[INSERT TITLE OF DOCTORAL DISSERTATION]

By

[INSERT AUTHOR'S NAME]

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 [insert your graduate program's official name]

Written under the direction of

[insert Dissertation Director's name]

And approved by

____

____

____

____

New Brunswick, New Jersey

[Month] [Year]

ABSTRACT OF THE DISSERTATION

[INSERT TITLE OF DOCTORAL DISSERTATION]

by

[INSERT AUTHOR'S NAME]

Dissertation Director:

[insert Dissertation Director's name]

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

Insert your acknowledgements here.

DEDICATION

Insert your dedication here (optional).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

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 (James Emil Flege, 1991; James Emil Flege & Eefting, 1988). 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 (Llama & Cardoso, 2018; Llama, Cardoso, & Collins, 2010), and have found that the impact of previously learned languages varied in their relative influence on the third language. 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.

At large, empirical studies in L3 acquisition have used low samples and their results have been used as the basis of predictive models. An issue with the use of few participants to make generalizations is that it increases the likelihood of false positive and false negative results due to an insufficient quantity of samples from a hypothetical population distribution to reliably infer information about that group. In addition to low samples, L3 studies typically use analyses such as an ANOVA, t-tests, non-parametric tests against zero, such as the Kruskal Wallis test. An issue with these choices of analyses is the lack of evidence for the null hypothesis within the larger umbrella of the frequentist approach to inferential statistics. That is, in the case of L3 acquisition, where the interest is which of two source languages impacts a third, the researcher is interested in the possibility of providing evidence that L3 performance is either L1 or L2 like. Often, the lack of a so-called "significant" p-value in frequentist analysis is taken as evidence of similar performance between groups or within subjects. This conclusion is not well justified, since it is unclear if null results are due to low samples with wide confidence intervals and how much noise is expected in the data. That is, the lack of a significant p-value alone is not evidence for the null hypothesis. When only a p-value is reported, and not detailed descriptive statistics, it is difficult to establish objective and precise criteria for evidence for practical equivalence.

Additionally, 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. L3 proficiency has varied in the body of literature, with some studies examining beginners, while others study intermediate and advanced learners. Language dominance and use has not always been considered in L3 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. Secondly, more fitting statistical analyses are used to evaluate equivalence for within and between group comparisons, such as Bayesian Regression with a region of practical equivalence. In terms of sample size, it is probable that low samples in L3 research have occurred as a result of the difficulty in finding subjects with similar enough backgrounds in three languages. To work around the issue of low availability in participants, bilinguals' first exposure to a third language is studied in the present dissertation. It is likely that relatively homogeneous 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 (groups with the same L3 but opposite order of acquisition). 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.

Additionally, first exposure to the L3 can be very valuable information when it comes to modeling third language acquisition. Broadly, Third Language Acquisition (L3A) refers to the acquisition of a third language by an individual who already speaks to languages. Current models of L3A do not always account for potential individual differences in CLI, and typically are restrictive in their predictions. By investigating first exposure, it is possible to gather evidence regarding whether there is a default state of L1 and L2 CLI at first exposure to an L3, and whether this relative language influence is variable at the individual level. If the goal of L3 models is to predict the trajectory of L3 acquisition, then it seems pivotal to uncover the potential variability of its starting point.

To deal with the potential impact of proficiency and dominance on L3 production and perception, L2 proficiency was measured using the LexTALE in English (Izura, Cuetos, & Brysbaert, 2016; Lemhöfer & Broersma, 2016) and Spanish (Izura et al., 2016). The Bilingual Language Profile (BLP) was used to measure language dominance and background (Birdsong, Gertken, & Amengual, 2012). Finally, participant groups will be exposed to either German or French, to examine how phonetic similarity across languages may play a role in the first exposure to an L3.

1.2 Cross-lingusitic language features

In order to gain insights into the relative influence of a first and second language the present dissertation makes use of both (relative) voice-onset time (VOT) in stop production and perception and spectral qualities of vowels, specifically the first and second formants, in vowel perception and production. In the following sections, the vocalic systems of Spanish, English, French, and German, as well as their respective uses of voice-onset time, are overviewed.

