

# An Empirical Study on the Influence of Spanish Experience on the Production of English Stops by Chinese English Majors

Siying Yao

International College of Chinese Language and Culture, Hunan Normal University, Changsha, China  
Email: 1747135014@qq.com

**How to cite this paper:** Yao, S. Y. (2025). An Empirical Study on the Influence of Spanish Experience on the Production of English Stops by Chinese English Majors. *Open Journal of Social Sciences*, 13, 372-388.  
<https://doi.org/10.4236/jss.2025.134022>

**Received:** December 30, 2024

**Accepted:** April 26, 2025

**Published:** April 29, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

This paper, based on the second language acquisition theories Perception Assimilation Model on L2 (PAM-L2) and Speech Language Model (SLM), as well as theories of Language Transfer and Language Attrition, selects 5 students who choose Spanish as a second foreign language among the students of the Foreign Studies College, Hunan Normal University, Class of 2019 as the participants of the experiment. The experimental stimuli are three groups of English words and Spanish words starting with the letters [t] [d] [p] [b] [v] [h]. There are 18 English words and 18 Spanish words in each group, and the words are similar in both languages. Based on the participants' output figures, Praat and xRecorder are used to extract acoustic parameters to analyze their English and Spanish stops, and Voice Onset Time (VOT) is used to demonstrate that there is a negative transfer effect of third language acquisition on second language acquisition.

## Keywords

PAM-L2, SLM, Language Transfer, Language Attrition, VOT

## 1. Introduction

### 1.1. Background of Study

Speech transfer is an important part of the study of trilingual transfer. Spanish learning has gained popularity recently among academics as a second foreign language for English majors. There are, however, comparatively few empirical investigations and the majority of the studies are theoretical. Therefore, based on this background and combined with theories Perception Assimilation Model on L2

(PAM-L2), Speech Language Model (SLM), Language Transfer and Language Attrition, this paper takes a production test to study the influence of Spanish experience on the production of English stops.

### 1.2. Purpose of Study

The main purpose of this paper is to study the influence of Spanish experience on the production of English stops by using Voice Onset Time. And based on the four theories and the influence of the language transfer, the study can identify the problem of similar stops in English and Spanish. It also explores how the trilingual experience affects bilingualism. And then the phonetic pronunciation problem can be continuously improved to be able to distinguish between the two different language stops.

### 1.3. Significance of Study

Phonetic learning is necessary. In the past Spanish teaching, it can be found that students like to use avoidance strategies when they cannot communicate well in Spanish. That being said, vocabulary and grammatical points can be “evaded”, but phonetics are not. Voice is the material shell of language, and it does not have such a huge system compared to vocabulary and grammar. Therefore, the study hopes that in teaching Spanish stops in third language, teachers can effectively analyze and correct the pronunciation problems encountered by students and help students minimize the influence of Spanish learning experience on English stops. At the same time, according to this study, it will not only enable students to fully visualize the pronunciation of speech sounds, as well as the physiological organs and principles of human pronunciation, but also to find out whether trilingual experience has a positive or negative transfer effect on the pronunciation of second language English stops.

## 2. Literature Review

### 2.1. Theoretical Framework

This paper mainly uses PAM-L2, SLM, Language Transfer and Language Attrition for analysis. This paper is based on four theories to analyze the influence of Spanish experience on the production of English stops.

In cross-linguistic speech recognition research, Best (1995: pp. 194-195) proposed the Perceptual Assimilation Model (PAM), which identifies three cases of perceptual assimilation of non-native phonemes by monolingual speakers and six cases of assimilation of two contrasting non-native phonemes and their degree of differentiation. Vowels are more variable than consonants in assimilation patterns (Best, Faber, & Levitt, 1996). So Best demonstrated that the PAM can also predict and explain the perceptual patterns of L2 for monolinguals in 2010. Neither the PAM nor the NLMM include a factor for L2 learning experience and are applicable to the initial stages of learning. However, Guion et al. (2000) found that the

PAM could largely be extended to later stages with appropriate modifications; Best & Tyler (2007) eventually proposed the PAM-L2. Catherine Best proposed PAM-L2 (Perceptual Assimilation Model on L2) in 2007. This model is based on the best previous PAM (Perceptual Assimilation Model) with some improvements and changes. PAM is based on pronunciation phonology, which maps non-native sounds to listeners based on pronunciation similar to the native language category. The subject of this model must be a specific speaker. At the same time, PAM-L2 can directly study the perceptual assimilation of speech in the process of second language acquisition. The focus of this model is whether the SLA (Second Language Acquisition) listeners exhibited perceptual learning of the L2 contrast that was initially indistinguishable. PAM-L2 takes functional monolingualism as a starting point and assumes that learners are actively learning a second language in a predominantly oral environment. PAM-L2 assumes that the perceptual system is common to all of the learner's languages. If certain L1 speech categories are sufficiently useful for discriminating L2 contrasts, then no additional learning is required for these contrasts. If, on the other hand, the learner cannot detect L1 contrasts for a particular pair of L2 phonemes, then perceptual learning is required to be able to detect L2 phoneme contrasts and to build an L2 vocabulary that preserves the phonological distinctions between these phonemes. How successful learners are in detecting new phonological contrasts in L2 depends on how the L2 phonemes were initially assimilated by the L1 phonological system. Many studies have shown that learners' cognition of a second language will change with the progress of learning. But the exact change depends on the similarities and differences between L1 and L2.

