must be able to produce them. As stated in an earlier section, English phonemes /p t k/ are not contrastive with [ph th kh]. While the difference between these two sets is not in meaning, there is a difference in terms of perception. In Spanish, [ph th kh] do not exist, so Spanish speakers should not need to perceive anything different from short VOT stops. A Spanish speaker must learn to perceive [ph th kh] as different from /p t k/. The differences are shown in Figure 1.

![Figure 1: VOT Comparison and Contrast](image-url)
Thus this study was initiated to examine the validity of the premise that native English speakers have an advantage over native Spanish speakers when it comes to acquiring new production boundaries in terms of word initial voiceless stops. This is pertinent when it comes to dealing with English speakers learning Spanish and Spanish speakers learning English.

6.2 THE SUBJECTS

Forty people were chosen to participate in the study at hand, with ten participants in each group. In an effort to establish some kind of homogeneity, all subjects were between the ages of eighteen and thirty, and a strong effort was made to use an equal number of males and females. Most subjects were students at the University of Texas at El Paso. All other subjects were either former UTEP students or were students at other universities, thus everyone had at least a high school education. All subjects were on various levels of language ability and had grown up with varying degrees of either English and/or Spanish.

6.3 MONOLINGUAL ENGLISH SPEAKERS

El Paso is very rich in Hispanic culture and includes many bilinguals who have spoken English and Spanish almost their entire life (Kilpatrick 2003). According to the U.S. Census Bureau, of the 678,346 people above the age of five, living in El Paso County, 485,660 people, or approximately 72 percent, speak Spanish at home. Usually, people from El Paso have friends who expose them to Spanish. Generally speaking, there is almost no way to avoid Spanish in El Paso. Street signs and
billboards are oftentimes in Spanish. In order to avoid such Spanish influence, nine of the ten monolingual English speakers chosen resided in London, East Texas, Utah and Kentucky. The tenth subject was a mother of four children who spends most of her time at home here in El Paso. One subject from Houston and four subjects from Utah had studied some Spanish in high school. However, they did not remember much, if any, of the Spanish that they studied. None of these subjects was able to understand or utter more than a couple of words in Spanish. Their ages ranged from eighteen to twenty-eight.

6.4 Monolingual Spanish Speakers

Due to the current level of violence in Mexico, finding true monolingual Spanish speakers in the age range from eighteen to thirty was next to impossible. Researchers from UTEP were not allowed to travel into Mexico to elicit data. Oftentimes, people of the older generations would be monolinguals, but most Hispanics in the age range eighteen to thirty have had a significant exposure to English. Although in Juarez, English as a foreign language is mandatory for every year in high school; many students do not take full advantage of the opportunity to learn English.

All monolingual Spanish speakers were chosen because they were students who grew up in Mexico and had limited abilities in English. Most of the students chosen were taking ESOL 1610, which is a low level English for non-native speakers. The other two subjects not enrolled in ESL classes were people whose communicative ability in English was very poor. None of the ten subjects could carry on a conversation in English. Although, a few of the subjects rated themselves fairly high
on speaking and listening ability, they still were not able to understand simple information given them about the study. Their ages ranged from eighteen to twenty-five.

6.5 L1 English L2 Spanish Bilinguals

All participants were chosen based upon four criteria. First, they needed to be from a predominately Hispanic area. Second, they needed to have a good command of spoken Spanish. Third, they were not to be conversant in a third language. Finally, participants were not chosen if they were balanced bilinguals. It was important to find native English speakers accustomed to producing long VOTs in English, who then had to learn to produce shorter VOTs in Spanish. Nine of these participants were students at UTEP. The tenth participant was a student at Brigham Young University who grew up in a similar Hispanic dominant area to El Paso. All ten participants had a good command of spoken Spanish, but some had better command than others. English ability was determined through a short conversation before testing. Effort was made to find subjects that were close in Spanish competency. Students from the higher level Spanish for non-native speakers were recruited but to no avail. Prospective participants were asked a few questions in Spanish in order to show competency. They were also asked to self-describe/rate/evaluate their level of Spanish proficiency. If they felt that their Spanish was either really strong or really weak, they were not asked to participate.
6.6 **L1 Spanish L2 English Bilinguals**

The same criteria applied for native Spanish bilinguals. As with the native English bilinguals, these native Spanish bilinguals were all from a Hispanic area. All ten participants had a good command of spoken English. None of the participants were conversant in a third language. None of them were balanced bilinguals. Every participant grew up speaking only Spanish. They all learned to speak English after they had grown out of early childhood, thus making them late bilinguals as well. All Spanish native bilinguals were students attending UTEP. Most were either in one of the final classes of the ESL program at UTEP or had recently finished the program.