1.2.1 Voice-Onset Time

Voice-onset time (VOT) is a phonetic measure in milliseconds of the release of the closure and its relationship with the onset of the vibration of the vocal folds (Lisker & Abramson, 1964). In particular, when the vocal folds begin to vibrate prior to the release, this is described as pre-voicing and is characterized by a negative VOT value. On the other hand, when the release occurs prior to the onset of voicing, a positive VOT is measured. Stops with a positive VOT are considered phonetically voiceless, where a negative VOT is phonetically voiced. VOT has been found to be a primary acoustic cue in the distinction of stop consonants in many of the world's languages. However, the use of VOT to make distinctions between sounds is language specific. For example, Spanish and French belong to a classification of languages referred to as true-voicing languages (Lisker & Abramson, 1964). This label suggests that the difference between a stop consonant that is phonemically voiced, such as /b/, is also phonetically voiced (negative VOT) and its counterpart that is phonemically voiceless, such as /p/, is also phonetically voiceless (positive VOT). True voicing languages stand in contrast to aspirating languages, such as German and English, in that the phonemic categories /p/ and /b/ exist within these languages, but are distinct in the way that they make use of VOT. That is, in aspirating languages, /p/ typically has a long-lag VOT (a longer duration in milliseconds) than the true-voicing languages' /p/. In the latter case, /p/ also has positive VOT, but typically with a much shorter duration. The phoneme /b/, on the other hand, is often realized with a short-lag positive VOT in aspirating languages, in which it likely has acoustic/phonetic overlap with the true-voicing language /p/. In other words, true voicing languages, like Spanish and French, use lead (phonetically voiced) and short-lag (phonetically voiceless) VOT to distinguish stop consonants, where aspirating languages (English and German) use short-lag and long-lag VOT. Overall, these cross-linguistic differences in VOT provide a useful continuous measure that can provide insights into cross-linguistic influence.

1.2.2 Vocalic systems

Another useful language feature to examine cross-linguistic influence is the vowel inventory of each language. Studies examining the articulatory space of vowels typically involve spectral measurements such as formant frequencies. In these studies, the frequency of the at least first and second formants are measured and compared for each vowel. The first formant (F1) is an acoustic correlate for vowel height, with higher values corresponding to lower vowels. The second formant (F2) represents vowel frontness, with higher values corresponding to a more fronted vowel. The chosen languages also vary in the size of their vowel inventories. The Spanish vowel space is the smallest of the present study and consists of 5 distinct vowel monophthong categorizations (Bradlow, 1995). English, French and German have larger relative vowel spaces, where English has 11 monophthongs, and German has 14 (Bradlow, 1995; Jongman, Fourakis, & Sereno, 1989). French has as many as 16 phonemic vowel sounds.

Cross-linguistic differences in vowel spaces are important in part due to evidence that language specific L1 categories seem to matter during L2 learning. That is, previous research has identified patterns in learning novel vowel contrasts, where some are easier than others and this ease is driven by cross-linguistic similarity and vowel inventory size between languages. (Escudero & Vasiliev, 2011; Leussen & Escudero, 2015).

The main question of the present dissertation is how cross-linguistic acoustic similarity and language status effects interact in L3 acquisition. In other words, it will be examined whether L1 and L2 phonological systems have equal impact on the L3 in perception and production in that they are equally affected by cross-linguistic acoustic similarity, regardless of the most similar acoustic sound to an L3 sound is an L1 or L2 sound. Alternatively, it is possible that a privileged status of either the L1 or L2 could cause perception and production errors, in which L3 learners assimilate L3 sounds to a sound that is not the best acoustic match.

Chapter 2: Literature Review

2.1 Models of third language acquisition

Research in third language acquisition has attempted to model the interplay between L1 and L2 language systems and their cumulative influence in the process of 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 have largely focused on morphosyntax.

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 the founding of the CEM 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, and, importantly, are helpful. Flynn et al. (2004) 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 (2012) 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, contrary to Flynn et al. (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 that it can not explain non-facilitative influence from previous languages in the L3 (Rothman, 2011). 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 cognitive 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 than 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 an early 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 Bardel & Falk (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 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 further 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.

This result was replicated in other studies. For instance, Bohnacker (2006) found that L1 Swedish-L2 English-L3 German speakers showed evidence of English like word order in L3 German V2 constructions, rather than facilitative L1 Swedish, Additionally, Falk & Bardel (2011) found that 44 L3 German learners with L1 English-L2 French or L1 French-L2 English behaved differently in tasks involving object pronoun placement. This distinct behavior was likened to L2 influence and suggested that order of acquisition was more deterministic in predicting CLI than typology. 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 L3 learners of a V2 language who spoke a V2 L1 and non-V2 L2, Falk, Lindqvist, & Bardel (2015), 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.