Another well-known model that is often used for analysis and production is the SLM (Speech Learning Model) proposed by Flege. It is a rich model containing four postulates and seven hypotheses that predict and explain the final outcome of L2 acquisition. In order to take into account for age-related restrictions on the capacity to generate L2 vowels and consonants in a native-like manner, Flege and his colleagues have created an SLM (speech learning model). At present, there are 3 points worth noting about SLM. Firstly, SLM considers that speech perception occurs at the level of phonological variants, rather than at the level of abstract phonological phonemes. This follows on from the CAH (contrastive analysis hypothesis) understanding of phonological variables, but also focuses more on the contrast at the concrete phonological level. Secondly, the SLM suggests that "listening precedes speaking", i.e. that learners only begin to establish new phonological categories once they perceive the difference between L2 and L1, and that the accuracy of L2 phonological output is limited by the adequacy of the phonological categories established for it (Flege, 1993; Flege, MacKay, & Meador, 1999). The SLM argues that many articulatory errors stem from inaccurate listening, but does not believe that all articulatory errors have an auditory origin. In other words, accurate L2 speech recognition is a necessary but not sufficient condition for correct speech production. Finally, the SLM proposes an "Interactive Hypothesis

(IH)” between L1 and L2 (Baker et al., 2008), which suggests that the extent of the interaction depends on the maturity of the L1 phonology at the time of L2 initiation. The older the age at which L2 is first learned, the stronger the interaction the stronger the interaction, the lower the learner’s ability to perceive the difference between L1 and L2, and the greater the effect on L2 listening development (Flege et al., 1999). Since the SLM is primarily focused on the ultimate attainment of L2 pronunciation, research conducted within its framework focuses on bilinguals who have been speaking their L2 for a significant amount of time, rather than beginners.

The third theory is Language Transfer. Interlingual or cross-linguistic effect are other names for language transfer. Transfer is the effect that differences or commonalities between the target language and any other language that has been acquired (or not yet fully acquired) have on the language learner’s language learning process (Odilin, 1989). The effect that one’s knowledge of one language has on one’s knowledge and use of another language (Jarvis & Pavlenko, 2008). Eris proposes four categories of language transfer: positive (positive transfer), negative (negative transfer), avoidance and over-representation. Language transfer is mainly positive and negative. Positive transfer describes how having the same or comparable forms in the mother tongue as the target language positively influences learning the target language. Negative transfer is when patterns or rules from the mother tongue are applied directly to the target language without taking into account the variations between the mother tongue and the target language. Language transfer can be divided into three main categories based on the dimension of direction: lateral transfer between transitional languages, such as bilingual to trilingual and trilingual to quadrilingual; forward transfer from the mother tongue to the target language; reverse transfer from the target language to the mother tongue; and transfer from the mother tongue to the target language. The systematic study of reverse transfer broadens the scope of migration research, showing that migration is an interlanguage interaction and not just a monopoly of the mother tongue. Current research on lateral transfer focuses on answering the question of the source language of transfer, i.e. whether the target language is more likely to be learned in the native or mediated language.

The fourth is the Language Attrition. Language Attrition is the idea that a bilingual or multilingual speaker’s ability to use the language will gradually diminish over time due to a reduction or cessation of use for some reason. Compared to acquisition theory, the study of language attrition is relatively recent, with the 1982 book *The Loss of Language Skills*, edited by Freed and Lambert, marking the real beginning of the study of language attrition. Language Attrition suggests that language acquisition includes not only a learning and acquisition aspect, but also a loss, an erosion aspect; not only an upward, developmental aspect, but also a rigid, degenerative aspect. In this sense, language attrition is the inverse process of language acquisition and is an important factor governing the ability to learn a foreign language. The study of language attrition in China started late and has

mostly been a review of foreign research findings. In the area of national languages, research has focused on mother tongue attrition and foreign language attrition. Ni et al. (2009) point out that learning a foreign language at the same time as the mother tongue system is not yet well constructed can cause mother tongue attrition in younger children. According to Zhang Jijia, the factors affecting mother tongue erosion include both objective and subjective causes, and ethnic education should pay attention to preventing mother tongue erosion. In terms of foreign language attrition, Ni (2015) measured 116 Chinese university graduates and found that foreign language vocabulary attrition is regressive, showing that foreign vocabulary learned first is attrited later. Tang et al. (2015) found that Chinese non-English major university students experienced an overall abrasion of reading ability after completing formal English instruction, with individuals with high levels of foreign language reading being less abrasive than those with low levels of foreign language reading. Occasional study of a foreign language at a later stage reduced the degree of foreign language attrition.

These theoretical models have different roles to play in the study of this paper. The PAM-L2 explores whether the acquisition of non-mother tongue sounds is assimilated in the context of second language learning experience. The SLM explores how trilingual experience affects second language learning. The Language Transfer theory explores whether trilingualism has a positive or negative transfer effect on second language. And the Language Attrition theory explores whether the ability of trilingual speakers to use a language decrease over time as a result of a reduction or recession in the use of that language. Then the figures measured by VOT are used to prove the above four theories.

## 2.2. Previous Study

The concept of “transfer” is not actually first proposed in the field of second language acquisition, it is actually an important concept of Learning Psychology. Psychologist Ellis once defined “transfer” as a hypothesis that the learning of task A will affect the learning of task B, and said that language transfer is “perhaps the most important concept in educational theory and practice.”

Reeder (1998) studied the acquisition of unvoiced vowels in Spanish by native English speakers and found that the speech acquisition of Spanish language learners improved with the learning stage, but from the perspective of acoustic parameters, it had not reached the same level as the native Spanish language. In the long run, the acquisition of vowels was more complete than other phonetic syllables. This finding supports a view in Flege’s SLM model that sounds different from L1 in L2 are easier to acquire than similar sounds.

