

©2023

Kyle Parrish

ALL RIGHTS RESERVED

THE PHONOLOGY OF THE ABSOLUTE INITIAL STATE OF L3 ACQUISTION

By

KYLE PARRISH

A Dissertation submitted to the

School of Graduate Studies

Rutgers, the State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in The Department of Spanish and Portuguese, Second Language

Acquisition and Bilingualism

Written under the direction of

Joseph Casillas

And approved by

---

---

---

---

New Brunswick, New Jersey

May 2023

## ABSTRACT OF THE DISSERTATION

The phonology of the absolute initial state of L3 acquisition

by

KYLE PARRISH

Dissertation Director:

Joseph Casillas

This is where the body of your abstract goes. The abstract should summarize your work. The abstract for a dissertation or document may be longer than one page; word count is more important than page length in this section.

## ACKNOWLEDGEMENTS

Couldn't find sections/acknowledgements.tex to include as your acknowledgements.

## DEDICATION

*I dedicate this dissertation to my loving wife, Marinna, and to my 3 fur children,  
Olive, Kevin, and Phyllis. I could not have done it without you.*

## TABLE OF CONTENTS

<b>ABSTRACT . . . . .</b>	<b>ii</b>
<b>ACKNOWLEDGEMENTS . . . . .</b>	<b>iii</b>
<b>DEDICATION . . . . .</b>	<b>iv</b>
<b>TABLE OF CONTENTS . . . . .</b>	<b>v</b>
<b>LIST OF TABLES . . . . .</b>	<b>vi</b>
<b>LIST OF FIGURES . . . . .</b>	<b>vii</b>
 <b>Chapter 1: Literature Review . . . . .</b>	 <b>2</b>
1.1 Introduction . . . . .	2
1.2 Models of L2 phonological acquisition . . . . .	3
1.2.1 The Speech Learning Model . . . . .	3
1.2.2 The Perceptual Assimilation Model . . . . .	3
1.2.3 The L2LP . . . . .	5
1.3 Models of third language acquisition . . . . .	5
1.3.1 The Cumulative Enhancement Model . . . . .	5
1.3.2 The L2 Status Factor Model . . . . .	5
1.3.3 The Typological Primacy Model . . . . .	6
1.3.4 The Linguistic Proximity Model and the Scalpel Model . . . . .	7
1.4 Previous literature in L3 phonology . . . . .	8
1.4.1 L3 Production Studies . . . . .	8
1.4.2 L3 Perception studies . . . . .	10
1.5 Methods and anylsis in previous work . . . . .	11
1.6 Cross-lingusitic language features . . . . .	13
1.6.1 Vocalic systems . . . . .	14
1.6.2 Use of Voice-Onset Time . . . . .	14
 <b>Chapter 2: Production at first exposure to an L3 . . . . .</b>	 <b>15</b>
2.1 Overview . . . . .	15
2.1.1 Research Questions . . . . .	15
2.2 Methods . . . . .	15
2.2.1 Participants . . . . .	15
2.2.2 Tasks . . . . .	15
2.2.3 Analysis plan . . . . .	17

2.3	Simulated Results . . . . .	17
<b>Chapter 3:</b>	<b>Vowel Production at first exposure to an L3 . . . . .</b>	<b>18</b>
3.1	Research Questions . . . . .	18
3.2	Methods . . . . .	18
3.2.1	Tasks . . . . .	18
3.2.2	Participants . . . . .	18
3.3	Simulated Results . . . . .	18
<b>Chapter 4:</b>	<b>Perception of L3 vowels . . . . .</b>	<b>19</b>
4.1	Research Questions . . . . .	19
4.2	Methods . . . . .	19
4.2.1	Participants . . . . .	19
4.2.2	Tasks . . . . .	19
4.3	Simulated Results . . . . .	20
<b>REFERENCES</b>	<b>. . . . .</b>	<b>21</b>