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 (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 perception and production. The present dissertation refers to the "absolute initial state" as "first exposure" to avoid potential misunderstandings regarding the amount of exposure that L3 learners have prior to the experimental tasks done in the present dissertation. 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 relative to one another. In a future study 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. The study examined adjective placement and meaning in the L3. The first group consisted of L1 Italian-L2 English speakers learning L3 Spanish, while the second group spoke L1 English and L2 Spanish and was learning L3 Brazilian Portuguese (BP). 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. These results suggested that both groups performed in a target-like manner on the experimental task and also supported the Cumulative Enhancement Model.

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 et al. (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 did not find a reliable difference between experimental conditions. 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 predictive 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, and it is unclear how the authors of the systematic review determined when to depart from the conclusions of authors of studies and when they decided not to.

Limitations of the TPM in addition to the lack of robust and cross-linguistic support 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 et al. (2020) describes this 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 over-interpretation can be seen in the use of single language properties and inappropriate statistical tests to justify narrative 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, some of the evidence gathered to date provides counter-evidence 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 can rule out the influence of both languages within the property examined, since small samples which provide inconclusive results (non-significant 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. This is due to the idea that, on one hand, it is not possible to determine whether group results have to do with sampling issues, and, on the other, clear criteria for co-activation of both languages known by a bilingual was not explicitly included in predictions and design of these studies. As a result, in the case of the L3 BP studies, it cannot be determined based on the results of studies to date that 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

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 (Westergaard, 2021; Westergaard, Mitrofanova, Mykhaylyk, & Rodina, 2017). In the founding article of the LPM, Westergaard et al. (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 the co-occurrence 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 the general perceived 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. In order to test whether bilinguals have access to two languages during L3 learning, Mitrofanova & Westergaard (2018) conducted a study on Norwegian-Russian bilinguals and Norwegian monolinguals in which they taught them some of an artificial language which was designed to contain Norwegian-like lexical items, but also Russian like case marking to distinguish subjects from objects. The results revealed that the bilingual group performed better than the monolingual group on a sentence/picture matching task, in which the monolinguals relied solely on word order when making decisions. These results suggested that the bilingual group was able to make use of both of their languages at first exposure. In a series of studies, Stadt, Hulk, & Sleeman (2016) examined L1 Dutch - L2 English - L3 French speakers in both English immersion and general Dutch-dominant educational settings. In the first study, the results of a Grammaticality Judgment Task indicated that third year the immersion group showed signs of heavy L2 influence in the L3, where the regular group showed evidence of CLI from both the L1 and the L2.

In a follow up study, the authors recruited 4th year students from the same educational background to investigate whether L2 influence would differ between 3rd and 4th year students (Stadt, Hulk, & Sleeman, 2018b). The results revealed that both 3rd and 4th year immersion groups behaved similarly and showed evidence of primarily English influence on their L3 judgments, though evidence for L1 influence was also found. In a second follow up study, the authors gave similar materials to the an L3 German group (Stadt, Hulk, & Sleeman, 2020). In this case, L3 German judgments were concluded to have been influenced by the more typologically similar L1 Dutch, rather than L2 English as found in the previous cases. Taken together, the results of these studies suggest that both typology and L2 Status may play a role in modeling the patterns of cross linguistic influence in an L3. Finally, these authors also conducted longitudinal study (Stadt, Hulk, & Sleeman, 2018a), again with L1 Dutch-L2 English-L3 French speakers. In this study, they found that the L1 was more influential in the beginning stages of L3 development, but that the L1 also maintained influence as L3 proficiency increased. In a more recent study, 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-facilitative influence in others. Additionally, the sources of this influence could be likened to either Russian or Norwegian. The results of studies, taken together, suggest that both languages are active in L3 learning, and it also appears that they remain active throughout L3 development.

Like other L3 models, the LPM has received criticism. Scholars have argued that the LPM makes vague predictions that create a problem in modeling L3 transfer acquisition (Bardel & Falk, 2021; Wrembel, 2021), 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 & Falk, 2021). However, Westergaard (2021) 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.

Additionally, it is notable that the empirical studies that serve as the basis of the LPM (Westergaard, 2021; Westergaard et al., 2017) utilize groups of simultaneous bilinguals, which do not allow for the examination of potential L2 status effects. Following this idea, Westergaard et al. (2017) suggest that an optimal design to examine both the individual 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.