Magloire & Green (1999) investigated how voice onset time (VOT) was affected by variations in speaking speed in Spanish and English. Three groups of subjects (English monolinguals, Spanish monolinguals, and early Spanish-English bilinguals) produced sentences containing voiced and unvoiced bilabial pauses at different speaking speeds. The results of this study suggest that short lag pauses pro-

duce minimal variation in speech rate, regardless of other contrastive speech categories in a given language. Therefore, based on this study, it is possible to further explore whether the experience of learning Spanish has an effect on the VOT of the English stops, consistent with native Chinese, English as the first foreign language and Spanish as the second foreign language.

Rose (2012), based on the 2010 research, followed Best's PAM model to explore the reasons, and found that it is difficult for learners to distinguish between the Spanish flash sounds and the d in English because they classify the two in the same category. The results also show that two Language learning experience and vocabulary context will affect the classification of second language consonants in the mother tongue.

Nagle (2019) examined the VOT of Spanish /b/ and /p/ during two semesters of primary language instruction based on both SLM and PAM-L2 theories. Twenty-six native Spanish learners completed two L2 production tasks five times and an English production task once with the aim of determining the frequency of their pre-production of English stops. The results included near-native learners and asymmetrical developers who improved production of /p/ but not /b/. But it only studied one pair of sounds.

From the review of existing domestic and foreign research, it can be concluded:

- 1) An essential aspect of the research of second language acquisition is the study of language transfer. Learning to speak well is an important step in learning a language and by examining whether the Spanish learning experience has an impact on the English stops, there are insights for second language teaching.
- 2) There are many improper substitutions for beginners, and the two voices substitute for each other during use. So, the influence of Spanish stops on English stops can be studied on the basis of previous research by measuring VOT.

### 3. Research Questions and Predictions

#### 3.1. Research Questions

In this paper, the following question is investigated through a production test:

- 1) What effect does the Spanish learning experience have on English stops when native Chinese English majors choose Spanish as a second foreign language?
- 2) Whether the use of VOT values can validate the conclusions of the four theories at the current stage of research?

In this paper, five English majors from Hunan Normal University whose second foreign language is Spanish are selected as subjects for this experiment, given the limitations of the experimental subjects. In the output experiment, a pronunciation test of 32 English words and Spanish words each is set up and recorded using the X-Recorder. After recording, the recording figures are extracted using Praat and the acoustic parameters are analyzed. VOT is used as acoustic parameter to investigate whether there is a positive or negative transfer of Spanish learn-

ing experience on English stops.

### 3.2. Research Predictions

1) When English majors learn Spanish stops, they can distinguish between English and Spanish stops more accurately if they are given enough time to differentiate and practice their pronunciation. However, if they don't remember and practice, after a while the two languages will still be confused.

2) English majors can learn Spanish as a second foreign language more precisely if they learn Spanish as a third language, are able to discern between the two languages, and learn Spanish as a third language.

3) The more time English majors spend studying Spanish stops, the more knowledge and expertise they get, and the more their pronunciation of English stops will be somewhat influenced by that of Spanish stops.

Additionally, English majors are influenced by their mother tongue, Chinese, as well as their second language, English, including things like pronunciation, stress, and sound length. Furthermore, Chinese speakers who major in English also have influences from their original language, including pronunciation, emphasis, and sound length. These elements are not taken into account in this paper. However, prior studies have shown that having Chinese as a first language has a significant impact not just on learning English as a second language but also on learning pronunciation in a third language, where positive or negative transfer might happen.

## 4. Methods

### 4.1. Participants

Due to the limited number of participants selected for this study, its findings do not apply to all English majors who wish to study. In addition, these participants are the first students of Spanish as a second language at our university. Also, only English is used as a second language, while Spanish is used as the target language. Therefore, the specific implementation of the teaching methods is different for other language learners and the pilot study is done.

The five participants in this study are undergraduate English majors from Hunan Normal University, consisting of four females and one male. They have the same teacher for their Spanish classes. Each of them has been learning Spanish for a year and a half. In the final test, they all score above 90. In order to more thoroughly examine the issues raised in this research, five English majors who have not studied Spanish are chosen as the control group for this experiment. While the control group is only required to read the English words, the experimental group is required to read the Spanish words twice, followed by the English words twice. Both groups are given one-to-one independent word pronunciation practice only for the adult English majors, and no sentence pronunciation practice is given. As mentioned above, this study has its limitations and we still have much work to do.

## 4.2. Stimuli

The experiment is a production test. The subject participates in the recording of the sound through the recording device xRecorder. The device records what the recorder reads and then converts it into text. After the text is cut by Praat, acoustic parameters and related figures are extracted to help study the material.

For this experiment, target words with predominantly [t] [d] [p] [b] [v] [h] word headings are selected. English words that conform to the rules of Spanish word formation and begin with a clear stop sound are selected. Mono-, di- and trisyllabic words are taken into account, both true English words and English pseudo-words (e.g., ultra-low frequency words such as Buscopan). The experimental group pronounce the English and Spanish words twice and the control group pronounce the English words.

**Table 1** is three sets of stimulus items, Spanish and English words with initials containing the stops [t] [d] [p] [b] [v] [h]. The Spanish words are shown before and the English words after; the subjects are told to read the Spanish words twice and then the English words.

**Table 1.** Stimulus items.

<b>P (English-Spanish)</b>	<b>B (English-Spanish)</b>
Piloto-Pilot	Billón-Billion
Pulsera-Pulse	Buscar-Buscopan
Parte-Party	Ballet-Ball
<b>T (English-Spanish)</b>	<b>D (English-Spanish)</b>
Tigre-Tiger	Director-Director-
Tumor-Tumor	Ducharse-Duchess
Taza-Tapa	Dato-Date
<b>K (English-Spanish)</b>	<b>G (English-Spanish)</b>
Kilómetro-Kilometer	Girar-Girard
Kurdo-Kurtosis	Guitarra-Guitar
Kaki-Kaki	Garaje-Garage

## 4.3. Procedure

The whole experiment is a production test. The test takes place in a quiet language laboratory at Hunan Normal University. The experiment does not tell subjects whether the word is pronounced as it is in English or Spanish and is pronounced in a word. Before the formal recording is made, the subjects are confirmed to be willing to participate in the experiment and are told that they can leave at any time. Upon entering the room where the recordings are collected, the microphone

is adjusted to the appropriate position and the task of the experiment will be briefly explained. For each word, the subject first says the word in a natural rhythm. If the subject thinks the pronunciation is incorrect during the experiment, the subject can repeat what he thinks is correct without having to start again.

