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 acquistion

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

LIST OF TABLES

LIST OF FIGURES

```r
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: Introduction

1.1 Introduction

The difficulty of learning a new language in 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 (**Flege & Eefting, 1988, Flege (1991)**). 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. Much less is known about the acquisition of a third language (L3A), particularly with regard to L3 production, perception, and phonological learning. The present dissertation explores the production patterns and perceptual categorization routines of adult bilinguals during the initial stages of L3 acquisition.

Empirical studies in L3 phonological acquisition have found evidence of multi-directional influence in language production (CITE), with the impact of previously learned languages on a third varying. In some cases, L2 influence on L3 production has been found, where in others, L1 influence or simultaneous influence of both languages on L3 production has been reported. It is unclear what factors could be at the root of these varied findings. On one hand, it might be sampling error. At large, empirical studies in L3 acquisition have used low samples to build models. One issue with this approach is related to sampling error. On the other, factors such as L2 and L3 proficiency, language dominance and choices of methods and analysis may play a role in the conclusions drawn by these studies.

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 homogeneous populations of bilinguals are in greater supply that 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. To deal with the potential impact of proficiency and dominance on L3 production and perception, L2 proficiency was measured using the LexTALE in English (citation) and Spanish (citation). The Bilingual Language Profile (BLP) was used to measure language dominance and background (citation).

Chapter 2: Models of third language acquisition

2.1 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. To date, the models of L3 acquisition of largely focused on morpho-syntax. 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.

2.1.1 The Cumulative Enhancement Model

The Cumulative Enhancement Model (CEM) (Flynn, Foley, & Vinnitskaya, 2004) was one of the first formal models of L3 acquisition in adulthood. The authors argued that L3 acquisition can provide unique insights about language learning that is not possible in L1 or L2 research alone. The key question in this article was whether the properties of the L1 maintain a privileged status in L3 acquisition, or if L3 acquisition is cumulative process, in which all grammatical properties of previously known languages impact subsequent language acquisition.

Within this seminal article, the authors conducted a study to provide evidence for the newly proposed CEM. The study examined the production of restrictive relative clauses in adults and children who spoke L3 English (L1 Kazakh-L2 Russian). The results revealed similar performance on the experimental task by adults and children, and that the production of relative clauses was influenced by the participants' L2 Russian. This influence was facilitative, since this syntactic structure is common in Russian and English, but not Kazakh. The authors note, however, that the results of this study could also be explained by a special status for the L2, rather than the L1.

Berkes & Flynn (2010) conducted a further study to empirically test the predictions of the CEM. German L1 and Hungarian L1-German L2 speakers were tested in English (their L2 and L3). The authors found evidence that the L3 group performed better than the L2 group, when the syntactic structure in question was L1-like, oppositely from (Flynn, Foley, & Vinnitskaya, 2004), which found facilitation when L2 and L3 elements were similar. The authors argue that the results of these two studies taken together suggest that both the L1 and the L2 may influence the L3 grammar, and that these languages' influence help to produce target-like L3 productions, when the L3 has a common feature with either the L1 or the L2. A proposed issue with this model is non-facilitation in the L3. Non-facilitation refers to when L3 performance is not target like and resembles performance in the L1 or L2. Several empirical studies have found evidence for non-facilitation in L3 tasks and would underly the predictions of further L3 models.

2.1.2 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. This prediction stems from the proposed cognitive similarity between the L2 and the L3. These proposed cognitive similarities stem from the Declarative-Procedural model, which posits that the grammar of late learned languages are largely subserved by the declarative memory system, whereas early learned languages are subserved by procedural memory. Procedural and declarative memory are long term memory systems which serve distinct general functions. The declarative memory is largely used to store explicitly known knowledge, such as factual information. The procedural memory, on the other hand, subserves implicit knowledge and procedures, such as riding a bike. During first language acquisition, learning is argued to be largely implicit and is seen as a procedure in this view.

Alternatively, L2 learning is associated with declarative memory and explicit learning of a grammar, at least to a larger extent that L1 learning (Paradis, 2009) As a result, the L2 Status Factor does not make predictions in L3 learning for simultaneous bilinguals, since it is argued that both of their languages are procedural. In summary, the L2

Status Factor predicts that, due to the cognitive similarity between a late-learned L2 and L3, the L2 will influence the L3 by default. Likewise, unlike the CEM, this influence will not always be facilitative, and is predicted to block access to the L1.

Several studies have found L2 influence in L3 tasks. In a low-sample study Bardel & Falk (2007) examined two sets of participants learning Swedish as an L3 (n = 5), and either Dutch or Swedish as the L3 (n = 4). The participants varied in their L1 and L2, and sometimes spoke three languages (which the authors classified as 2 L2s). This design was chosen to vary the order of acquisition of V2 and non-V2 languages, which differ in their placement of negation. An inspection of individual language background by the participants in p@bardel_role_2007] (p 471-472) reveals that 4 of the participants spoke a non-V2 L1 and a V2 L2, and the remaining 5 spoke a V2 L1 and non-V2 L2. The researchers were investigating whether cross-linguistic similarity (the presence or absence of V2 in the L1 or L2) affected the production of V2 in the L3. The results provided evidence that that the group with the V2 L2 outperformed the non-V2 L2 group. The authors took this result as evidence for a privileged status of the L2 in L3 production, despite its lack of facilitation. These results also contradict the predictions of the CEM, which suggested that L3 learning is a cumulative process that takes advantage both L1 and L2 grammars in L3 learning.