## LIST OF TABLES



## LIST OF FIGURES

```
library(ggplot2)
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --

## v tibble  3.1.1      v dplyr   1.0.5
## v tidyr   1.1.3      v stringr 1.4.0
## v readr   1.4.0      v forcats 0.5.1
## v purrr   0.3.4

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

## Chapter 1: Literature Review

### 1.1 Introduction

The difficulty of learning a new language adulthood has a well documented history. This is especially true in the case of phonological acquisition. For example, many studies over the course of the past five decades have demonstrated that adult second language (L2) learners often produce and perceive the sounds of the target language in a non-native manner (cite some flege stuff).

This difficulty arises, in part, because bilinguals often navigate complex communicative situations in which they produce and perceive speech from both of their languages in real time.

Recent studies suggest that bilinguals learn to modify their perceptual categorization routines in accordance with the phonetic realizations of the language they believe they are hearing (cite recent language set studies). Much less is known about the acquisition of a third language (L3A), particularly with regard to L3 production, perception, and phonological learning.

The present investigation explores the perceptual categorization routines of adult bilinguals during the initial stages of L3 acquisition.

Gap in the literature: Much of L3 phonological acquisition is not well known and is still under studied - lack of available resources and participants have lead to low sample sizes

The present dissertation aims to address the issue of low sample sizes and varied statistical analyses by, firstly, recruiting bilinguals who do not yet know a third language, and, secondly, by using more fitting statistical analyses made more appropriate by a larger sample size. It is likely that the relatively homogenous populations of bilinguals are in greater supply than trilingual populations, particularly when suggested methodological practices are to be used, such as the use of mirror-image groups. With

the increase in sample size, more precise observations and conclusions may be drawn about the very starting point of L3 phonological acquisition which are far less likely to be explained by statistical limitations such as sampling error.

## **1.2 Models of L2 phonological acquisition**

Many accounts for L2 phonological acquisition exist. In general, these models all have in common that language-specific L1 categories drive L2 speech development. In the section that follows, a brief overview of each model, along with its evidence, will be covered.

### **1.2.1 The Speech Learning Model**

The original SLM posits..

SLM studies...

The revised SLM...

L3 SLM

### **1.2.2 The Perceptual Assimilation Model**

Relative to the SLM, the Perceptual Assimilation Model (SLM) focuses on the perception of sound contrasts by L2 learners (CITATIONS). The PAM involves various scenarios in which the cross-linguistic inventories of specific languages predict how easily L2 sounds will be to acquire. For instance, if a L2 learner's native language contains a contrast that also exist in the L2, then this contrast will be easily perceived in the L2 (two-category assimilation). In the event that two native language categories correspond the a single L2 category (single category assimilation), discrimination is predicted to be intermediate. In the event that a native category must be split into two categories (X assimilation) discrimination is predicted to be poorer.

Evidence for the predictions of the PAM has been found in studies which involve the

presentation of sounds to naive learners of an L2 (first exposure), or to more experienced L2 learners. Experiments typically present the subjects with native-language vowel categories in written form and auditory stimuli of L2 vowel sounds and participants are tasked with choosing the closest matching native-language vowel category given the options, and to rate the goodness of fit of this decision.

This methodological paradigm has found evidence that **two category assimilation** is difficult for L2 learners. For example,

Evidence for PAM...

**Escudero and Chládková (2010)** found that Spanish L1 speakers assimilated SSBE /ae/ and /a/ to Spanish /a/ (that discrimination of this contrast was difficult).

**Escudero et al. (2014)** Salento Italian L1 speakers assimilated SSBE /ae/, /a/ and /<sup>h</sup>/ to Salento Italian /a/.

**Escudero & Vasiliev (2011)** Spanish speakers assimilated Canadian English /ae/ and /e/ to Spanish /a/

**Escudero & Williams (2011)** Spanish listeners categorized Dutch /a/ and /a:/ in terms of their native /a/

The PAM in L3...

The present dissertation tests whether assimilation acquired during second language acquisition apply to third language sound perception. In other words, it tests whether L3 perception mirrors L2 perception in similar ways that has been observed in L3 production. Specifically, it tests whether L3 speakers categorize L3 sounds similar to L2 sounds, and whether phonetic discrimination of sounds that would be phonemic in the L1 is accessible in the perception of L3 words at first exposure, or whether there is an initial blocking or the L1 or L2 bias effect.