Before the formal test, the researcher informs the subject of the use of the xRecorder and has it imitate an experiment in advance, informing the subjects of the norms of use and the rules of recording to ensure that the subjects' voices and content are normative recordings during the experiment. During the experiment, subjects read the material that appears on the screen. The experiment is divided into a round, the round is conducted with individual pronunciation of words. And the experimental group is required to read both English and Spanish words, while the control group is required to read only English words. All words are pronounced with stop initials, for a total of 36 words.

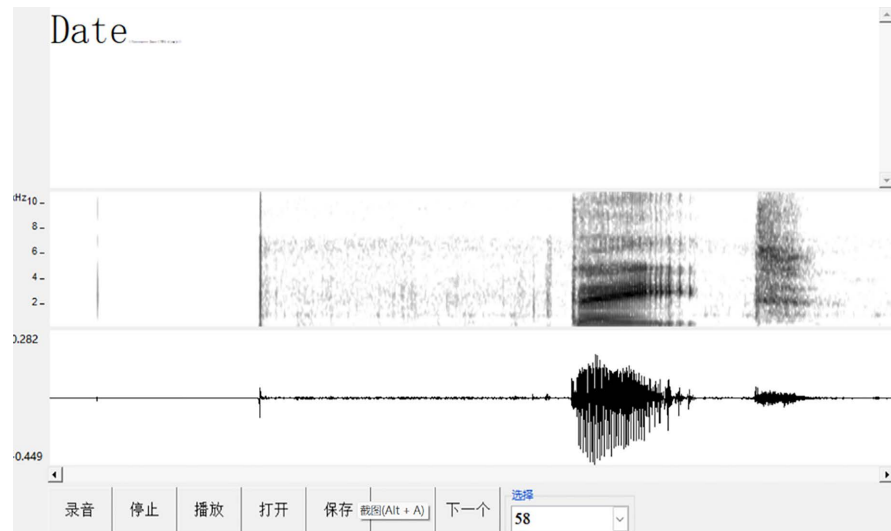
During the experiment, the subjects read aloud the material that appears on the screen. The subjects' figures are saved in the same folder and the experimental figures are imported into Praat for the extraction of acoustic parameters. Voice Onset Time is determined by locating the start of the first vocal cycle. The VOT is calculated by subtracting the time of voice onset from the burst time after all markers after all labeling is finished. VOT is the abbreviation for Voice Onset Time. [Lisker and Abramson \(1964\)](#) study defined VOT as the time interval between the burst that marks release of the stop closure and the onset of quasi-periodicity that reflects vibration. It is regarded as one of the most important acoustic parameters and has been used extensively in measuring word-initial stops. The VOT value of a stop refers to the time between the release and the onset of the vocal fold vibration of the following vowel, which is reflected in the spectrogram as the time between the formation of the punch straight bar and the first regular ripple due to the sudden release of energy. According to Lisker and Abramson, voice onset time serves as a device for "separating the stop categories of a number of languages in which both the number and phonetic characteristics of such categories are said to differ." ([Lisker & Abramson, 1964](#)) When the vocal fold vibration precedes the release, the VOT value is negative, i.e., prevoiced; when the vocal fold vibration occurs almost simultaneously with the release or slightly after it, the VOT value is positive, i.e., short-lag; when the vocal fold vibration occurs after the release, the VOT value is positive, i.e., long-lag.

#### **4.4. Data Analysis**

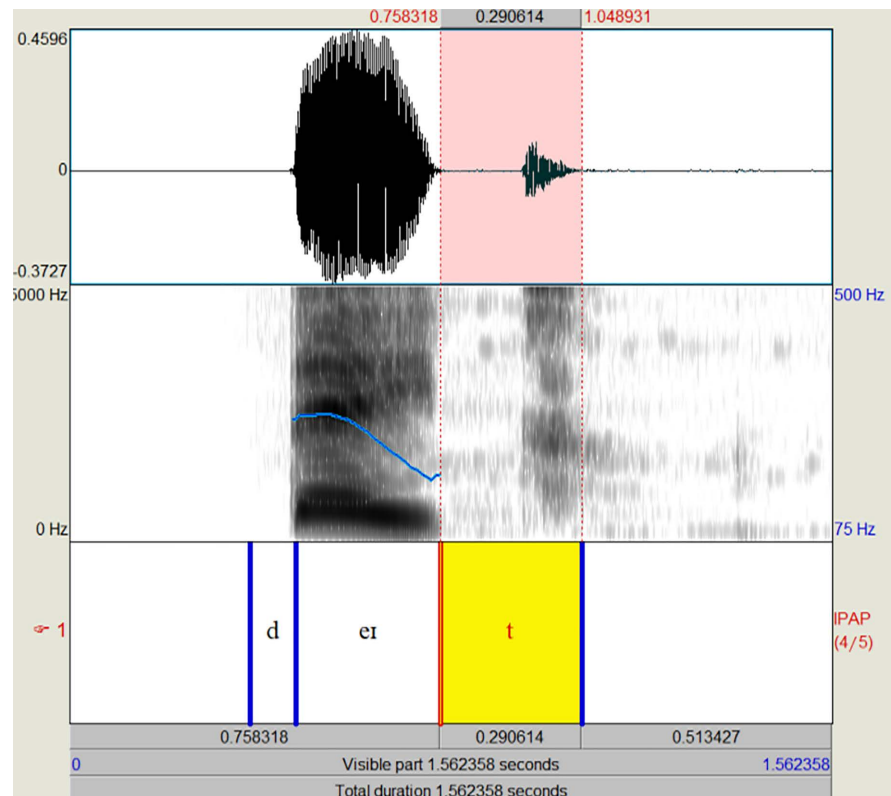
In the preparation of the data for this experiment, the experimental figures are collected and batch labelled using xRecorder and Praat.