The L2 Status Factor model has seen some empirical support in studies examining L3 syntax. **Bayona (2009)** examined the acquisition of middle and impersonal passive constructions in L2 and L3 Spanish. The L2 group consisted on mostly English L1 speakers, and the L3 group consisted of English L1 and French L2 speakers. The author found that the L3 Spanish group was more accurate in rejection or acceptance of target-like use of middle and passive constructions in L3 Spanish than the L2 group. These results suggest that the L2 influenced that L3, and provided evidence for both the L2 Status Factor and the Cumulative Enhancement Model. In another early study, Leung (2005) investigated the acquisition of determiner phrases and adjective word order in French as an L2 (L1 Vietnamese) and L3 (L1 Cantonese-L2 English). Using a battery of tests, the authors concluded that the L3 group experienced influence from

their L2 in both determiner phrases, which was facilitative in this case, and adjective order, which resulted in non-target like pre-nominal French adjectives. In this case, the results could arguably be accounted for by the L2 Status Factor, while the CEM cannot explain non-facilitation. However, the author's narrative conclusion of these results did not claim that there was sole access to the L2, or that the L1 was blocked. Rather, they suggest that L3 learning is not simply another case of L2 learning, in which the L1 is the sole influence. Interestingly, later studies found that L1 access was possible when learners had higher metalinguistic awareness of cross-linguistic similarity between the L1 and the L3. In a study of XX learners of L3 XXXX, *citation*, participants were given a survey which measured their awareness of cross-linguistic features in their known languages. The results revealed that participants who scored higher on the metalinguistic survey also showed evidence of facilitative influence of their L1. This outcome led the authors to revise the predictions of the L2SF to include the potential for L1 influence when metalinguistic awareness is high.

"Bohnacker (2006), who finds non-facilitative influence of L2 English non-V2 word order (XSV) on the L3 German of L1 Swedish learner"

add falk and bardel 2011

2.1.3 The Typological Primacy Model

Later studies in L3 morphosyntax would first aim to demonstrate that influence of previously known languages in L3A is not always facilitative, and would then suggest an alternative explanation and methodology for the results found in previous studies supporting L2 influence in L3 acquisition. In one such study, Rothman & Cabrelli Amaro (2010) aimed to demonstrate that the influence of previously known languages is not always helpful, as the CEM predicts, by examining null subjects in L3 learners. The study included two groups of Spanish-English bilinguals who were learning either L3 French or L3 Italian. Importantly, all participants spoke L2 Spanish, which allows for null subjects, and spoke either L3 French (obligatory overt subjects) or L3 Italian (optional null subjects). As a result, dropping a subject in L3 French would result

in non-target like production, where the opposite would be true in L3 Italian. The authors found that the data from this study provided evidence that the L2 was the influence on L3 production in both L3 groups, and provided evidence against the CEM, since the French group experimented non-facilitative effects from a previously learned languages, and the L2 Status Factor, since neither group showed evidence of L1 influence. Despite these results, the authors posited that it was possible that the perception of cross-linguistic structural similarity (or so called phycho-typology **Kellerman, 1983**) could also be a potential explanation for the results obtained in (Rothman & Cabrelli Amaro, 2010).

Eventually, this perspective was formalized as The Typological Primacy Model (Rothman, 2010, 2011, 2013, 2015) (TPM). TPM, 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, and that which language transfers to the L3 is determined by psychotypology after exposure to input. Rothman (2015) explains that L3 input is parsed in a hierarchical manners in which the lexicon is the primary cue parsed to determine cross-linguistic similarity between the L3 and the languages known by the L3 learner, followed by phonological or phonotactic cues, functional morphology and finally by syntactic structure. Importantly, the transfer of one grammar holistically is predicted to occur during the initial stages of acquisition. This idea of initial stages is described by Rothman as "very early in the L3 process" Rothman (2015, p. 180). At what the TPM refers to the "absolute initial state," it suggests that both languages are in principle available to the L3 learner, but does not elaborate or make predictions as to how features of the L1 and L2 will affect absolute initial state. As a result, the TPM does not make clear predictions in the context of the present dissertation, since the learners examined here are intended to perceive and produce the L3 at their very first exposure. Nonetheless, the TPM would best be able to account for results in which L3 speakers produce and perceive the L3 in a practically equivalent manner to their L1 or L2 in a cross-sectional design, rather than solely the L2.