6.7 **Tokens**

Eighteen English tokens and eighteen Spanish tokens were embedded into two texts respectively. A token in this study is a word consisting of an initial voiceless stop /p t k/ followed by a vowel. There were six tokens of each type of voiceless stop. Occasionally, some tokens were either inaudible, due to soft speech, or unreadable, due to outside interference, such as coughs or noises from the heater. The last ten seconds of one subject’s speech was not recorded due to equipment failure caused by a low battery. Such tokens were not included into the analysis. In total, monolingual speakers were given eighteen tokens, while bilingual speakers were given thirty-six tokens. Eliminating any bad tokens that did not record correctly, forty-nine in total, there were a total of 1031 tokens.

6.8 **Procedures**

All participants were told that a study would be undertaken to investigate how speakers perform when speaking a second language. After each participant read and signed a consent form, they were given a questionnaire to help show how they
learned their second language and how comfortable they felt in their second language. Once the paperwork was finished, participants were given a text that contained the eighteen tokens. Monolingual English speakers read only the English text, while the monolingual Spanish speakers read only from the Spanish text. Both sets of bilinguals read from both of the Spanish and English texts. Their speech was recorded onto a computer using a microphone. The recorded speech was then uploaded into a software program called Praat.

Praat was designed by Paul Boersma and David Weenink. Simply stated, this program is used to analyze speech. It can do phonetic analysis as well as sound manipulation, and it supports speech synthesis, which is used by many people, such as Stephen Hawking⁴, It is through this program that VOT values can be determined. Once a recording was finished, each token was studied to determine the VOT of the initial voiceless stop.

The VOT value was determined by focusing in on the actual token in Praat. By zooming in to about 20 milliseconds, the wave lengths from the initial burst of the voiceless stop until the start of the vowel sound became clear. Once these wavelengths were in view, the exact time of the initial burst until the start of the vowel was highlighted. This identified the exact voice onset time. VOT measurements were taken to the nearest millisecond. The initial burst was evidenced by a sudden increase in wavelength, and the vowel was evidenced by a sudden noticeable pattern in the wavelengths. After the measurements were complete, overall mean values were calculated for each group of participants, as well as mean values as they pertain to

⁴ Stephen Hawking is a famous physicist confined to a wheelchair due to illness. He has next to no mobility and is only able to speak using speech synthesis.
place of articulation and following vowel features, see Figures 2 and 3. Note the
difference in time between the initial burst and the onset of voicing within the two
figures.

Figure 2: Waveform and spectrograph of English /tu/
Figure 3: Waveform and spectrogram of Spanish /ta/

6.9 STUDY LIMITATIONS

As is the case with most research projects, a number of unforeseen problems surfaced at the outset and continued as the study progressed. The first problem entailed finding qualified participants who were equally proficient in their L2 languages. The highest hurdle was to find participants of the same proficiency in L1 English L2 Spanish. Spanish students who qualified into upper level Spanish courses were recruited with no success. Lacking satisfactory incentives, these students did not wish to make themselves available for the study. Otherwise, potential willing participants either spoke Spanish too fluently or were not able to hold a conversation in Spanish. Thus, participants who were not quite fluent in Spanish as well as
students who were barely able to hold a conversation in Spanish were asked to participate. This discrepancy created a somewhat uneven proficiency level for L1/L2 participants.

Technical difficulties with the vocal recordings created additional problems with primary data development. A buzz from the microphone interfered with one sampling of recordings making it difficult to discern the specific VOT values. Likewise, the buzzing created extra waveforms that sometimes caused the initial burst of consonants to be undetectable. Also, there were occasions when other outside noises, e.g. coughing and sneezing, made the discernment of VOT nearly impossible.

Finally, the actual measuring of VOT was problematic for two reasons. First, velar stops, /k/, can often show two separate bursts (Thomas 2011). My measurements were calculated using the first burst, although later it was made known that Cho and Ladefoged (1999) measured VOT using the second burst (Thomas 2011). Second, according to Thomas (2011), aspiration regularly makes the onset of the following vowel ambiguous. In the present study, the onset of the vowel was determined to be where the vocal vibrations begin to show a pattern.
Chapter 7: Results

7.1 Introduction

VOT values were calculated using percentages and averages. No other statistics outside of these were used in this study\(^5\). These percentages and averages showed certain patterns developing for each group of participants. While the average VOT for monolingual Spanish speakers was 23 ms., the average VOT for monolingual English speakers was 75 ms. The L1 Spanish L2 English bilinguals had a VOT average of 26 ms. in Spanish and 52 ms. in English. The L1 English L2 Spanish bilinguals show an average of 64 ms. in English and 29 ms. in Spanish.

It was necessary to divide the tokens into smaller groups in order to analyze separate features. In addition to overall average for each participant, the tokens were divided by initial consonant: velar, alveolar, and bilabial. Kilpatrick (2003) maintained that a study on VOT would not be complete without considering vowel height. So in an effort to follow this notion, tokens were also divided not only by and +low and –low vowels, but also by +back and –back vowels.