There are 36 words in this experimental data, 18 English words and 18 Spanish words. For this experiment, the phonetic symbols in each of the 18 English words and Spanish words are labelled in an Excel sheet using the International Phonetic Alphabet. After marking, the data is copied and pasted into the xRecorder script.

And subjects read the material that appeared on the screen, as **Figure 1**. And the basic acoustic parameters are extracted using Hirst's analyze tier script in order to better measure the VOT of English and Spanish stop sounds. The recordings of the participants are saved originally as .m4a files, which are then converted into .wav files using Adobe Premiere. The .wav files can then be analyzed by Praat, and the VOT can be extracted using its text grid function.



**Figure 1.** Example of a test English word “Date” shown in the xRecorder.



**Figure 2.** Example of an English word “Date” shown in the waveform and spectrogram.

Speech events such as burst and VOT are labeled using Praat. Waveform and spectrograph are visually shown on the screen, as **Figure 2**. And they are displayed to aid labeling. The burst is identified as the peak of an individual spike from a cluster of spikes that make up the transient noise of constriction release, and is recognizable by the first clear deviation from the zero crossing in the waveform. VOT is identified by locating the beginning of the first voicing cycle. After all labeling is finished, VOT is calculated by subtracting the time of voice onset from the burst time. The voice-onset times are then saved in Microsoft Excel for further analysis.

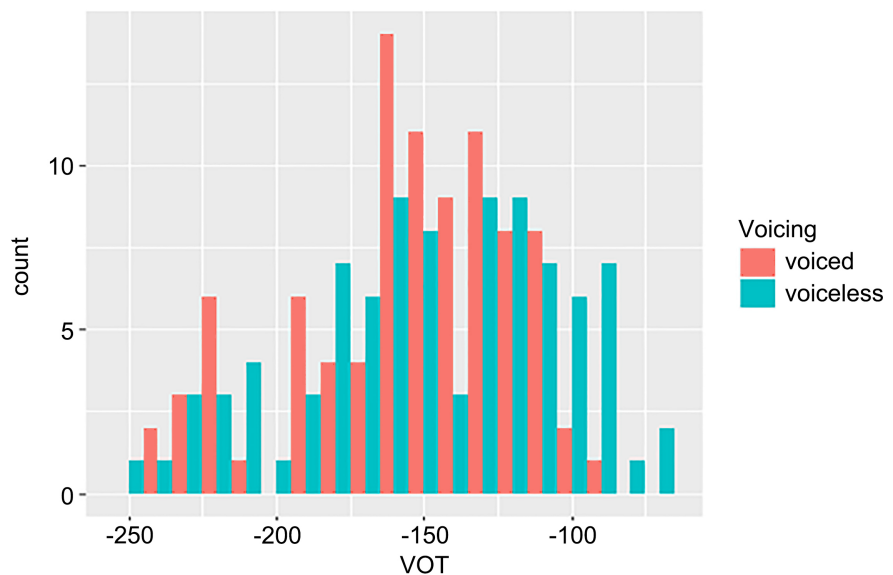
## 5. Results

### 5.1. Data Visualization

This experiment uses R to produce descriptive statistics on the output of Spanish and English stops for 5 subjects. The script for this R run is in **Appendix 1**.

#### 5.1.1. Overall Results for Spanish VOT in Experimental Group

The VOT for Spanish stops are visualized in **Figure 3**. According to the Spanish pronunciation characteristics, the stimulus [p] [t] [k] in this experiment is voiced to [b] [d] [g]-like sounds, i.e., the Spanish voiceless sounds essentially appear as voiced sounds. Based on the characteristics of the VOT, the Spanish stops [p] [t] [k]-[b] [d] [g] are then prevoices and have negative values. The horizontal coordinates in this graph are the VOT values, and the 6 Spanish stops have a VOT range of around  $-250$  ms to  $-100$  ms.

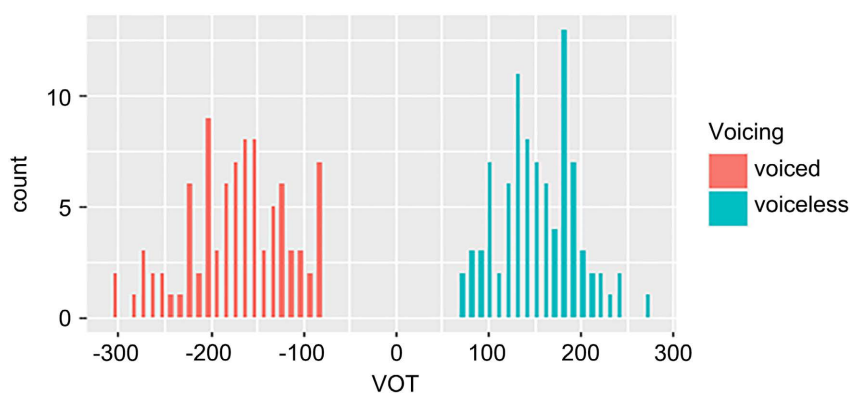


**Figure 3.** VOT of Spanish stops.

#### 5.1.2. Overall Results for English VOT in Experimental Group

**Figure 4** visualizes the VOT values of 5 subjects for the 6 English stops [p] [t] [k]-[b] [d] [g]. The horizontal coordinates in this chart are the VOT values, which

range from  $-300$  ms to  $300$  ms, and the vertical coordinates are the intervals. As shown, there is a clear distinction between voiceless and voiced English stops (voiceless are in blue and voiced are in red). The voiceless [p] [t] [k] are concentrated between  $0$ ms and  $300$  ms and have positive values, while the voiced sounds [b] [d] [g] are concentrated between  $-300$  ms to  $0$  ms and have negative values. Depending on the nature of the VOT, the voiced sounds are generally prevoices, while the voiceless sounds are generally short-lag or long-lag.