2.1.5 The Scalpel Model

The Scalpel Model (Slabakova, 2017) is an additional model of L3A which overlaps in many of its predictions with the LPM. Like the LPM, the Scalpel Model rejects the idea of wholesale transfer of either L1 or L2 proposed by the TPM. Rather, it suggests that, given L3 input, the L1 and L2 combined grammar should successfully extract facilitative features from the input. In this view, non-facilitation is driven by misleading, processing complexity, and frequency of a construction. That is, more complex constructions and less frequent constructions are predicted to be associated with non-facilitative CLI.

The major divergence of the Scalpel Model from other L3 modes lies in its focus on the role of input in driving non-facilitation. Evidence for this claim comes from 2 empirical studies. Both studies examined three groups of participants L1 Basque-L2 Spanish-L3 English, L1 Spanish-L2 Basque-L3 English, and L1 Spanish-L2 English (García Mayo & Slabakova, 2015; Slabakova & García Mayo, 2015). García Mayo & Slabakova (2015) tested the use of null objects, which are not allowed in English, they are in Basque and sometimes in Spanish. Using speaker judgments, the results showed that null objects were often correctly rejected in L3 English. The second study, examining the same groups, focused on topicalization (Slabakova & García Mayo, 2015). In this case, L3 English speakers did not show evidence of target like behavior on grammaticality judgment tasks. The results of these studies suggest that one feature, topicalization, is difficult to acquire, where null objects was easier to acquire in L3 English. The author suggests that this difficulty could be explained the likely lower frequency of topicalization in L3 English input relative to the lack of null objects. Overall, the predictions of the Scalpel Model at first exposure might suggest that facilitation would occur in the event that the L3 stimulus has a clear bias to the L1 or the L2. That is, the Scalpel Model and the CEM should share similar predictions for the behavior of L3 learners at first exposure, but differ as the L3 develops. The Scalpel Model does not seem to predict or concur with the prediction of the LPM that L3 performance on experimental task should fall between L1 and L2 performance.

2.2 Models of L2 phonological acquisition

Notably, much of the evidence for the L3 models to date come from studies which examine the acquisition of (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 (Rothman, 2010, 2011, 2013, 2015). 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. Though a model of L3 phonological acquisition has yet to be proposed, the expansion of L2 models of phonological acquisition has been proposed (Wrembel, Marecka, & Kopečková, 2019).

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 these models apply to L3 speech development when L2 categories are also theoretically available 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.

2.2.1 The Speech Learning Model

The original SLM (J. 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 (J. Flege, 1995). In this view, acquisition of L2 segments occurs after increased exposure to these segments, where segmental complexity modulates the rate of learning (James Emil Flege, Aoyama, & Bohn, 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, & Bannister (2005) conducted a study in which they successfully 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 & Flege (1992) 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 as long lag /t/ as /t/, despite the lack of a long-lag category in their native inventory.

2.2.1.1 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 bilinguals, where speech learning begins when the phonetic categories of the L1 have been established. According to James Emil Flege et al. (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 for speech learning, and suggests that non-native production and 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.

2.2.2 The Perceptual Assimilation Model

Relative to the SLM, the Perceptual Assimilation Model (SLM) focuses on the perception of sound contrasts by L2 learners (Best & Tyler, 2007). 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 exists 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 two non-native sounds are equally good exemplars on an L1 category (uncategorized 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 closet 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, Escudero & Chládková (2010) found that Spanish L1 speakers assimilated SSBE /ae/ and /a/ to Spanish /a/ (that discrimination of this contrast was difficult). Additionally, Escudero, Sisinni, & Grimaldi (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 assimilated the two categories in Dutch /a/ and /a:/ to their single Spanish category /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 similarly 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.

2.2.3 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 (Leussen & Escudero, 2015). 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 empirical studies. Namely, that these studies have provided evidence that L2 learners experience difficulty with the contrasts of /r/ and /l/ in Japanese, (Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004) “beat” and “bit” in Spanish and Portuguese (J. E. Flege, Frieda, & Nozawa, 1997) “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 based on Stochastic Optimality Theory (Boersma, 1998).