### 1.2.3 The L2LP

The L2LP is similar to the PAM and predicts...

Evidence for the L2LP comes from...

The L2LP in L3...

## 1.3 Models of third language acquisition

Research in third language acquisition has attempted to model the interplay between L1 and L2 language systems in the acquisition of a third. Among questions asked by third language models is whether the L1 or L2, or a combination of both languages, serves as the basis in L3 acquisition. This question is complicated in the context of multilingualism due to the widespread diversity in bilingual populations that include wide variation in ultimate attainment in adult L2 learners, and, in the case of phonological acquisition, wide variation in the production patterns of L2 segments. The predictions and theoretical underpinnings of models of L3 acquisition are covered in the following section, as well as the empirical evidence for each model.

### 1.3.1 The Cumulative Enhancement Model

The Cumulative Enhancement Model (CEM) suggests that each language known by a bilingual may be used to enhance the process of L3 acquisition. In this view, both features of the L1 and features of the L2 are in principle available to the L3 learner from the beginning of the acquisition process.

Evidence for this model was found...

An issue with this model is non-facilitation, as argued by Rothman.

### 1.3.2 The L2 Status Factor Model

In contrast to the CEM, the L2 Status Factor Model (L2SF) predicts that the L2 will influence the L3 by default, due to the proposed cognitive similarity between

the L2 and the L3. These proposed cognitive similarities stem from the Declarative-Procedural model, which posits that late learned languages are largely subserved by the declarative memory system, whereas early learned languages are subserved by procedural memory.

Evidence for the L2 Status factor has been reported...

Counterevidence to this model has also been found in empirical studies that would serve as the basis for subsequent models.

### 1.3.3 The Typological Primacy Model

The Typological Primacy Model, like the L2SF, predicts that a single language will influence the L3, but differs in that this language may be either the L1 or the L2. The choice of language is thought to be driven by individually perceived psycho-typology. In this view, the choice of language is also not thought to be conscious, but rather driven by input in L3 learning.

There has been some support in the literature for the TPM, but this evidence often simultaneously supports other models. For instance, in a recent systematic review of 92 studies in L3 acquisition, @puig-mayenco\_systematic\_2020 concluded that, in support of the TPM, that either the L1 or L2 influenced L3 performance in 59 out of 92 studies. On the other hand, 29 of the studies out of 92 found that the L2 influenced the L3. However, these findings were not mutually exclusive; the authors coded 25 total studies as being explained both by L2 status and typology, meaning that the results of these studies reported that the L2 transferred to the L3, but could not rule out the possibility that psycho-typological transfer could also explain the results, since the studies did not use mirror image groups (i.e. L3 groups with the same languages, but the opposite order of acquisition).

Counter-evidence and issues with this model include lack of cross-linguistic support, and lack of specificity when it comes to the term “initial stages.” For example, evidence for the TPM has been found primarily in studies which contain Romance languages

and a non-Romance language. To the author’s knowledge, no study has yet shown support for the TPM’s predictions outside of language combinations which contain two romance languages and a non-romance language which also employ the criteria suggested by Rothman (CITE) of mirror-image groups. The studies coded in the systematic review by Puig-Mayenco et al. as providing evidence for L2 Status and Typological proximity cannot be said to conclusively provide evidence for the TPM for the same reasons that they do not provide conclusive evidence for the L2SF, since L2 status and typological primacy could explain the results. In short, the TPM does not currently have support from a study which uses mirror-image groups and does not involve 2 romance languages.

A second potential shortcoming of the TPM is the vagueness associated with the term “initial stages” of L3 acquisition. Rothman (2015) describes this time as some 20-25 hours of instruction, before which time access to either language is possible. This vague criterion makes it quite difficult to derive the predictions of the TPM in the case of learners who first encounter an L3.

### 1.3.4 The Linguistic Proximity Model and the Scalpel Model

These models propose that both language systems may influence the L3, but also that they may not.

Unlike the TPM and the L2SF, the Linguistic Proximity Model