Several published studies include conclusions in which the TPM is supported. In the

discussion of Rothman & Cabrelli Amaro (2010), the authors suggested that it could not be determined whether L3 performance was due to L2 effects or psychotypology.., As a result, the authors suggested the use of language combinations in which the order of acquisition of L1 and L2 are reversed. These groups, sometimes referred to as mirror-image groups, would include groups such as L1 Spanish-L2 English and L1 English-L2 Spanish speakers who learned the same L3. In this view, between-group performance should be similar if transfer is driven by psychotypology, and groups should behave differently if L2 status is deterministic. For example, if mirror-image groups had been used in Rothman & Cabrelli Amaro (2010) with L1 Spanish-L2 English and L1 English-L2 Spanish groups, the acceptance or rejection of L3 null subjects would allow for the examination of the predictive power of L2 status and psychotypology Using this methodology, Rothman (2011) aimed to investigate whether L3 Brazilian Portuguese would be influenced by Spanish both when it was and was not the L2. **add more detail on this study** The first group consisted of L1 English-L2 Spanish speakers, while the second group spoke L1 Spanish and L2 English. Based on the absence of a main effect for group in a one-way ANOVA, Rothman (2011) concluded that there was similar performance between groups and used this as evidence for the basis of the TPM.

Following the formal introduction of the TPM, further studies included mirror-image groups in their design and aimed to test the predictions of the model relative to the L2 Status Factor. Several of these studies involve the acquisition of L3 Brazilian Portuguese (BP) by Spanish-English bilinguals in both orders of acquisition. Montrul, Dias, & Santos (2011) examined the production of clitic and object expressions in L3 BP by mirror-image groups of Spanish-English bilinguals. Using an oral production task and a written acceptability judgment task, (AJT), it was found that L3 BP was influenced by Spanish whether it was the L1 or the L2 of the participants. Examining the same language combination and using similar mirror-image groups, Giancaspro, Halloran, & Iverson (2015) also found investigated the use of differential object marking (DOM) in L3 BP. The results suggested that Spanish influenced L3 BP whether it was

the L1 or the L2. Like Montrul, Dias, & Santos (2011), Parma (2017) also investigated L3 BP clitic development and expanded upon previous studies by including both a perception and production task. Another novelty of Parma (2017) was the inclusion of an online measure, a self-paced reading task, to measure comprehension, while the study also included a story-telling task to measure production. The results of the comprehension task were inconclusive, and it could not be conlcuded that the experimental conditions were being processed at meaningfully difference speeds by any group. The production task, on the other hand, found evidence of errors in clitic production in L3 BP by both L1 and L2 Spanish speakers that could be likened to Spanish influence. The results of these studies provided evidence contrary to the predictions of the L2 Status Factor and suggest that Spanish influences L3 BP. Following these initial results in the language combination of L3 BP with English and Spanish, other studies aimed to test whether either the L1 or the L2 could influence performance on L3 tasks. The evidence outside of the BP/Spanish/English triad has been more limited and controversial. Some accounts in favor of the TPM's predicitve power have aimed to accumulate evidence in its favor. For instance, in a recent systematic review of 92 studies in L3 acquisition, Puig-Mayenco, González Alonso, & Rothman (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. Importantly, the findings were not coded in a mutually exclusive manner, since 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). A closer examination of this systematic review reveals, however, that the coding procedure did not follow a clear objective criterion. That is, there are instances in which the coding provided by the authors in the appendix contradicts the narrative conclusions of the cited studies. **update totals** In other words, the cumulative evidence for L1 or L2 influence on L3 behavior as predicted by the TPM

is unclear outside of 2 Romance Language and English triads.

Limitations of the TPM in addition to the lack of robust and cross-linguistic include overstated predictive power and over interpretation of results. The potential problem when it comes to the predictive power of the TPM is the vagueness associated with the term "initial stages" of L3 acquisition. Puig-Mayenco, González Alonso, & Rothman (2020) 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. The problem of overinterpretation can be seen in the use of single language properties and inappropriate statistical tests to justify narritive conclusions. For example, the TPM is rather explicit in its prediction that languages transfer holistically to the L3, rather than on a property-by-property basis that the CEM would suggest. While, the evidence gathered to date provides counterevidence to sole L2 influence on the L3, it is unclear how these studies support holistic transfer. In other words, it does not seem that the body of evidence in L3 BP can rule out the influence of English within the property examined, since small samples which provide inconclusive results (non-signigicant p-values) do not entail equivalence within or between subjects. Such small samples and choices of methods do not allow for potential gradient and small effects of co-activation to be observed, since, on one hand, it is not possible to determine whether group results have to do with sampling issues, and, on the other, clear critria for co-activation of both languages known by a bilingual was not explicitly included in predictions and design of these studies. As a result, it cannot be determined based on the results of studies to date that, in the case of the L3 BP studies, Spanish is solely activated and that English is not activated. Telling this would require much larger samples and clear criteria for interpreting how results support models prior to data collection, which could include potential for non-binary interpretation of data.