An important 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 assimilation 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 relatively 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 & Boersma (2004), 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 discriminate outside of lexical items (Curtin, Goad, & Pater, 1998). Other studies show that some lexical items can also be reliably distinguished by L2 learners that could not be told apart pre-lexically (Cutler, Weber, & Otake, 2006; Weber & Cutler, 2004). As a result, an important tenet of the revised L2LP (Leussen & Escudero, 2015) 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, Leussen & Escudero (2015) suggest that 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. 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.)

2.3 Previous literature in L3 phonology

In the research to date, L3 models have focused largely on empirical results stemming from experiments in morpho-syntax to derive predictions about how the L1 and L2 affect L3 acquisition. Fewer studies have been conducted in L3 phonology, and it remains unclear whether L3 models make specific predictions for phonology and morphosyntax. Despite this lack of specificity in the models when it comes to phonology, some trends have emerged in the body of research. Specifically, it appears that the L1 and L2 likely both affect L3 perception and production, with the L2 often providing a greater level of influence in the beginning stages of acquisition, even when it is not facilitative. This trend cannot readily be accounted for by the TPM, the L2SF, the CEM nor the Scalpel Model, leaving only the LPM which can predict co-activation of all languages known by a bilingual in L3 acquisition. The following sections cover the empirical studies that have been done in L3 production studies across L3 proficiency levels, and the fewer studies carried out in L3 perception, as well as their relevance to L3 models.

2.3.1 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_language_1998. 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 to 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 was rated as having 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, but also provided evidence that this effect diminishes as L3 proficiency increases. 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 L3 phonology 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 (Wrembel, 2010). Similar findings have also been reported in vowel production (Kamiyama, 2007) and vowel reduction and speech rhythm (Gut, 2010). L2 influence has also been found in VOT productions. Llama et al. (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.

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 reported by Wunder (2010) in L3 Spanish speakers, and by Blank & Zimmer (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, 2012), or in advanced L3 learners (Llama & Cardoso, 2018). Importantly, these studies report L3 VOT values which fall between L1 and L2 values. Following the suspicion that intermediate values might have to do with either sampling issues or proficiency effects, Parrish (2021) examined Mexican Spanish-English bilinguals who produced voiceless stop-initial 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, and that suggested that intermediate values were less likely to have been seen in previous studies as a result of small samples or proficiency effects. However, a subsequent analysis of the data suggested that wide individual variation existed, in which some participants produced L3 French as L1 Spanish like, and other produced intermediate, L2-like values. This result suggests that higher samples could reveal group trends and provide better insights into individual variation in crosslinguistic influence, as opposed to assuming that a single group trend exists.

2.3.2 L3 Perception studies

Overall, studies in L3 perception have been much more scarce relative to studies in production. These few studies have often used methods found in L2 speech acquisition research, such as studies testing the predictions of the Perceptual Assimilation Model. For instance, Wrembel et al. (2019) examined the categorization and discrimination of L3 vowels by 10 young trilinguals who spoke L1 German-L2 English-L3 Polish. To test categorization, a cross-linguistic similarity task was used in which participants heard minimal pairs of sounds and had to rate how similar sounds were on a 1-7 Likert scale. The results showed evidence that participants assimilated L3 sounds to both L1 and L2 categories, but preferred the L2. In a second experiment, and AX discrimination task was given to participants to evaluate whether retroflex and palato-alveolar silibant discrimination, a feature of Polish, could be accessed in L3 words. The results revealed that discrimination of the L1 Polish contrast was very good (84% accuracy), suggesting that L3 learners retain access to L1 sound contrasts in L3 words. Additionally, this language specific phonetic discrimination was attended to by even L3 beginners. Balas (2018) also used the PAM as a perceptual framework to work in and adapted it to L3 learners. The study recruited three groups of Polish L1 speakers, including two L3 groups (L1 Polish-L2 English-L3 Dutch and L1 Polish-L2 English-L3 Dutch). The third group spoke only English as an L2. All three groups were listened to Dutch vowels and were asked to categorize them given Polish vowel categories. Unfortunately, the L3 groups were not given L2 English categories as options during this task, so the results of this study cannot directly provide evidence that L3 learners categorize L3 sounds using both the L1 and L2. The same study also conducted an AXB discrimination task of 8 Dutch vowels and found that discrimination was at ceiling for all vowels involved.