*(westergaard\_crosslinguistic\_2017?) suggests that transfer occurs on a property-by-property basis, rather than holistically. The LPM suggests that there is a “full transfer possibility,” meaning that any individual structure may transfer at any time, but that it also may not. Scholars have argued that this vague prediction creates a problem in modeling L3 transfer acquisition (bardel\_l1\_2020?; wrembel\_multilingual\_2020?), since it is unclear when transfer of a particular structure occurs and when it does not. Likewise researchers have argued that, unlike the TPM and the L2SF, the LPM is not easily falsifiable (bardel\_l1\_2020?). Given the lack*



*of predictive power of the LPM, it may only receive post-hoc support.*

Evidence for these models.. Westergaard et al. (2017), Mitronova & Westergaard (2018)

Issues come from lack of predictive power of the models.

## 1.4 Previous literature in L3 phonology

The models discussed to this point, with the exception of Wrembel's model, have focused largely on morpho-syntax. The findings in L3 phonology have largely yielded mixed results and do not quite have a comprehensive model. Despite the lack of a model, some patterns emerge from the body of research. The following sections cover the empirical studies that have been done in L3 production studies across L3 proficiency levels, and the fewer studies done in L3 perception.

### 1.4.1 L3 Production Studies

The findings in empirical studies of L3 phonological cross-linguistic influence have varied. One of the first studies to examine progressive influence of the L1 or L2 on L3 production was the seminal case study of **Williams and Hammarberg (1993)**. This study elicited the production of an adult L1 British English, L2 German, and L3 Swedish speaker in the L2 and L3 upon her arrival in Sweden. The speech samples were rated for native-likeness by native speakers of German and Swedish respectively, with low ratings (i.e. non-nativeness) being elaborated upon. In the event of non-native speech, the raters guessed where the speaker in the recording might be from. The informant had near native productions in her L2 German, while her L3 Swedish was rated as being non-native like and to be German-accented. The experiment was repeated after 6 months in Sweden, however, and the Swedish raters then judged the informant's Swedish to be British English accented.

This study constituted evidence of an L2 status effect in the initial stages of L3 phonological acquisition, in which the second learned language influence L3 production.

This notion has been called the ‘foreign language effect’ (**meisel\_transfer\_1983?**), which refers to the idea that speakers who learn a second non-native language are biased to sound unlike a native speaker of their native language.

The default L2 status effect has received some empirical support in the literature. In a study of global accent production, heavier L2 influence in L3 productions was found by L1 Polish, L2 German and L3 English speakers based on ratings of EFL instructors. Specifically, 53% of L3 productions were identified as having come from a German native speaker (**wrembel\_l2-accented\_2010?**). S vowel production (**kamiyama\_acquisition\_2007?**) vowel reduction and speech rhythm (**gut\_cross-linguistic\_2010?**).

Similar patterns of L2 influence have been found in VOT productions. (**llama\_influence\_2010?**) examined L3 Spanish VOT production by French-English mirror-image bilingual groups and found that both groups had L2-like productions of the L3.

Williams & Hammarberg 1998

Hammarberg and Hammarberg 1993, 2005, Cenoz 2001 Hammarberg 2001 Bannert 2005, Jessner 2006, Fernandes-Boëchat and Siebeneicher Brito 2008,

Other findings in L3 production, however, have yielded mixed results. Several studies have found that acoustic properties of the participants’ productions fall between L1 and L3 values, suggesting that both the L1 and the L2 have some influence on L3 productions, rather than solely one language. For instance, (**wrembel\_vot\_2014?**) measured VOT and aspiration in all languages of participants with two different language combinations: (1) L1 Polish, L2 English, and L3 French; (2) L1 Polish, L2 English, and L3 German. The results showed that each language had a specific stop-value, and that the L3 VOT productions were intermediate, falling between the L1 and L2 values. Similarly, (**wrembel\_cross-linguistic\_2011?**) examined thirty-two learners of L3 French with L1 Polish and L2 English who were recorded reading lists of

words in carrier phrases. As in previous studies (**wrembel\_vot\_2014?**), combined transfer from the L1 and the L2 in VOT productions was found.

Findings of combined L1 and L2 influence in VOT productions were also reported by (**wunder\_phonological\_2010?**) in L3 Spanish speakers, and by (**blank\_transferencia\_2009?**) in L3 English speakers who spoke L1 Brazilian Portuguese and L2 French. Other studies have found an L1 influence on production despite L3 proficiency (**wrembel\_foreign\_2012?**), or in advanced L3 learners (**llama\_revisiting\_2018?**).\*