2.1.4 The Linguistic Proximity Model and the Scalpel Model

The prediction that a sole language will influence L3 production and perception is not shared by all models. More recent models of L3 acquisition, like the CEM, predict that both the L1 and the L2 are available to influence an L3. For example, the Linguistic Proximity Model (LPM) predicts that there is full transfer potential (FTP) of either linguistic system, but that this occurs in a gradient fashion and on a property-by-property basis **cite**. In the founding article of the LPM, Westergaard, Mitrofanova, Mykhaylyk, & Rodina (2017) posed several research questions. In addition to investigating whether a sole language influences an L3, they also examined whether CLI comes from the more typologically similar language, and whether this influence is facilitative. To provide evidence that could aid in answering these questions, the authors recruited 22 Norwegian-Russian simultaneous bilinguals who spoke English as an L3. The participants completed a binary Grammaticality Judgment Task related to verb movement. The results of the tasks indicated that, compared to L1 Norwegian speaking children, bilingual children were able to benefit from their knowledge of Russian when learning L3 English. At the same time, the bilingual participants did not perform as well as L1 Russian children learning English. The authors interpreted these results, taken together, as influence from both the L1 and the L2 in the L3 English of the Russian-Norwegian bilingual children. That is, the intermediate score of the Russian-Norwegian simultaneous bilingual children learning L3 English relative to the comparison groups was taken as evidence of facilitative influence from their Russian and non-facilitative influence from their Norwegian.

The LPM, like the TPM and CEM, predicts that similarities between languages plays a major role in L3 acquisition, rather than order of acquisition, as the L2 Status Factor predicts. The LPM departs from the TPM in that it predicts that abstract structural properties causes CLI, rather than general typological proximity. In other words, it suggests that cross-linguistic influence is decided on a feature-by-feature basis, rather than generalizing whole language predictions. Additionally, the LPM predicts that all languages are available to the L3 learner throughout the learning

process, unlike the TPM.

A methodological consideration advocated for by the LPM is the use of subtractive groups. Unlike mirror image groups, which seek to compare two trilingual groups, subtractive groups compare L3 learners to L2 learners. In the case of the studies to date, L3 learners of English who speak Russian and Norwegian were compared to Russian L1/English L2 and Norwegian L1/English L2 groups. It has been argued that this design allows for gradient effects of both the L1 and L2 to be observed. That is, intermediate values in L3 performance on experimental tasks relative to the L2 performance of the comparison groups is taken as evidence of co-activation of both languages. On the other hand, if no difference could be found between L2 and L3 learners of the same language, the LPM would consider this evidence for the influence of a single language on the L3.

Following the introduction of the model, several studies have tested the predictions of the LPM.

Mitronova & Westergaard (2018) "Norwegian-Russian bilinguals and Norwegian monolinguals in a sentence/picture-matching experiment using an artificial language (Aliensk)" The language contained accusative/nominative case marking distinctions, a feature from Russian, and lexical items that were Norwegian-like. The results showed that bilinguals were better than monolinguals at matching sentences to pictures than Norwegian monolinguals, who appeared to rely on solely on word order.

Stadt, Hulk, and Sleeman (2016, 2018a, 2018b, 2020),

Ben Abbes (2016, 2020)

Kolb et al. under review <- accepted international journal of bilingualism Lloyd-Smith (2020)

Mitrofanova N and Westergaard M (2018) Wholesale vs. property-by-property transfer: Acquisition of morphological case in an artificial L3. Poster presented at

2nd International Symposium on Bilingual and L2 Processing in Adults and Children (ISBPAC), Braunschweig, May 24-25.

Westergaard, Marit, Natalia Mitrofanova, Yulia Rodina & Roumyana Slabakova. Submitted. Full Transfer Potential in L3/Ln acquisition: Crosslinguistic influence as a property-by- property process. The Cambridge Handbook of Third Language Acquisition and Processing.

Kolb, N., Mitrofanova, N. & Westergaard, M. (under review). Cross-linguistic Influence in Child L3 English: An Empirical Study on Russian-German Heritage Bilinguals.

Jensen et al. (2021) investigated similar groups to Westergaard et al. (2017), in which simultaneous Russian-Norwegian bilinguals were compared to subtractive L1 Russian-L2 English and L1 Norwegian-L2 English groups, but included seven total linguistic properties. The results indicated that the L3 group experienced facilitative influence in some cases, but non-facilatative influence in others. Additionally, the sources of this influence could be likened to either Russian or Norwegian.

Scholars have argued that the LPM makes vague predictions 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?**). However, Westergaard et al. (submitted) argue that the LPM does make specific predictions, since it is predicted that accuracy should fall between L1 and L2 values on experimental tasks. The authors state that, when using subtractive groups, if the L3 group performance falls above or below the one of the two L2 groups, then then model would be falsified.

It is notable that the empirical studies that serve as the basis of the LPM **westergaard studies** utilize groups of simultaneous bilinguals, which do not allow for the examination of potential L2 status effects. Following this idea, Westergaard,

Mitrofanova, Mykhaylyk, & Rodina (2017) suggest that an optimal design to examine both the invididual contributions of languages, as well as potential language status effects, would be the so-called fully combined design. This suggestion entails the use of both mirror-image groups, as seen in the TPM studies (e.g. L1 Spanish-L2 English-L3 BP and L1 English-L2 Spanish-L3 BP), and subtractive groups (e.g. L1 Spanish-L2 English-L3 BP, L1 Spanish-L2 BP and L1 English-L2 BP) used in the LPM studies. As a result, a fully combined design would result in the use of 6 total groups, in which a mirror image design contains L1 speakers of both background languages who learn the L3 as an L2.