\(^5\) Due to a lack of statistical knowledge as well as time constraints, it was determined to forego any advanced statistical analyses.
Table 1: Overall Averages for each group (ms)

<table>
<thead>
<tr>
<th></th>
<th>English values (ms)</th>
<th>Spanish values (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monolingual Spanish</td>
<td>n/a</td>
<td>23.6</td>
</tr>
<tr>
<td>L1 Spanish L2 English Bilingual</td>
<td>52.7</td>
<td>25.8</td>
</tr>
<tr>
<td>Monolingual English</td>
<td>73.7</td>
<td>n/a</td>
</tr>
<tr>
<td>L1 English L2 Spanish Bilingual</td>
<td>63.8</td>
<td>29.7</td>
</tr>
</tbody>
</table>

7.2 **Monolingual Spanish speakers**

Monolingual Spanish speakers recorded an average VOT fairly similar to those of the study by Kilpatrick (2003), her VOT averages being 5 ms shorter than these. The overall average for monolingual Spanish speakers is 23 ms. While the lowest VOT was 11 ms, the highest VOT was 46 ms. 46 ms is higher than the accepted averages of this group, but this could be due to participants living in a bilingual area as well as to the nature of the study. The data could be skewed somewhat since all the data comes from readings from the participants instead of from natural conversation.

When the tokens were divided in accordance with initial consonant, the data revealed /k/ to contain the longest VOTs, followed by /t/ and /p/ respectively. The average VOT for /k/ initial tokens is 29 ms for this group, with a range from 21 to 39 ms. The average for /t/ initial tokens is 19 ms, with values ranging from 15 to 27 ms. The average /p/ initial tokens is also 19 ms, with values ranging from 14 to 28 ms.
The shortest VOT values for monolinguals were for /p,t,k/ preceded by –back vowels, 20 ms. Conversely the longest VOT values were preceding +back vowels, 26 ms. +Low and –low vowels showed less variation.

Table 2: Averages for /p/, /t/, /k/ (ms)

<table>
<thead>
<tr>
<th></th>
<th>English Values (ms)</th>
<th>Spanish Values (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>/p/</td>
<td>/t/</td>
</tr>
<tr>
<td>Monolingual Spanish</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>L1 Spanish L2 English Bilinguals</td>
<td>49.9</td>
<td>56.6</td>
</tr>
<tr>
<td>Monolingual English</td>
<td>72.2</td>
<td>78.8</td>
</tr>
<tr>
<td>L1 English L2 Spanish Bilingual</td>
<td>61.5</td>
<td>67.7</td>
</tr>
</tbody>
</table>

Table 3: voiceless stops preceding - Low vowels vs. + Low vowels (ms)

<table>
<thead>
<tr>
<th></th>
<th>English Values (ms)</th>
<th>Spanish Values (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-Low</td>
<td>+Low</td>
</tr>
<tr>
<td>Monolingual Spanish</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>L1 Spanish L2 English Bilinguals</td>
<td>51.8</td>
<td>52.7</td>
</tr>
<tr>
<td>Monolingual English</td>
<td>73.1</td>
<td>74.3</td>
</tr>
<tr>
<td>L1 English L2 Spanish Bilingual</td>
<td>60.1</td>
<td>66.9</td>
</tr>
</tbody>
</table>
### Table 4: back vowels vs. –back vowels (ms)

<table>
<thead>
<tr>
<th></th>
<th>English Values (ms)</th>
<th>Spanish Values (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+back</td>
<td>-back</td>
</tr>
<tr>
<td>Monolingual Spanish</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>L1 Spanish L2 English Bilinguals</td>
<td>52.4</td>
<td>52.3</td>
</tr>
<tr>
<td>Monolingual English</td>
<td>74.3</td>
<td>73.1</td>
</tr>
<tr>
<td>L1 English L2 Spanish Bilingual</td>
<td>68.1</td>
<td>60.6</td>
</tr>
</tbody>
</table>

### 7.3 **Monolingual English Speakers**

The data from monolingual English speakers showed an average VOT of 74 ms. These values were about 4 ms longer than those in the study by Kilpatrick (2003). VOT values for monolingual English speakers ranged from 65 ms to 95 ms. One participant who had a VOT average 95 ms, because he was reading his passage slowly, was asked to read at a natural speaking pace. He replied that that was his natural speaking pace. After conversing with him, it was evident that he was telling the truth.

The average VOT for /k/ initial tokens is 72 ms for this group, with a range from 52 to 89 ms. The average for /t/ initial tokens is 79 ms, with values ranging from 63 to 103 ms. The average /p/ initial tokens is 72 ms, with values ranging from 52 to 99 ms. This contrasts from the general rule that aspiration with /k/ initial words is longer than both /p/ and /t/ initial words.

---

6 The difference of 4 ms. is most likely insignificant, but without doing a statistical analysis, it is hard to be completely sure.
Monolingual English speakers showed almost no variation with respect to vowel environment, unlike the L1 English L2 Spanish bilinguals.