Parrish (2021) examined Mexican Spanish-English bilinguals who produced French words in isolation at first exposure to the language. The results found that the relative VOT of the L3 fell between their own L1 and L2 values, in line with previous research (citations). However, a subsequent analysis of the data revealed that wide individual variation existed in the data, in which some participants produced L3 French as L1 Spanish like, and other produced intermediate, L2-like values.

#### 1.4.2 L3 Perception studies

Balas (2019), Liu

\*Perception studies in L3 acquisition have been much more scarce than those on production, and primarily in young participants. One of few L3 perception studies in adults is (**liu\_effects\_2019?**), which examined the perceptual boundary of a VOT continuum in trilinguals. The participants were L3 Spanish speakers who spoke L2 English and L1 Chinese. Though the authors focused their analysis on regressive transfer and comparisons to all speakers, the reported boundaries in the study (n = 10, 28ms for Chinese, 24.6ms for English and 23ms for Spanish) suggest that the participants were using their English boundaries for L3 Spanish categorization.

Other studies in L3 perception, which have examined young trilinguals, have found a wide range of results. For instance, (**wrembel\_cross-linguistic\_2020?**) found that cross-linguistic influence is structure dependent and varies among individuals, while

(balas\_perception\_2019?) found CLI to be modulated by a complex combination of factors such as markedness of the segment under examination, proficiency in the L2 and the L3, and L1 typology. Finally, another study on L3 perception in young learners (Polish L1, English L2, German, L3) revealed that the L2 predominantly influences L3 perception in both rhotic sounds and devoicing of word final stops (wrembel\_cross-linguistic\_2020?). These variable findings of the source of influence in L3 perception and production, including single language influence, combined languages influence, or structure-dependent influence, point to a lack of homogeneity in multilinguals.\*

## 1.5 Methods and anylsis in previous work

The widely varied findings in previous work do not have a clear cause, but may be related to issues related to sampling issues combined with methodological choices. In this subsection, an overview of sample sizes in previous work will be given, followed by potential issues related to small samples. Subsequently, common methods in the body of research used to analyze L3 data will be discussed, followed by their potential shortcomings.

Overall, the sample sizes in the body of research across domains have been quite small. An analysis of the sample used in the empirical studies in the systematic review of Puig-Mayenco et al. (2019) revealed that the average sample size per group was **X**, (sd = **Y**). In the empirical studies in L3 phonology cited in the body of this dissertation, the average sample size per group was **X** (sd = **Y**). For a full list of studies included in L3 group calculation, see **list**, and for a full list of studies in the Puig-Mayenco et al. systematic review see Puig-Mayenco et al. (2019). **figure** shows the distribution of sample sizes in L3 research by domain. As the figure suggests, the majority of L3 research has been constrained by relatively small sample sizes.

Unfortunately, the tradition in L3 research has involved model building with small samples. The use of small samples are associated with higher sampling error, and, as a result, a higher risk of type 1 error. In other words, a single study with a small

sample cannot rule out the possibility that their results are due to sampling error, or a non-representative pool of participants from an assumed population distribution. As **brysbeart** argues, low samples lead to low statistical power, and in turn provide a metaphorical blurred picture of our desired outcome.

In addition to issues associated with low sample size, L3 research to date has used statistical methods which provide dubious evidence for their claims. Among these issues is the use of an Analysis of Variance (ANOVA) as a basis for determining whether groups or individuals perform experiment tasks in a practically equivalent manner rather than a statistical Test of Equivalence (**Lakens**). At the heart of this issue is a criticism which may apply to frequentist methods of statistical analysis in general; testing against the null hypothesis. If the null hypothesis is rejected, evidence is provided that the difference between or within groups is non-zero. On the other hand, if a non-zero difference is not found, there is not evidence for practical equivalence. Such an assumption has been made in the L3 literature, and in particular in L3 model building.