Chapter 3: Models of L2 phonological acquisition

3.1 Models of L2 phonological acquisition

Notably, much of the evidence for the L3 models to date come from studies which examine the acquisition (morpho)syntactic features. The models do not always spell out specific predictions when it comes to L3 phonology. In the case of the TPM, full transfer could be taken to mean the transfer of an entire language system to the L3, including phonology, but this stipulation is not explicitly spelled out in the recent articles articulating the predictions and motivations of the TPM (**cite**). In the LPM, phonology is also not directly addressed, but there does not seem to be any reason why the same predictions should not apply to phonology as syntax. Specifically, the LPM should predict that L3 performance should be intermediate and fall between L1 and L2 performance on experimental tasks. If predictions of L3 models are analogous in phonology and syntax, then the predictions and expansion of models of speech learning

Many accounts for L2 phonological acquisition exist. In general, these models all have in common that language-specific L1 categories drive L2 speech development. It is still not well known how the predictions of the predictions of these models apply to L3 speech development when L2 categories are also theoretically avialable to the L3. In the section that follows, a brief overview of each model, along with its evidence, will be covered. Following the model introductions, its relationship and proposed expansion to L3 phonology will be discussed.

3.1.1 The Speech Learning Model

The original SLM **Flege (1995)** focuses on the acquisition of segments in the L2 and emphasizes the importance of cross-linguistic similarity in L2 speech learning. In this view, new segments are predicted to be easier to learn, where segments which have close matches in the L2 will be much harder to acquire at a native-like level. Essentially, the SLM proposes that L1 and L2 sounds are linked through a process

referred to as "interlingual identification," in that the L2 sound encountered as first exposure are seen as either good or bad phonetic variants of a native category **(Flege, 1995)**. In this view, acquisition of L2 segments occurs after increased exposure to these segments, where segmental complexity modulates the rate of learning **(Flege, 2021)**. Additionally, the SLM posits that allophonic variation and segment position matter, as opposed to the learning of a phoneme generalizing across positions.
For instance, **Iverson, Hazan, and Bannister (2005)** conducted a study in which they succesfully trained Japanese speakers to better identify /r/ and /l/ in word initial position, but that this training was not effective for word-medial or word-initial clusters containing liquids.

The original SLM also suggested that the age of first exposure (the younger the better), and L2 experience were important predictors to L2 learning success. The SLM makes a distinction between identification as opposed to categorization. The importance of this distinction involves the learning of phonetic variants during the learning process, which lead to new L2 phonetic category formations. For example, using a two-alternative forced choice task, **Bohn and Flege (1993)** asked Spanish monolinguals to identify English stops and provided them with pre-voiced, short lag and long lag stops as tokens. Importantly, Spanish monolinguals identified english stops identified long lag /t/ as /t/, despite the lack of a long-lag category in their native inventory.

3.1.1.1 The revised SLM

The SLM was recently revised, and has updated some assumptions and predictions. The revised model maintained many of the same assumptions as the original, and importantly focuses on sequential bilinugals, where speech learning begins when the phonetic categories of the L1 have been established. According to **Flege (2021)**, the updated tenets of the Revised Speech Learning Model (SLM-r) are that, first, phonetic categories are formed based on their statistical regularities in the input, such that greater exposure to a particular phonetic cue should be correlated with the re-tuning

of that cue towards the input. Second, all learners make use of the same learning mechanisms in L2 learning that they do in native language learning. That is, the SLM-r refutes the idea of a critical period (**CITE**) for speech learning, and suggests that non-native production in perception is due to differences the quality and quantity of input between native speakers and L2 learners and L1 effects. Another important update to the SLM-r is the idea that perception and production co-evolve. In other words, the SLM-r does not predict that perception will precede production.

The revised model has not yet received empirical support, due to its recent revision. It is outside the scope of the present dissertation to evaluate the predictions of the SLM-r when it comes to statistically driven re-tuning of phonetic categories. However, the phonetic categories that are in place in a bilingual at the first exposure to L3 learning could be treated as an integrated inventory when it comes to L3 learning. Based on the idea that the same mechanisms are in place in L1 and L2 phonetic category formation, the same mechanisms should be in place for L3 phonetic category formation. In this view, the acoustic similarity of an L3 segment to an L1 or L2 segment should predict how difficult learning an L3 segment will be. In this view, there is not reason to predict a blocking of one language system (L1 or L2) provided that these systems have well-established phonetic categories.

3.1.2 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 experiened 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 closet mathcing 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, **Escudero and Chládková** (2010) found that Spanish L1 speakers assimilated SSBE /ae/ and /a/ to Spanish /a/ (that discrimination of this contrast was difficult). Additionally, **Escudero et al. (2014)** provided evidence that Salento Italian L1 speakers assimilated SSBE /ae/, /a/ and /ˆ/ to Salento Italian /a/, while **Escudero & Vasiliev (2011)** Spanish speakers assimilated Canadian English /ae/ and /e/ to Spanish /a/. Finally, **Escudero & Williams (2011)** found that Spanish listeners categorized Dutch /a/ and /a:/ in terms of their native /a/

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 similary 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.