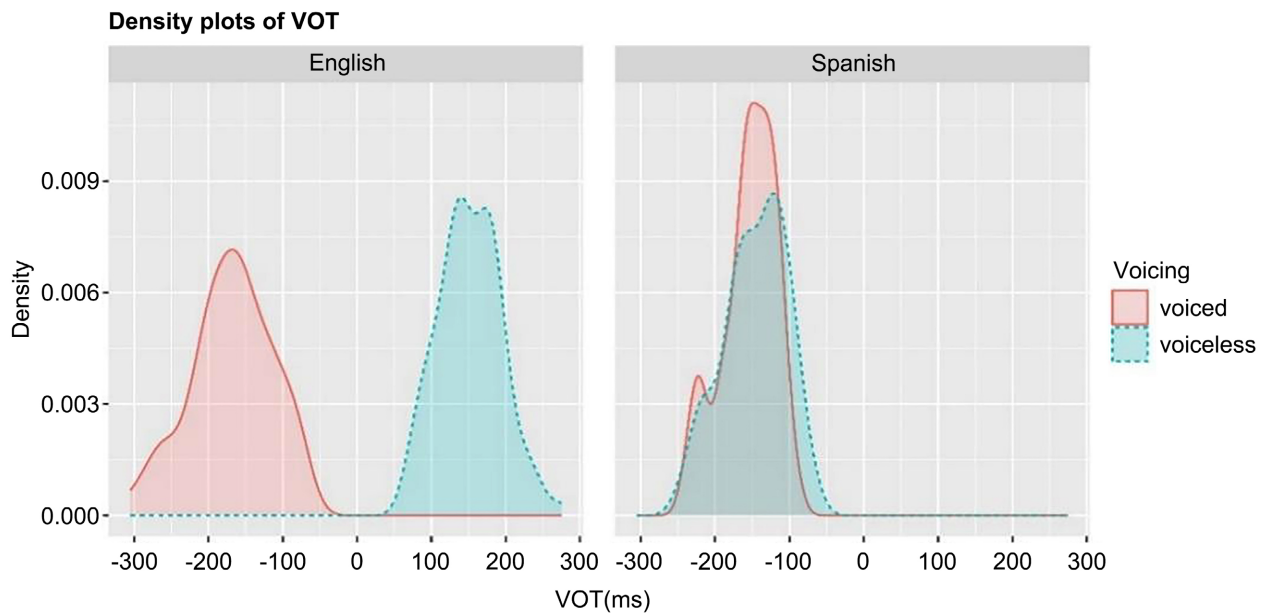
**Figure 4.** VOT of English stops.

In the voiced sounds [b] [d] [g], the VOT is mainly concentrated between  $-250$  ms to  $-150$  ms; in the voiceless sounds [p] [t] [k], the VOT is mainly concentrated between  $100$  ms to  $20$  ms.

### 5.1.3. Comparative Analysis

The aim of this experiment is to test whether the experience of learning Spanish as a second language has a positive or negative transfer effect on the output of English majors in stops. A brief description of the Spanish and English stops coming out of the R run is given above, followed by a comparative analysis approach to study the VOT values of Spanish and English stops [p] [t] [k]-[b] [d] [g] output.

Density plots, also known as Kernel density plots, they're are used to understand the distribution of data. It is considered as an effective way to present the variable distribution over the given time period. The density plot's peak gives the data of concentrated values over the time period. For a better comparative analysis, this production test was visualized using a script of density plots in R for the Spanish and English stops' VOT, as shown in **Figure 5**. The density plot has horizontal coordinates for VOT values and vertical coordinates for density ranging from  $0.000$  to  $0.009$ . In the English VOT the peaks for voiced sounds are concentrated between  $-200$  ms and  $-100$  ms, and for voiceless sounds there are two peaks between  $100$  ms and  $150$  ms and  $150$  ms and  $200$  ms respectively. In the Spanish VOT, since the voiceless sounds are voiced, the Spanish voiced and voiceless sounds are distributed between  $-300$  ms and  $0$  ms, with the peak of the voiced sounds concentrated around  $150$  ms and the voiceless sounds concentrated between  $-150$  ms and  $-100$  ms.