### 7.4 L1 Spanish L2 English Bilinguals

The results from these bilinguals showed a wide range of averages. With respect to English, the lowest VOT average was 26 ms. while the highest was 75 ms. The participant with the highest VOT average in English also produced the highest VOT average in Spanish, 30 ms. The lowest average in Spanish for these bilinguals was 18 ms. The overall average for this group was 26 ms. in Spanish and 52 ms. in English.

Concerning English, the average VOT for /k/ initial tokens was 58 ms, with a range from 41 to 90 ms. The average for /t/ initial tokens was 56 ms, with values ranging from 21 to 78 ms. The average /p/ initial tokens was 42 ms, with values ranging from 16 to 71 ms. The ranges for these bilinguals are substantial.

With respect to Spanish, the average VOT for /k/ initial tokens was 34 ms, with a range from 25 to 40 ms. The average for /t/ initial tokens was 23 ms, with values ranging from 15 to 28 ms. The average /p/ initial tokens was 20 ms, with values ranging from 13 to 28 ms.

These bilinguals showed no variation with different vowel environments except for +back and –back vowels when speaking Spanish. The difference was 32.3 ms and 26.4 ms. respectively.
7.5 L1 English L2 Spanish Bilinguals

The VOT values for this group of bilinguals were less varied. The values ranged from 21 to 38 ms. with regard to the Spanish tokens. In English, their averages varied from 55 ms. to 68 ms. It is interesting to note that the ranges in VOT for these bilinguals were similar in both their Spanish and English, while the ranges of L1 Spanish L2 English bilinguals were especially different between their English and Spanish. The overall average for this group was 29 ms in Spanish and 64 ms in English.

Pertaining to English, the average VOT for /k/ initial tokens was 64 ms. for these bilinguals, with a range from 59 to 74 ms. The average for /t/ initial tokens was 68 ms, with values ranging from 60 to 79 ms. The average /p/ initial tokens was also 61 ms, with values ranging from 53 to 79 ms.

L1 English L2 Spanish bilinguals did not vary as much in range in their L2 as did the L1 Spanish bilinguals. The L1 English bilinguals had an average VOT for /k/ initial tokens as 34 ms, ranging from 25 to 42 ms., with respect to their Spanish tokens. The average for /t/ initial tokens was 25 ms, with values ranging from 19 to 30 ms. The average /p/ initial tokens was 28 ms, with values ranging from 14 to 44 ms.

When comparing vowel environment, +back vowels yielded the longest average time at 68.1 ms, and –low vowels yielded the shortest average time at 60.1 ms. Vowel environment for their Spanish values was almost as diverse as their English values ranging from 26.4 ms up to 32.3 ms.
Chapter 8: Discussion and Conclusion

8.1 DISCUSSION

The results convey a number of important particulars occurring with both L2 English and L2 Spanish language learners. First, as both groups of learners become more proficient in their L2, their VOT values improve in accordance with their L2. For instance, the L2 Spanish learners are better able to decrease their VOT values as they become more proficient in Spanish. Conversely, the L2 English learners are better able to increase their VOT values as their English proficiency increases.

Just like the results reported by Kilpatrick (2003), these results show that acquiring native VOT values generally equal to monolingual speakers is nearly impossible. This is not to say, though, that bilinguals are not able to distinguish between their native language and their L2, be it English or Spanish. The overall average VOT of L1 English L2 Spanish bilinguals is 64 ms in English and 29 ms in Spanish. This is a difference of 35 ms. The overall average for L1 Spanish L2 English is 26 ms in Spanish and 52 ms in Spanish, with a difference of 26 ms.

An interesting result is the compromised VOT values of both groups of bilingual speakers. The native Spanish bilingual speakers exhibited VOT values that are about 2 ms, or 9 percent longer than those of monolingual Spanish speakers. On the other hand, native English bilinguals produced VOT values 10 ms shorter than those of monolingual speakers, or a difference of 14 percent.

The L1 Spanish bilinguals produced VOT values that were almost 5 ms longer than those produced by Spanish monolinguals with words containing an initial /k/.

This is approximately 12 percent longer. Initial /p/ words yielded a 17 percent (4 ms)
increase in VOT, while initial /t/ words did not vary much between monolingual and bilingual.

When uttering words beginning with /k/, L1 English bilinguals produced VOT values 8 ms, or 11 percent shorter than those of monolingual speakers. Initial /t/ words yielded a reduction of 14 percent (9 ms), and average VOT values of initial /p/ words were 15 percent (10.5 ms) shorter when produced by bilinguals compared to monolingual English speakers.

Figure 4: Averages for /ptk/ (ms)

Following Kilpatrick’s maintenance that a study on VOT would not be complete without regarding vowel height, such was considered. -Low vowels included were /i/, /u/, /o/, and /e/, while the only low vowel was /a/. Monolingual Spanish speakers showed a 12 percent difference between low and minus low, with -low vowels averaging 23.3 ms, and low vowels averaging 20.7 ms. Spanish bilinguals showed
virtually no difference between the two sets of vowels—25.8 ms for -low and 25.9 ms for +low vowels in English. In the bilinguals’ English set, there was a slightly larger difference when comparing -low and +low respectively, 51.8 ms to 52.7 ms, a difference of 1.7 percent.