3.1.3 The Second Language Linguistic Perception Model

The Second Language Linguistic Perception Model (L2LP_ is a computational model of L2 speech learning and is similar to the PAM in that it focuses on sound contrasts as the basis for L2 speech learning, rather than single segments (**van Leussen & Escudero, 2015**, Escudero, 2005, 2009). In this view, it is difficult for L2 learners to make contrasts which are not present in their L1. In their revision of the L2LP,

van Leussen & Escudero support this claim with several empircal studies. Namely, that these studies have provided evidence that L2 learners experience difficult with the contrasts of /r/ and /l/ in Japanese, (**Aoyama et al., 2004**) "beat" and "bit" in Spanish and Portuguese (**Flege et al., 1997; Rauber et al., 2005**) "bet" and "bat" in Dutch (**Broersma, 2005**). The authors argue that these results suggest that linguistic experience is at the heart of L2 learning. Specifically, cross-linguistic comparisons of L1 categories and L2 categories are thought to predict the ease of L2 category learning. The L2LP, unlike the PAM and the SLM, also aims to model the entire learning process, rather than the beginning stages of L2 speech learning. This entire learning process is predicted computationally and is baed on Stochastic Optimality Theory (Boersma, 1998).

An important tenet of the postulate of the L2LP is the optimal perception hypothesis. Essentially, this proposal suggests that the initial perception of L2 sounds is the result of L1 acquisition. The developmental of L2 learners is predicted similarly to PAM, where a single category assimulation from PAM, in which a native category must be split, is called a new category scenario in the L2LP. This scenario is predicted by both models to be difficult for the L2 learner. On the other hand, the PAM and the L2LP predict that the case when two L1 sounds correspond well to two L2 sounds that discrimination of these sounds will be relavitely easier. This is referred to as a single category assimilation in PAM and a similar scenario in the L2LP. The final scenario of the L2LP is the subset scenario. In this case, a single L2 phoneme is perceived as two L1 categories. The same case is referred to as uncategorized or categorized-uncategorized in the PAM, and both models predict that discrimination of these sounds will be better than the case of new scenarios, but not as good as discrimination of subset scenarios.

Empirical support has been found for these predictions. In two-category assimilation/similar scenario, **Escudero and Boersma (2002)**, native Dutch speakers assimilated the Spanish /i/ to their native Dutch /i/, and the Spanish /e/ to their Dutch /I/.

However, the L2LP predicts that the similar (two-category assimilation) scenario may lead to inappropriate lexical contrasts, and argues that pre-lexical and lexical contrasts should be taken into account when it comes to sound discrimination. Evidence for this claim stems from studies which found that L2 learners could not perceive a contrast in lexical items that they could disriminate outside of lexical items (**Curtin et al., 1998**). Other studies show that some lexical items can also be relaibly distinguished by L2 learners that could not be told apart pre-lexically **(Weber and Cutler, 2004; Cutler et al., 2006; Escudero et al., 2008)**

As a result, an important tenet of the revised L2LP (**cite**) is that meaning-driven learning predicts the developmental path of L2 phoneme perception. The L2LP simulates the entire trajectory of L2 learning based on the various learning scenarios. In the revised L2LP, **van Leussen & Escudero, 2015** suggest that, despite past null results (**Weiand, 2007**), category reduction is possible when it is driven by meaning based learning.

The predictions of the L2LP which are important for the present dissertation are the Full Copying hypothesis (**Escudero, 2005**). If L3 phonological learning is another instance of L2 learning, then the L2LP and the TPM should share the prediction that the initial state of L3 learning is the end state of either L1 or L2 learning, but not both. In this view, bilingual participants who are first exposed to an L3 should produce and categorize L3 sounds as similarly to a single language, rather than producing or perceiving some sounds as L1-like and others as L2-like. Additionally, these participants' behavior on the experimental tasks should a) resemble either their own L2 behavior or L1 behavior or b) in the case of L1 influence, resemble a monolingual comparison group who is first exposed to an L2 (e.g. Spanish L1, English L2, exposed to German, and Spanish L1 exposed to German should behave similarly if Spanish is language which is "fully copied" at first exposure.)

Chapter 4: Previous literature in L3 phonology

4.1 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.

4.1.1 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, 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, 2010). S vowel production (**kamiyama_acquisition_2007?**) vowel reduction and speech rhythm (Gut, 2010).

Similar patterns of L2 influence have been found in VOT productions. Llama, Cardoso, & Collins (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,

LPM support Wrembel 2015, Jabbari and Pourmajnoun 2016, Lloyd-Smith, Gyllstad and Kupisch 2017

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, 2014) measured VOT and aspiration in all languages of participants with two different language combinations: 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, 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, 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 re-

ported 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 & Cardoso, 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 suggested 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.

4.1.2 L3 Perception studies

To date, few studies in L3 phonology has incorporated the predictions of the perceptual assimilation model

Wrembel et al. (2019) examined the categorization and discrimination of 10 young trilinguals who spoke L1 German-L2 English-L3 Polish. The results showed evidence that participants assimilated L3 sounds to both L1 and L2 categories, but preferred the L2. Additionally, language specific phonetic discrimination were attended to by even L3 beginners. **Balas (2019)**

Balas (2019), Liu

Werker (1986), Patihis, Oh and Mogilner (2015), Kopečková (2015), Cabrelli Amaro (2016), Onishi (2016), and Wrembel, Marecka and Kopečková (under review).