For example, in his seminal article introducing the Typological Primacy Model, Rothman (2011) concluded that mirror-image groups of Spanish-English bilinguals did not perform differently on the L3 Brazilian Portuguese acceptability judgment task and took this as evidence for similar performance between the groups, and evidence of typological similarity effects in L3 judgments. There are two possible issues with this conclusion. First, sampling error, it can't be said that the low samples did not draw primarily from participants who were influenced by their Spanish, second a wide confidence interval from the low sample size would make providing evidence for statistical equivalence in a test of equivalence likely impossible. A power analysis of the descriptive data in Rothman (2011) reveals that the statistical power in a test of equivalence with the report sample sizes (12 and 15) is zero. This suggests that, given the means and standard deviations reported in this study and equivalence bounds of  $-.4$  and  $.4$  standard deviations from the mean, 100 simulations (random sampling from distributions with those means and standard deviations) of the repeated

experiment yield a statistically equivalent result **0** percent of the time. On the other hand, non-zero differences are found **X** percent of the time. These post-hoc analyses of the data from Rothman (2011) suggest that sufficient evidence was not provided for practical equivalence between groups. **brysbeart** argues that, for between group comparisons, samples of at least 80 per group are necessary.

It seems that large sample sizes are quite difficult to find in L3 research. Plonsky (2015) suggests that combating issues associated with low sample size is possible in ways other than simply increasing sample size. For instance, he suggests that the use of descriptive statistics, including effect sizes and confidence intervals, would be an improvement in L2 research in general. This advice is in line with the idea that frequentist analysis, and linguistic research, has relied on p-values to determine the presence of a so-called statistical difference. Plonsky, along with others, have argued that the use of p-value alone to make real-world inferences is problematic due to issues associated with sampling error and the presence of the magnitude of an effect.

In order to address these potential issues, the present dissertation recruited bilinguals, rather than trilinguals, at first exposure to a third language in order to pull from a likely higher and more homogeneous population of participants. This higher sample, coupled with the use of frequentist tests of equivalence and Bayesian inference allow for both a categorical and gradient interpretation of the data. In doing so, less reliance is put on a narrative interpretation of the results, and the results lend themselves to a more objective outcome.

## 1.6 Cross-linguistic language features

In order to gain insights into the relative influence of a first and second language the present dissertation examines (relative) voice-onset time (VOT) in L3 production, the formant values of vowels in perception and production. In the following sections, the vocalic systems of Spanish, English and French, as well as their respective uses of voice-onset time, are overviewed.

### 1.6.1 Vocalic systems

The Spanish vowel space is the smallest of the present study and consists of 5 distinct vowel monophthongs (**LIST THEM**) categorizations. English and French have a larger relative vowel space.

### 1.6.2 Use of Voice-Onset Time

The description of cross-linguistic uses of VOT.

*Voicing in these segments can be measured acoustically using voice-onset time (VOT). VOT refers to the time interval between the release of a stop consonant and the onset of vocal fold vibration (**lisker\_cross-language\_1964?**). For example, in Spanish, the difference between [p] and [b] is manifested as a difference in VOT. Where [b] is voiced (a negative VOT), [p] is voiceless (a positive VOT). Languages which contain this distinction, in which the realizations of /p/ and /b/ are phonetically voiceless and phonetically voiced, respectively, are referred to as “true-voicing languages” (**lisker\_cross-language\_1964?**). Relevant to the present study, Spanish, French, and Hungarian fall under this category. English, on the other hand, contrasts stop consonants with only positive VOT via a long versus short lag distinction. That is to say, both /p/ and /b/ are phonetically voiceless, but /p/ is realized as a long lag stop and is aspirated, and /b/ is a short-lag stop that is not aspirated (**lisker\_cross-language\_1964?**).*

## **Chapter 2: Production at first exposure to an L3**

### **2.1 Overview**

The present study investigates the production of words in a language by bilinguals that they do not yet know. In so doing, the present dissertation aims to investigate the very initial stage of acquisition by providing an overview of the initial relative influence of both L1 and L2. In this study, sequential, late, Spanish-English bilinguals of both order of acquisition will be given an aspirating L3 (German) or a true-voicing L3 (French) to shadow.

#### **2.1.1 Research Questions**

RQ1: Will Spanish-English bilinguals produce L3 words with Spanish or English like VOT?

RQ2: Will Spanish-English bilinguals produce vowels with F1 and F2 values more like English or Spanish?