This study took vowel environment one step forward by comparing –back vowels and +back vowels. The monolinguals showed a difference between –back and +back vowels with 26 ms and 20 ms respectively or a 23 percent difference. The bilinguals showed a difference of –back and back vowels with 28ms. and 24.2 ms. respectively or a difference of 13.6 percent.

Perhaps the largest difference came between English monolinguals and bilinguals with respect to +back and –back vowels. Monolinguals showed next to no difference with 74.3 ms for back vowels and 73.1 ms for –back vowels or 1.6 percent. The bilinguals, and the other hand, produced an average VOT of 68.1 ms for back vowels and 60.6 ms for –back vowels or a difference of 12.4 percent. There is a relatively large disparity between bilinguals and monolinguals here, and are apparent by looking at figure 5.
Individually, there is a large amount of variation within the bilingual groups when the L2 is spoken. The group with more variation was the L1 Spanish L2 English speakers in English. The lowest VOT was 26 ms, and the highest was 75 ms. 75 ms was higher than any VOT value average for L1 English L2 Spanish bilinguals and most English monolinguals. The variation within the L1 English L2 Spanish bilinguals VOT values was less, and their VOT varied from 21 ms to 38 ms.

Kilpatrick claimed that her native Spanish subjects, when asked if initial voiceless stops were different in English and Spanish, said that they were equal. This was not the case with subjects from this study. Bilingual, native Spanish subjects claimed a clear difference existed between initial voiceless stops in Spanish and English. This suggests a conscious understanding of a difference, even if these voiceless stops are not produced equally every time.
Another claim made by Kilpatrick was that the VOT values of her monolingual English subjects were not affected by Spanish. VOT values of L1 English L2 Spanish bilinguals were up to 15 percent shorter than those of monolingual English speakers. Compare this to the L1 Spanish bilinguals whose VOT values were up to 17 percent longer than those of monolingual Spanish speakers. This suggests that there is a fairly even amount of interference for both groups of bilinguals.

Do L1 Spanish L2 English bilinguals show an ability to acquire English VOT values approximately equal to L1 English L2 Spanish bilinguals acquiring Spanish VOT values? This question still cannot be answered completely. By comparing the percentages of all bilinguals and monolinguals, there appears to be a slight advantage for English speakers learning Spanish. In order for this to really be determined, a study containing a larger sample size would be needed.

8.2 CONCLUSION

In keeping with Kilpatrick’s findings, this study likewise suggests that bilinguals understand that a difference in VOT exists between English and Spanish. Most bilingual participants were not able to consistently produce native-like VOT values, but both groups of bilinguals showed a noticeable variance between English VOTs and Spanish VOTs.

Cynthia Kilpatrick’s study examined VOT with the expectation of explaining what VOT norms are in El Paso and how they differ, not only between English and Spanish, but between monolinguals and bilinguals. She did not take into account dividing the bilingual group into two separate groups. She determined that bilinguals failed to produce VOT values within the range of monolingual speakers of either
English or Spanish. What she failed to present was whether or not L1 English bilinguals or L1 Spanish bilinguals would be more successful in producing native-like VOT values in their L2.

There are pedagogical implications involved. With both English and Spanish speakers showing difficulty acquiring native-like VOT values in their L2, extra class time should be spent helping students with their VOT acquisition. This can be done by pronunciation practice, which is not as common as fluency practice is in a Speaking class.

The results from the present study, as well as the results of Kilpatrick's study suggest that bilinguals establish separate categories for stops for English and Spanish. If this is true, is there an approach to teaching second language learners how to better establish these separate categories though the teaching of pronunciation? Although, if native English or Spanish speakers prove able to acquire native Spanish or English-like VOT easily, then teaching them how to produce those VOTs would not be as advantageous. The data from this study suggests that there could be an advantage to teaching VOT pronunciation to both English and Spanish learners. This study suggests that there is not a true advantage for either sets of bilinguals, because the data do not seem to show significant contrast among the groups of participants. Not only did this study utilize a small sample of bilinguals and monolinguals, no rigorous statistics were used to analyze the data. The data from the present study, without any direct test of perception versus production, has shown differences among all four groups of participants. The question now is, are these differences really significant, and if so, to what extent? If a study were to be done using a more substantial amount of participants and generating more statistics
to show how different are those differences, then a more conclusive analysis could be completed. Also, another study could be performed that follows an interview style sociolinguistic interview. Interviewing subjects and asking them about strong emotional memories would elicit much more natural speech, as opposed to only having participants read various selections. This approach would be more time consuming. First, a researcher must design the interview by devising an assortment of questions that could possibly elicit natural data. There will be trial and error as many of the questions will most likely not be effective. Once the questions have been created, the interview can take place. All utterances need to be recorded and transcribed. Every instance of every voiceless stop should be taken into account. Once all transcribing finishes, the data can be analyzed using Varbrul analysis, or variable rule analysis. There are software programs available for this. The hope is that a study such as this one would either give evidence to the suggestion that neither group of bilinguals has any advantage or that native English speakers really do have an advantage over native Spanish speakers.