*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, Gut, Kopečková, & Balas (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, Gut, Kopečková, & Balas, 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.*

Chapter 5: Methods and anylsis in previous work

5.1 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 empircal 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 as 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 in 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 interpreation 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.

Examples of non-main effects as equivalence:

with ancova "The non-significant value for the 2-way language group*level interaction (p = .410) suggests that the proficiency effect was statistically equivalent for the two language groups,

Moreover, the 3-way language group*level*type of relative interaction (p = .088) gave a non-significant value, which suggests that the effect of proficiency is not significantly different across levels of the two studies and sentence type. " Berkes and Flynn (2010, p. 153)

L3 Study: HungarianL1/GermanL2/EnglishL3: Berkes and Flynn (2010, p. 156) As evidence for similar performace on all 3 types of relative clause production: lack of sig. p-value, no main effect or sig. pariwise comparison

5.2 Cross-lingusitic 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.

5.2.1 Vocalic systems

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

5.2.2 Use of Voice-Onset Time

The description of cross-lingusitic 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 & Abramson, 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 & Abramson, 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 & Abramson, 1964).

5.3 Bringing L2 speech models and L3 models together and evaluating predictions

By using the methods used to test the prediction of models of L2 phonological acquisition (The SLM, the PAM, and the L2LP), more nuanced evidence may be obtained to evaluate the predictive power of L3 models. With the revision of the SLM, at least three L2 speech models advocate for the study of naive or beginning learners in L2 speech learning research. The present dissertation adopts this point of view and applies it in a third-language context in perception and production. By measuring the perception and production patterns of the first (or, at least, very early) exposure to a third language, coupled with measurements in each language, an the variation of cross-linguistic influence can be observed/studied in L3 perception and production. If, as the SLM-r predicts, category formation is driven by input and retuning of L1 categories in L2 acquisition, and the same mechanisms that are used in L1 phonetic category formation in L2 phonetic category formation, then it is likely that the SLM would predict that those same mechanisms are at play in L3 phonetic category formation, and that this process is guided by L3 input.

5.4 Perception

Following this logic, the question becomes whether L1 or L2 categories are retuned to L3 categories, and what conditions determine which language category is initially chosen and its rate of retuning. The present dissertation focuses on which category is chosen as the initial L3 category, while the rate of change in these categories is left for future research. One possibility which may influence whether an L1 or L2 category influences an L3 is the acoustic similarity of the L3 segment relative to an individuals' L1 and L2 phonetic categories. The TPM predicts that phonetic cues do play a role in L3 input parsing, in that they are parsed and used in order to make a decision of which language system to holistically transfer. However, the TPM would (likely) not predict that two segments would be assimilated to two distinct languages, and it would not predict that the same segment would be categorized as different sounds.

Additionally, if one language holistically impacts the acquisition of a third at first exposure, then the behavior of L3 learners should resemble L2 learners, provided the L1 is the source of influence in L3 acquisition, or should be practically equivalent in a within-subject comparison. In the case of the predicted behavior of L2 influence on L3 productions, L2 production and perception should be practically equivalent in a within-subject comparison.

The predictions of the Linguistic Proximity Model lack predictive power, but would be able to account for the same L3 sound being categorized differently by the same subject, and by different subjects, and for two different L3 sounds being categorized as an L1 and L2 category respectively.

Chapter 6: Production at first exposure to an L3

6.1 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.

6.1.1 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?

6.2 Methods

6.2.1 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.

6.2.2 Tasks

Oragnization paragraph: An elicted production task & A shadowing task in the L3.

6.2.2.1 Bilingual Language Profile

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

6.2.2.2 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.

6.2.2.3 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.

6.2.2.4 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.

6.2.3 Analysis plan

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

6.3 Simulated Results

Chapter 7: Vowel Production at first exposure to an L3

7.1 Research Questions

7.2 Methods

7.2.1 Tasks

7.2.2 Participants

7.3 Simulated Results

Chapter 8: Perception of L3 vowels

8.1 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?

8.2 Methods

8.2.1 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).

8.2.2 Tasks

8.2.2.1 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.

8.2.2.2 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.

8.2.2.3 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.