An additional line of research in L3 perception studies has pondered whether L3 learners have a general perceptual advantage over their L2 learning counterparts. In an early study, Werker (1986) investigated mono, bi and multilingual participants' discrimination of non-native perceptual contrasts and did not find any evidence that the bi or multilingual group had any perceptual advantage. Patihis, Oh, & Mogilner (2015) examined Korean stops by naive mono and multilingual listeners and found that phoneme discrimination was feature dependent, rather than bilingualism providing a general advantage. Antoniou, Best, & Tyler (2013) investigated the contrast of Ma'di stops by early English-Greek bilinguals, English monolinguals and Greek monolinguals. The researchers wanted to know whether language model affected categorization and discrimination. They found that categorization is modulated by language mode, but discrimination of novel sounds was not. In particular, the Greek monolingual group was most successful in the discrimination of the novel contrast, where the English monolingual group was the least accurate. The bilingual group, regardless of language mode, displayed intermediate ability to discriminate the novel sounds. In a further study, Antoniou, Liang, Ettlinger, & Wong (2015) taught mono and bilingual participants an artificial language based on either English or Mandarin. The results revealed that bilingual participants outlearned the monolingual groups overall, but universally more difficult unfamiliar L3 segments do no seem to be learned more easily by bilinguals. Wrembel, Gut, Kopečková, & Balas (2020) also found that cross-linguistic influence is structure dependent and varies among individuals. Onishi (2016) reported a bilingual advantage, but, rather than feature dependent, suggested that this advantage was global. The study involved speakers of Korean who spoke L2 English and L3 Japanese. The author found a correlation between successful discrimination of L3 contrasts and L2 experience, and argued that this correlation is evidence for a bilingual advantage.

These findings suggest that crosslinguistic similarity of both a bilingual's languages play a role in the ease of acquiring a novel, L3 sound. Additionally, research focused on a bilingual advantage in phonetic learning can inform L3 models, since it implies that bilinguals have access to the phonetic categories in both languages that they speak, even at first exposure, whether or not the advantages of bilingualism on phonetic learning are global or feature dependent. Few studies have investigated the role of VOT in L3 perception. On such study is Z. Liu, Gorba, & Cebrian (2019), which examined the perceptual boundary of a VOT continuum in L1 Mandarin, L2 English and L3 Spanish. Using a /pi-bi/ continuum, the authors found that the perceptual boundaries of each language were 28ms for Chinese, 24.6ms for English and 23ms for Spanish. Despite the goal of the study being to examine regressive transfer and to compare the L3 speakers to monolinguals, descriptive evidence of L2 effects can be inferred from the reported means. Additionally, J. Liu & Lin (2021) examined perception and production of word-initial voiced and voiceless stops in speakers of L1 Mandarin, L2 English and L3 Russian or Japanese. The authors found that voiced stops were effectively perceived, but not reliably produced, where voiceless stops were both successfully perceived and produced.

2.4 Methods and anylsis in previous work

The widely varied findings in previous work do not have a clear plausible correlate, 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. Finally, the manner in which the present dissertation addresses these issues will be covered.

Overall, it is unclear whether sufficient sample sizes have been used in the body of research to date. Of the empirical studies reported in the present dissertation, none used a power analysis or otherwise justified their sample size per group. 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 (Brysbaert, 2020). 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 Brysbaert (2020) argues, low samples lead to low statistical power, and in turn provide a metaphorically 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 interpretation of the results of various inferential statistical methods, such as the lack of a main effect in an 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, 2017). 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 two groups of L3 learners 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. Firstly, the lack of a power analysis does not rule out sampling error. It cannot be argued that the sample used in this study was high enough to reliably detect an effect (or lacktherof). Secondly, a wide confidence interval from the low sample size would make providing evidence for statistical equivalence in a test of equivalence likely impossible. Unfortunately, the reporting of results in many L3 studies to date do not allow for post-hoc power analyses to be run, since means and standard deviations are seldom reported. It remains unclear whether studies in L3 research are sufficiently powered to be able to generalize, and, by extension, to build models.

Despite the lack of empirical data available to evaluate the statistical power of L3 studies to date, it is probable that the sample sizes used are not statistically powered. Brysbaert (2020) argues that, for between group comparisons, samples of at least 80 per group are often necessary, but a power analysis should be carried out to justify sample size. To the present author's knowledge, no L3 study has been able to recruit this many participants per group. For instance, the seminal study motivating the TPM (Rothman, 2011) recruited just 11 and 15 in two groups. 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 significant statistical difference. Plonsky (2015), 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 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.

2.4.1 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, and 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 reasonable to suppose 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.