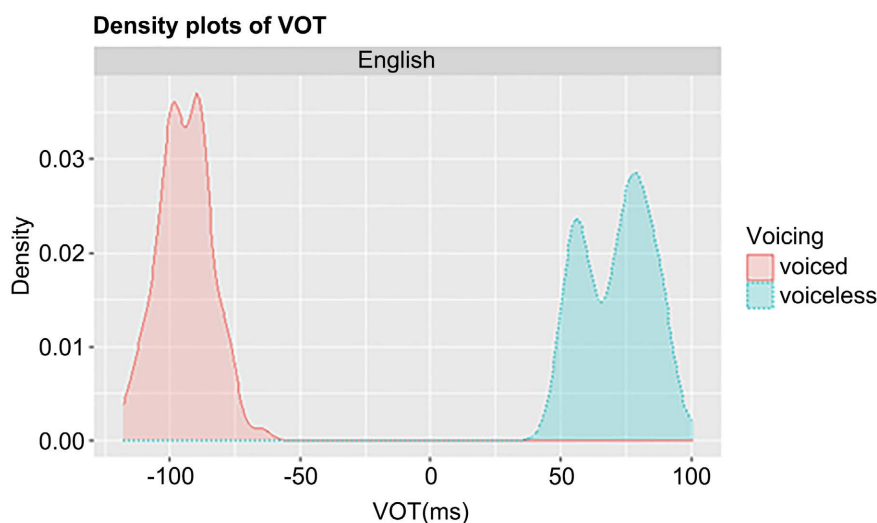


**Figure 5.** Density plots of VOT of Spanish and English stops in the experimental group.

In this paper, the VOT values of English stops produced by 5 English majors who are not studying Spanish are used as a control group, as shown in **Figure 6**. In the English VOT the two peaks for voiced sounds are both concentrated between  $-100$  ms and  $-75$  ms, and for voiceless sounds there are two peaks between  $50$  ms and  $75$  ms and  $75$  ms and  $100$  ms respectively. Compared to the VOT values for English stops derived from the experimental group, the VOT values for the control group are smaller. And in conjunction with previous research, we can learn that in Lisker, Abramson's study as shown in **Figure 7**, it is known that the VOT values for the voiceless sounds [p] [t] [k] are all positive, between  $50$  ms and  $80$  ms for the aspirated voiceless sounds, and between  $0$  ms and  $25$  ms for the unaspirated voiceless sounds. And the VOT for the voiced sounds [b] [d] [g] are negative and range from  $-105$  ms to  $-80$  ms. In Nathan's study as shown in **Figure 7**, the English stops are also characterized by positive VOT values between  $80$  ms and  $120$  ms for the aspirated voiceless sounds [p] [t] [k] and between  $0$  ms and  $25$  ms for the unaspirated voiceless sounds. In contrast, the VOT of the voiced sounds [b] [d] [g] are negative and range from  $-80$  ms to  $-70$  ms.

In this experiment, the two cases of aspirated and unaspirated of voiceless sounds are not taken into account and are only divided into two broad categories of voiced and voiceless sounds. Therefore, in this paper, the VOT values of English stops with Spanish learning experience are compared with the control group and the VOT values of English stops in previous studies. And based on the comparison of VOT figures, we can conclude the followings. Firstly, the VOT of English stops is longer for those with experience of learning Spanish. Secondly, the range of the VOT of English stops and the VOT of Spanish stops with Spanish language experience is approximately the same. Thirdly, according to language transfer theory, the Spanish learning experience has a negative transfer effect on the English stops.

Based on the PAM-L2 and SLM theories, pronouncers with experience of learning Spanish tend to assimilate English and Spanish stops into one pronunciation category when pronouncing English stops.



**Figure 6.** Density plots of VOT of English stops in the control group.

Researchers	/p/	/t/	/k/	/b/		/d/		/g/	
				Voiced sound	unaspirated voiceless sound	Voiced sound	unaspirated voiceless sound	Voiced sound	unaspirated voiceless sound
Lisker, Abramson (1964)	58	70	80	-101	1	-102	5	-88	21
Nathan (1987)	82.5	105	117	-75.5	2.2	-87.7	0	-72	24.7

**Figure 7.** VOT of English stops of Liker, Abramson and Nathan.

## 6. Discussion

These inferences are made based on the graphs displayed in the descriptive statistics:

1) **Figure 3** shows that a number of Spanish voiceless sounds are pronounced as prevoiced sounds, which is evidence of overlearning.

2) **Figure 4** shows that in most cases, the subjects pronounced the English voiced sounds as an early pronunciation, which is not very consistent with native speakers. In reality, the English initial velar stops do not need to be pronounced in advance and are similar to Chinese. Therefore, it is possible that this situation is influenced by the learning of Spanish. In addition, the subjects do not pronounce the English voiceless stops as voiceless without a breath after learning Spanish, but still use a longer breath, suggesting that the Spanish voiceless stops do not affect the subjects' output of the English voiceless stops. It is possible that this is because the subjects, being Chinese, are sensitive to air delivery and there-

fore do not readily abandon the feature of air delivery.

3) **Figure 5** shows that the Spanish stops are pronounced as an early-pronounced voiced sound. This is typical of over-learning. It is likely that the teacher repeatedly emphasized this feature of early pronunciation in class and then the students pronounced it early whenever they encountered a stop regardless of whether it is a voiceless or a voiced one.

4) As can be seen in **Figure 6**, the value of prevoice is actually lower for those who have not studied Spanish. It is around  $-100$  ms; for those who have learnt Spanish, it is around  $-200$  ms.

In the production test, the VOT of English stops from previous studies are analyzed in comparison with the experimentally derived the VOT of English stops. The analysis of the assimilation categorization results shows that all 5 subjects categorize the English voiced and voiceless sounds mainly into the Spanish voiced category, i.e., they perceptually assimilated two phonological categories in the foreign language into the same native phonological category. This is the type of single category assimilation predicted by the PAM-L2 to be the most likely to produce perceptual blending.

Further acoustic analysis shows that the range of VOT values for the English stops [p] [t] [k]-[b] [d] [g] are approximately the same as that of the VOT of Spanish stops for both the voiced and voiceless sounds, suggesting that the English stops pronunciation is influenced by the Spanish stops pronunciation and that this influence is a negative transfer effect.

According to the language attrition, when English majors learn Spanish stops, they can distinguish between English and Spanish stops more accurately if they are given enough time to differentiate and practice their pronunciation. However, if they don't remember and practice, after a while the two languages will still be confused. And if English majors learn Spanish as a third language, can find the difference between the two languages and can distinguish them clearly, then English majors can learn Spanish as a first foreign language more accurately.