8.3 Simulated Results

REFERENCES

Bardel, C., & Falk, Y. (2007). The role of the second language in third language acquisition: The case of Germanic syntax. *Second Language Research*, *23*(4), 459–484. http://doi.org/10.1177/0267658307080557

Flynn, S., Foley, C., & Vinnitskaya, I. (2004). The Cumulative-Enhancement Model for Language Acquisition: Comparing Adults' and Children's Patterns of Development in First, Second and Third Language Acquisition of Relative Clauses. *International Journal of Multilingualism*, *1*(1), 3–16. http://doi.org/10.1080/14790710408668175

Giancaspro, D., Halloran, B., & Iverson, M. (2015). Transfer at the initial stages of L3 Brazilian Portuguese: A look at three groups of English/Spanish bilinguals. *Bilingualism: Language and Cognition*, *18*(2), 191–207. http://doi.org/10.1017/S1366728914000339

Gut, U. (2010). Cross-linguistic influence in L3 phonological acquisition. *International Journal of Multilingualism*, *7*(1), 19–38. http://doi.org/10.1080/14790710902972248

Jensen, I. N., Mitrofanova, N., Anderssen, M., Rodina, Y., Slabakova, R., & Westergaard, M. (2021). Crosslinguistic influence in L3 acquisition across linguistic modules. *International Journal of Multilingualism*, 1–18. http://doi.org/10.1080/14790718.2021.1985127

Leung, Y.-K. I. (2005). L2 vs. L3 initial state: A comparative study of the acquisition of French DPs by Vietnamese monolinguals and Cantonese–English bilinguals. *Bilingualism: Language and Cognition*, *8*(1), 39–61. http://doi.org/10.1017/S1366728904002044

Lisker, L., & Abramson, A. S. (1964). A Cross-Language Study of Voic-

ing in Initial Stops: Acoustical Measurements. WORD, *20*(3), 384–422. http://doi.org/10.1080/00437956.1964.11659830

Llama, R., & Cardoso, W. (2018). Revisiting (Non-)Native Influence in VOT Production: Insights from Advanced L3 Spanish. *Languages*, *3*(3), 30. http://doi.org/10.3390/languages3030030

Llama, R., Cardoso, W., & Collins, L. (2010). The influence of language distance and language status on the acquisition of L3 phonology. *International Journal of Multilingualism*, *7*(1), 39–57. http://doi.org/10.1080/14790710902972255

Meisel, J. M. (1983). Transfer as a second-language strategy. *Language & Communication*, *3*(1), 11–46. http://doi.org/10.1016/0271-5309(83)90018-6

Montrul, S., Dias, R., & Santos, H. (2011). Clitics and object expression in the L3 acquisition of Brazilian Portuguese: Structural similarity matters for transfer. *Second Language Research*, *27*(1), 21–58. http://doi.org/10.1177/0267658310386649

Paradis, M. (2009). *Declarative and Procedural Determinants of Second Languages* (Vol. 40). Amsterdam: John Benjamins Publishing Company. http://doi.org/10.1075/sibil.40

Parma, A. (2017). Cross-Linguistic Transfer of Object Clitic Structure: A Case of L3 Brazilian Portuguese. *Languages*, *2*(3), 14. http://doi.org/10.3390/languages2030014

Puig-Mayenco, E., González Alonso, J., & Rothman, J. (2020). A systematic review of transfer studies in third language acquisition. *Second Language Research*, *36*(1), 31–64. http://doi.org/10.1177/0267658318809147

Rothman, J. (2010). On the typological economy of syntactic transfer: Word order and relative clause high/low attachment preference in L3 Brazilian Portuguese. *IRAL - International Review of Applied Linguistics in Language Teaching*, *48*(2-3). http://doi.org/10.1515/iral.2010.011

Rothman, J. (2011). L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research*, *27*(1), 107–127. http://doi.org/10.1177/0267658310386439

Rothman, J. (2013). Cognitive economy, non-redundancy and typological primacy in L3 acquisition: Initial stages of L3 Romance and beyond. In S. Baauw, F. Drijkoningen, L. Meroni, & M. Pinto (Eds.), *Romance Languages and Linguistic Theory* (Vol. 5, pp. 217–248). Amsterdam: John Benjamins Publishing Company. http://doi.org/10.1075/rllt.5.11rot

Rothman, J. (2015). Linguistic and cognitive motivations for the Typological Primacy Model (TPM) of third language (L3) transfer: Timing of acquisition and proficiency considered. *Bilingualism: Language and Cognition*, *18*(2), 179–190. http://doi.org/10.1017/S136672891300059X

Rothman, J., & Cabrelli Amaro, J. (2010). What variables condition syntactic transfer? A look at the L3 initial state. *Second Language Research*, *26*(2), 189–218. http://doi.org/10.1177/0267658309349410

Westergaard, M., Mitrofanova, N., Mykhaylyk, R., & Rodina, Y. (2017). Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism*, *21*(6), 666–682. http://doi.org/10.1177/1367006916648859

Wrembel, M. (2010). L2-accented speech in L3 production. *International Journal of Multilingualism*, *7*(1), 75–90. http://doi.org/10.1080/14790710902972263

Wrembel, M. (2011). Cross-linguistic Influence in Third Language Acquisition of Voice Onset Time. In *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 2157–2160). Hong Kong: City University of Hong Kong.

Wrembel, M. (2014). VOT Patterns in the Acquisition of Third Language Phonology. In *Proceedings of the International Symposium on the Acquisition of Second Language Speech* (Vol. 5). Concordia Working Papers in Applied Linguistics,.

Wrembel, M., Gut, U., Kopečková, R., & Balas, A. (2020). Cross-Linguistic Interactions in Third Language Acquisition: Evidence from Multi-Feature Analysis of Speech Perception. *Languages*, *5*(4), 52. http://doi.org/10.3390/languages5040052