As Kilpatrick (2003) explained, El Paso is a good area to study bilingual speakers of English and Spanish, because there is such a large population of bilingual speakers in the El Paso and Juarez area. However, El Paso is not a good place to study monolinguals of either English or Spanish, because there is too much inference of both Spanish and English on each other. Any future studies should attempt to find participants who are true monolinguals not living in a bilingual community.

In doing this study, a few questions have surfaced. Does VOT differ between geographic or national regions speaking the same language? Would someone from
Spain produce VOT significantly different from someone from Venezuela? Likewise, does VOT differ between bilinguals who have learned their second language in different locales? Does a community or a formal study of language affect compromised VOT values? These questions are worth further study.

In completing this study, the difficulty for bilinguals in producing authentic VOT values in their L2 has become apparent. We have seen that bilinguals’ L1 can be influenced by their L2, independent of whether that L1 is English or Spanish. It is also apparent that both L1 Spanish and L1 English bilinguals produce voiceless stops with longer VOT in English and shorter VOT in Spanish.
References


Labov, William. Sociolinguistic Patterns. Philadelphia: University of


Appendix 1
Native English Speakers

Please complete following questionnaire to the best of your ability.

Name ___________________________________________ Age ________________________________

Hometown (City, State) ______________________________________________________________

What is your mother’s native language? ________________________________________________

What is your father’s native language? ________________________________________________

Have you studied Spanish? If yes, how old were you when you began to learn/study Spanish?
_____________________________________________________________________________________

In what contexts did you learn Spanish?

(check all that apply)

__ Home/family   __ Work

__ School   __ Media (TV/radio/internet/newspaper)

__ Friends

Please rate your abilities in Spanish in the following areas by circling the best choice:

Reading: Excellent Good Average Poor

Writing: Excellent Good Average Poor

Speaking: Excellent Good Average Poor

Listening: Excellent Good Average Poor

Which language do you speak most of the time at home? _________________________________

With friends? ______________________________________________________________________

Which language (English or Spanish) do you prefer to use? ______________________________
Appendix 2

Hispano hablantes de nacimiento

Por favor llene el siguiente cuestionario lo mejor que pueda.

¿Qué idiomas usaban sus padres o tutores para hablar con usted durante su niñez y adolescencia?

¿Ha estudiado inglés en la escuela? Sí, sí, ¿Qué edad tenía cuando comenzó a aprenderlo y estudiarlo?

¿En qué contextos(s) aprendió este idioma? (favor de marcar todos los que sean relevantes)

__ Casa/amigos    __ Trabajo
__ Escuela       __ Medios informativos (TV/radio/Internet/etc.)
__ Amigos

Favor de evaluar sus habilidades en el idioma inglés seleccionando la mejor opción.

Lectura: Excelente Bueno Regular No muy bien
Escritura: Excelente Bueno Promedio No muy bien
Habla: Excelente Bueno Promedio No muy bien
Escucha: Excelente Bueno Promedio No muy bien

En casa, ¿en qué idioma habla?

¿Con amistades y amigos?

¿Qué idioma prefiere usar, inglés o español?
Appendix 3

English Text

Two boys named Tim and Pat decided to play by the pond next to the park. After playing for a while, Pat, not knowing how to swim, fell into the pond. Without hesitating, Tim pulled Pat out of the pond. All the kids came to see the commotion. They had never seen such a performance out of Tim. He was usually too busy eating tacos and pizza all day. Pat’s mom, Tammy, was so thankful that she gave Tom a coke, candy and a piece of blue cake, which stained his teeth.

Spanish Text

Hace mucho fui camping con mis amigos y nos encontramos con un pueblo. Ese pueblo era hermoso porque nos recordaba a los viejos tiempos. No habían coches y el único medio de transporte es el caballo. Al explorar el viejo pueblo, nos dio hambre. Llegamos a un puesto de comida y pedimos cinco ordenes de tacos y de postre cinco piezas de pan dulce con tazas de te. Después entramos a una tienda donde vendían antigüedades. Ahí vi una hermosa tela y la compré para mandarle hacer un vestido a mi mamá. Aunque no fuimos camping, nuestro día fue todo un éxito.
Appendix 4
Informed Consent

University of Texas at El Paso (UTEP) Institutional Review Board
Informed Consent Form for Research Involving Human Subjects

Title: Resetting VOT in a Bilingual Region

Principal Investigator: Dan Morgan

UTEP: Languages and Linguistics

In this consent form, “you” always means the study subject or participant.

We ask you to read the following text so that you are informed of the nature of this research study and what is expected of you if you participate in it. Federal regulations require written informed consent. If you want to participate, signing this form will indicate that you understand what this research study is about and that you want to participate in it. It also means that you decided to participate in a free and informed manner.