### **2.2 Methods**

#### **2.2.1 Participants**

8 total groups of participants will take part in the experiment. There will be 4 bilingual groups with opposite orders of acquisition, Spanish L1-English L2, and English L1-Spanish L2, with each of these groups hearing either L3 German or L3 French, and 4 monolingual English and monolingual Spanish groups, assigned to either French or German. The bilingual groups will contain 75 participants per group, where the monolingual groups will contain 50 participants per group, for a total of 500 participants.

#### **2.2.2 Tasks**

Organization paragraph: An elicited production task & A shadowing task in the L3.



### **2.2.2.1 Bilingual Language Profile**

To measure language use and language dominance, the Bilingual Language Profile (BLP) will be used.

### **2.2.2.2 LexTALE**

The LexTALE will be used to measure English and Spanish proficiency. The LexTALE is a lexical decision task, in which participants see either words or pseudowords on a screen one at a time. Participants are supposed to then decide whether the word presented is a real word in the language or a pseudoword. The task is intended to be a measure of vocabulary size and thus a proxy of proficiency. ADD SENTENCE JUSTIFYING THE LINK BETWEEN VOCAB AND PROF. Either the English and Spanish versions of the task will be used in the study, depending upon the L2 of participant.

### **2.2.2.3 Elicited Production Tasks**

In the elicited production task, the participants read words which appeared on the screen one at a time and submitted their recording of the read words by clicking on a record button to begin the recording, a stop button to end the recording, and an upload button upon completion of the recording. Participants were able to listen to their recording after completing a trial and were able to re-record in the event of an error.

### **2.2.2.4 Shadowing Task**

For the shadowing task, participants were given a French word on screen and were able to play a recording of that word up to three times. The participant would then record themselves repeating the word that was uttered one at a time in a procedure identical to the elicited production task. The experiment was programmed and given in the online platform Labvanced. In order to control for language mode, the language specific word lists were presented in separate sessions, and the order of the languages (Spanish

first or English first) was counter-balanced across participants. All participants ended the experiment with the L3 shadowing task.

### **2.2.3 Analysis plan**

The results will be analyzed using a Bayesian Multilevel Regression where the outcome will be VOT and the fixed effect, categorical predictor will be language, and the fixed effect continuous predictor variables LexTALE scores, and a dominance metric.

## **2.3 Simulated Results**

## **Chapter 3: Vowel Production at first exposure to an L3**

### **3.1 Research Questions**

### **3.2 Methods**

#### **3.2.1 Tasks**

#### **3.2.2 Participants**

### **3.3 Simulated Results**

## Chapter 4: Perception of L3 vowels

### 4.1 Research Questions

RQ1: Will Spanish-English bilinguals assimilate L3 sounds to L1 or L2 categories?

RQ2: Will Spanish-English bilinguals perceive differences on an acoustic continuum using L1 or L2 boundaries?

### 4.2 Methods

#### 4.2.1 Participants

4 total groups of participants will take part in the experiment. There will be 2 bilingual groups with opposite orders of acquisition, Spanish L1-English L2, and English L1-Spanish L2, and monolingual English and monolingual Spanish participants. The bilingual groups will contain 75 participants per group, where the monolingual groups will contain 50 participants per group, for a total of 400 participants. These sample sizes were justified by a power analysis, in which a power level of .8 was desired (CITATION).

#### 4.2.2 Tasks

##### 4.2.2.1 AX Discrimination Task

The /i/ - /u/ continuum will be used, in which French /y/ is an intermediate. At this time, a second continuum is also being considered.

##### 4.2.2.2 Category Identification Task

The purpose of this task is to identify whether French segments are assimilated to Spanish or English categories. Four conditions are proposed for vowels (1) The same vowel sound is found in both languages, such as /i/, which can be found in all 3 languages. (2) A sound which is found in English, but not in Spanish, such as /e/ in “bed,” or the short /I/. (3) Only Spanish, such as the rounded back-vowel /o/, and

finally, (4) a vowel which is not present in English or Spanish, but is in French, such as /y/.

The stimuli will be created by splicing natural utterances of the French vowels together, in closed syllables, with a consonant (e.g. gVg). The stimuli will be played 5 repetitions each, which will create 20 tokens per participant.

#### **4.2.2.3 Analysis plan**

The results will be analyzed using a Bayesian Logistic Regression where the outcome of language choice (English or Spanish) will be modeled as a function of presented vowel sound, LexTALE score, and a continuous language dominance metric.

### **4.3 Simulated Results**

## REFERENCES