The SLM argues that accurate perceptual targets lead to accurate production such that learners who establish a new L2 category for a similar L2 sound will produce it more accurately than learners who perceive it as equivalent to a native category. The five subjects with three semesters of Spanish learning experience would have a longer VOT when pronouncing the English stops compared to the VOT without learning the Spanish stops, but they could still distinguish the voiced and voiceless sounds in English. Thus, learners may have equated the Spanish and English stop categories, but also perceived some phonological differences between them. However, individual results only partially fit with this account. First, at least a few individuals produced natively-like VOT in Spanish, which suggests that they had quickly discerned VOT differences between English and Spanish stops. According to this PAM-L2, learners may have assimilated Spanish and English stops as two distinct phonetic categories linked by a common phonological representation, which arguably allowed them to approximate the phonetic characteristics of

Spanish. On the other hand, participants may have assimilated Spanish and English stops as two instances of a single phonetic category.

## 7. Conclusion

To show that Spanish experience has a negative transfer effect on English stop pronunciation, this paper presents a research analysis through an empirical, i.e., production test. The four theories of language transfer, language attrition, PAM-L2 and SLM are used in this paper to show the research questions necessary for this paper. In the experiment, 36 stimulus items are used, of which 18 are English words and 18 are Spanish words, and the words are similar. The experimental data are collected and analyzed using Prrat and xRecorder, and the processed data are visualized and analyzed using R. In the analysis, the VOT of the English stops from the previous study is compared with the VOT of the English stops from the experiment, and it is found that the VOT of the English stops from the experiment is longer than that of the control group.

Given the difficulties in collecting experimental figures, this experiment is a pilot study. Therefore, more subjects should be engaged in future studies to enhance the validity and persuasiveness of the argument. Besides, although learners in the present study have similar language learning histories, they nevertheless exhibit a high degree of variability in their acquisition of L2 Spanish stops. It seems probable that multiple factors, including phonetic (dis)similarity, aerodynamics, and even individual differences in articulatory flexibility and willingness to communicate, affect both the rate and shape of L2 phonetic learning. So, these factors need to be taken into account in future research.

## Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

## References

- Baker, W., Trofimovich, P., Flege, J. E., Mack, M., & Halter, R. (2008). Child—Adult Differences in Second-Language Phonological Learning: The Role of Cross-Language Similarity. *Language and Speech, 51*, 317-342. <https://doi.org/10.1177/0023830908099068>
- Best, C. T. (1995). A Direct Realist View of Cross-Language Speech Perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 171-204). York Press.
- Best, C. T., & Tyler, M. D. (2007). Nonnative and Second-Language Speech Perception. In O.-S. Bohn, & M. J. Munro (Eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege* (pp. 13-34). John Benjamins Publishing Company. <https://doi.org/10.1075/llt.17.07bes>
- Best, C. T., Faber, A., & Levitt, A. (1996). Assimilation of Non-Native Vowel Contrasts to the American English Vowel System. *The Journal of the Acoustical Society of America, 99*, 2602-2603. <https://doi.org/10.1121/1.415316>
- Flege, J. E. (1993). Production and Perception of a Novel, Second-Language Phonetic Contrast. *The Journal of the Acoustical Society of America, 93*, 1589-1608. <https://doi.org/10.1121/1.406818>

- Fllege, J. E., MacKay, I. R. A., & Meador, D. (1999). Native Italian Speakers' Perception and Production of English Vowels. *The Journal of the Acoustical Society of America*, *106*, 2973-2987. <https://doi.org/10.1121/1.428116>
- Guion, S. G., Fllege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An Investigation of Current Models of Second Language Speech Perception: The Case of Japanese Adults' Perception of English Consonants. *The Journal of the Acoustical Society of America*, *107*, 2711-2724. <https://doi.org/10.1121/1.428657>
- Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistics Influence in Language and Cognition*. Routledge.
- Lisker, L., & Abramson, A. S. (1964). A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *WORD*, *20*, 384-422. <https://doi.org/10.1080/00437956.1964.11659830>
- Magloire, J., & 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*, 158-185. <https://doi.org/10.1159/000028449>
- Nagle, C. L. (2019). A Longitudinal Study of Voice Onset Time Development in L2 Spanish Stops. *Applied Linguistics*, *40*, 86-107. <https://doi.org/10.1093/applin/amx011>
- Nathan, G. S. (1987). On Second-Language Acquisition of Voiced Stops. *Journal of Phonetics*, *15*, 313-322. [https://doi.org/10.1016/s0095-4470\(19\)30583-2](https://doi.org/10.1016/s0095-4470(19)30583-2)
- Ni, C. B. (2015). Analysis of the Regression of Foreign Lexical Attrition. *Foreign Language Learning Theory and Practice*, *No. 4*, 21-28, 93.
- Ni, C.B., & Liu, Z. (2009). First Language Attrition: The State of the Art. *Contemporary Linguistics*, *No. 3*, 224-232, 285-286.
- Odling, T. (1989). *Language Transfer: Crosslinguistic Influence in Language Learning*. Cambridge University Press. <https://doi.org/10.1017/cbo9781139524537>
- Reeder, J. T. (1998). English Speakers' Acquisition of Voiceless Stops and Trills in 12 Spanish. *Texas Papers in Foreign Language Education*, *3*, 101-118.
- Rose, M. (2012). Cross-Language Identification of Spanish Consonants in English. *Foreign Language Annals*, *45*, 415-429. <https://doi.org/10.1111/j.1944-9720.2012.01197.x>
- Tang, C. J., Ma, Y. L., Tang, T., & Deng, X. Y. (2015). A Survey on Language Attrition of Non-English Majors of China Based on Sample Survey of Their Reading Competence. *Journal of Xi'an International Studies University*, *No. 3*, 69-73.