Introduction
We are asking you to take part voluntarily in the research project described below. Please take your time making a decision and feel free to discuss it with your friends and family. Before agreeing to take part in this research study, it is important that you read this text. Please ask the researcher team to explain any words or information that you do not clearly understand.

Why is this study being done?
This study is being done to investigate how speakers perform when speaking a second language. The plan is to test 40 adults, including 20 native English speakers and 20 native Spanish speakers.

You are being asked to participate in the study because you are an adult aged 18 years to 30 years, and because your first language is either Spanish or English.

If you decide to enroll in this study, your involvement will last approximately 10 minutes.

What is involved in the study?
You will read a paragraph, and your speech will be recorded on the computer. The speech will be later analyzed. Once analyzed, any recordings of your speech will be destroyed.

What are the risks and discomforts of the study?
There are no known risks associated with this research.
What will happen if I am injured in this study?

Participation in this project involves no more risks than those associated with daily activities.

Are there benefits to taking part in this study?

There will be no direct benefits to you for taking part in this study. However, your participation will further our understanding of how Spanish and English language learners are able to deal with sounds that are different from their own language.

What other options are there?

You have the option not to take part in this study. There will be no penalties involved if you choose not to take part in this study. If after starting to participate, you do not want to continue participating for any reason, the session is discontinued. You are allowed to stop participating at any time.

Who is paying for this study?
The study has no external financial support.

What are my costs?
There are no direct costs to participate.

Will I be paid to participate in this study?
You will not be paid for taking part in this research study.

What if I want to withdraw from this study?

Taking part in this study is voluntary. You have the right to choose not to take part in this study. If you do not take part in the study, there will be no penalty. If you choose to take part, you have the right to stop at any time. However, we encourage you to talk to a member of the research group so that they know why you are leaving the study.

Who do I contact if I have questions or problems?
You may ask any questions you have now. If you have questions later, you may email me at dwmorgan@miners.utep.edu
If you have questions or concerns about your participation as a research subject, please contact the Institutional Review Board Office at (915) 747-8841 or irb.orsp@utep.edu.

What about confidentiality?
1. Your part in this study is confidential. None of the information will identify you by name.

Your responses will be automatically saved to a password protected data file on the computer so that we can review the data later. Once the data has been reviewed, the data will be destroyed, so no recordings will be set aside for later use.

Records are confidential, and participants' names (first and last) are removed when any data are discussed. Your responses will be identified by a code number and never by your name. Please inform us if at any point you decide you do not want us to use your data in any form; otherwise, we will assume that we have your permission to use them.

Research participants in language experiments tend to talk to experimenters and have conversations with them. If participants provide other personal information during the session in addition to their responses to the computer, that information will also be kept confidential and be treated in the same way of the rest of the experimental data.

We would very much like to have you participate in our study. If you want to participate, please complete and sign the attached authorization statement.

Thank you. We very much appreciate your participation.

This research is being conducted by Dan Morgan in fulfillment of a thesis requirement within the department of Languages and Linguistics at UTEP.

Authorization Statement

I have read each page of this document explaining the study. I know that being in this study is voluntary and I choose to be in this study. I know I can stop being in this study without penalty. I will get a copy of this consent form now and can get information on results of the study later if I wish.

Participant Name: ________________________________________    Date:  __________________

Participant Signature:  __________________________________________

Consent form explained and debriefing conducted by:   _______________________________  

________________________________________ (signature)
Appendix 5

Panel de Revisión Institucional (Institutional Review Board) de la Universidad de Texas en El Paso (UTEP)

Forma de Autorización para Investigaciones que Involucran Sujetos Humanos.

<table>
<thead>
<tr>
<th>Título del Protocolo:</th>
<th>Cambio de sincronización del inicio de sonoridad entre hablantes bilingües</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigadora Principal:</td>
<td>Dan Morgan</td>
</tr>
<tr>
<td>UTEP:</td>
<td>Departamento de Lenguas y Lingüística (Languages and Linguistics)</td>
</tr>
</tbody>
</table>

En esta forma de consentimiento, “usted” se refiere siempre al sujeto de estudio (el participante).

Le pedimos que lea el siguiente texto para que se informe sobre la naturaleza de esta investigación y sobre qué se espera de usted si participa en ella. Las leyes Federales de los Estados Unidos requieren que le demos una forma de consentimiento por escrito. Si quiere participar, firmar esta forma de consentimiento indicará que entiende de qué se trata esta investigación y que quiere participar en ella. También querrá decir que decidió participar de manera libre e informada.

Introducción

Le pedimos que participe voluntariamente en la investigación descrita abajo. Por favor, tómese el tiempo que necesite para tomar una decisión y discútala con sus amigos y parientes si quiere. Antes de aceptar participar en esta investigación, es importante que lea este texto. Por favor, pídale al equipo de investigación que le explique cualquier palabra o información que no entienda con claridad.

¿Por qué se está haciendo este estudio?

El propósito de la investigación es analizar cómo los hablantes tratan de desempeñarse en un segundo idioma. El plan pone a prueba a 40 adultos, incluyendo 20 hablantes de Inglés y 20 hablantes de Español.

Se le pide su participación en esta investigación porque es un adulto entre 18 y 30 años y porque su primer idioma es ya sea el Inglés o el Español.

Si decide inscribirse en la investigación, su participación durará aproximadamente 10 minutos.

¿Qué pasaría si me hago daño durante la investigación?

Esta investigación no involucra más riesgo que aquellos asociados con actividades diarias.

¿Hay beneficios por participar en esta investigación?

55
No habrá beneficios directos para Ud. por participar en esta investigación. Sin embargo, su participación nos ayudará a entender el conocimiento del cambio de sincronización del inicio de sonoridad entre hablantes bilingües.

¿Qué otras opciones hay?

Ud. tiene la opción de no participar en esta investigación. No sufrirá ninguna penalidad si decide no participar en esta investigación. Si después de empezar a participar, no disfruta el juego o quiere dejar de participar por cualquier razón, la sesión se acaba. Puede parar de participar cuando quiera.

¿Quién paga esta investigación?

El estudio no tiene ningún apoyo financiero externo.

¿Cuáles son mis costos?

No hay ningún costo directo.

¿Me pagarán por participar en esta investigación?

No se le pagará por participar en esta investigación.

¿Qué pasa si quiero salirme o se me pide que me salga de la investigación?

Participar en esta investigación es voluntario. Tiene el derecho de decidir no participar en esta investigación. Si no participa en esta investigación, no sufrirá ninguna penalidad.

Si decide participar, tiene el derecho de parar en cualquier momento. Sin embargo, le pedimos que hable con un miembro del equipo de investigación para que sepan por qué está dejando la investigación.

¿A quién llamo si tengo preguntas o problemas?

Puede preguntar lo que quiera ahora. Si tiene preguntas después, puede escribir a Dan Morgan, Investigador Principal, un mensaje electrónico a dwmorgan@miners.utep.edu.

Si tiene preguntas o está preocupado/a por su participación en la investigación, por favor contacte el Institutional Review Board (IRB) en UTEP al (915-747-8841) o por correo electrónico a irb.orsp@utep.edu.

¿Qué pasa con la confidencialidad?

Su participación en esta investigación es confidencial. Ninguna información lo/la identificará por su nombre.

Sus respuestas serán guardadas automáticamente en un archivo de datos en la computadora para que podamos revisarlos hasta lo terminamos. Después de terminar, sus datos serán destruidos.

Los récordes son confidenciales, y los nombres y apellidos de los participantes son removidos cuando se discute cualquier dato. Sus respuestas serán identificadas por un código y nunca por su nombre o apellido. Por favor infórmenos si en cualquier momento Ud. decide que no quiere que usemos sus datos en alguna forma; de lo contrario asumiremos que tenemos su permiso para usarlos.

Los participantes en investigaciones sobre el lenguaje tienden a hablar con los investigadores y a conversar con ellos. Si los participantes nos dan información personal durante las sesiones más allá de sus respuestas en la computadora, esa información también será confidencial y la trataremos igual que el resto de los datos de esta investigación.
Nos gustaría muchísimo que participara en nuestra investigación. Si quiere darnos su consentimiento, por favor complete y firme el documento de autorización adjunto.

Muchas gracias. Estamos muy agradecidos por su participación.

Documento de Autorización

He leído cada página de este documento acerca de la investigación. Sé que participar en esta investigación es voluntario y decidí participar en esta investigación. Sé que puedo parar de participar en esta investigación sin ninguna penalidad. Me darán una copia de este documento de consentimiento y puedo obtener información sobre los resultados de la investigación después si quiero.

Usted recibirá una copia de este documento para sus récordes:

Nombre del/de la Participante: _________________________   Fecha: _________________________

(firme)

Forma de consentimiento explicada y estudio resumido por ______________________________

________________________________________________

(firma)
Vita

Dan Morgan was born on September 23, 1981 in El Paso, TX. He graduated from Franklin High School in 2000. After serving a two year LDS mission in the country of Brazil, he enrolled at Brigham Young University in January of 2003. While at BYU, he was inducted into the National Society of Collegiate Scholars. He graduated from Brigham Young University in the winter of 2007 with a B.A. in Linguistics along with a minor in Portuguese. After a short period of substitute teaching and working as an office temp worker, he returned to his studies at The University of Texas El Paso in the Graduate School in January of 2009. He worked as a teaching assistant during his time at UTEP. During his first year, he taught two Portuguese classes, and he tutored students of ESL and Arabic. During his second year, he taught two ESL 1610 classes. As he was finishing his thesis during the summer of 2011, he worked as an adjunct instructor at El Paso Community College.

Permanent address: 4201 Siete Leguas
El Paso, TX 79922

This thesis was typed by Dan Morgan