Following this logic, the question becomes whether L1 or L2 categories, or a combination of both, are retuned to L3 categories. Additionally, it is unclear what conditions determine which language category, whether L1 or L2, is initially chosen and its rate of retuning to become more target (L3) like as L3 input exposure increases. 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 language categories. 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 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 3: The Perception of the new language vowels by Spanish-English bilinguals

3.1 Introduction

The body of research on phonetic and phonological learning has uncovered a crucial role for the proximity of the native language of a given speaker and the language they are learning. For example, the empirical basis for two major models of L2 phonology, the Speech Learning Model, and the Perceptual Assimilation Model, posit key roles for the native phonological structure of a speaker when predicting the ease or difficulty of acquiring specific L2 sounds. For example, it has been shown that speakers of languages with a long-short lag VOT distinction in stop consonants have difficulty acquiring native-like VOT values in their L2 if it is a true voicing language. Additionally, it is difficult for speakers to "break up" a category into two smaller categories, such as Spanish speakers who have difficulty acquiring the /i/-/ɪ/ contrast (James Emil Flege, Bohn, & Jang, 1997). Both of these difficulties arise, presumably, due to the native language structure of the learner.

These models do not address whether this native phonological structure has a special status relative to later acquired phonemes. Specifically, it is not clear whether both the native and L2 phonological systems may both be possible sources of influence during new language learning. The present chapter tests whether this is the case by carrying out two experiments. In the first experiment, German and French vowels are categorized given both English and Spanish categories. The vowel categorization task is designed to determine whether bilingual speakers will account for new language sounds using both English and Spanish. In experiment 2, the discrimination of English vowel contrasts in L3 words are tested. This experiment avoids categorization and measures discrimination of sounds with the aim of showing whether phonetic discrimination is possible in a new language when the sound contrast exists in either the L1 or the L2. If speakers of categorize and discriminate new language sounds using both Spanish and English phonology, this would suggest that it is not simply native language phonology which predicts new language learning, but that all previously known languages may play a role in the learning of a new language sound system.

3.1.1 Perception/perceptual assimilation in L2 and L3 speech

Previous research has found that the ease of learning sounds in a new language depends upon your language background and experience. In particular, it has been found that the phonemic and phonetic qualities of ones native language can predict the ease or difficulty of sound contrast perception in both accented L1 and L2 speech. For example, when they were asked to categorize vowel sounds of Standard Southern British English (SSBE), speakers of Salento Italian categorized three SSBE vowel sounds as their native /a/ (/æ/, /ɑ/ and /ʌ/), and also did not discriminate these three sounds efficiently (Escudero et al., 2014). This finding suggests that experience with Italian, a 5-vowel system, makes the discrimination of languages with larger vowel inventories more difficult than it would be for a speaker of a language with a similarly large vowel system. Other language pairings also show evidence for this phenomenon, such as Spanish listeners categorized Dutch /a/ and /a:/ as their Spanish /a/ (Escudero & Vasiliev, 2011), and Spanish speakers categorizing Canadian English /æ/ and /ɛ/ to Spanish /a/ (Escudero & Vasiliev, 2011), where speakers of languages with a smaller vowel inventory would categorize more than one vowel sound in another language as one in their native language. In other cases, speakers have matched two novel-language vowels to two of their native categories. Escudero & Chládková (2010) found that Spanish speakers assimilated English vowel sounds to their Spanish /a/ and /e/, although they assimilated British /æ/ to their Spanish /a/ and English /æ/ to their Spanish /e/, while matching /a/ in both dialects to their Spanish /a/. Taken together, these studies suggest that native language categories impact the categorization of sounds in a new language.

Best and Tyler formalized the idea that the ease of the perception of any new language is driven by the phonemic and phonetic differences between the inventories of the speaker's native language and an unknown language. Originally, proposed to account for non-native (foreign accented) speech perception, the Perceptual Assimilation Model (*PAM; Best & Tyler (2007)) has also been extended to L2 learning (PAM-L2). The model predicts that when languages (the one known and the one being learned) are more similar in terms of how sounds contrast, then they will be easier to effectively perceive. For instance, in what the PAM calls Two-Category Assimilation, two L1 categories would be hypothetically matched to two L2 categories, and discrimination of that L2 contrast would be predicted to be very good. On the other hand, in the event that one L1 category corresponds to two L2 categories, discrimination is predicted to be poor (Single-category Assimilation).

Despite the general empirical support for the model, the PAM has not explicitly discussed whether bilingual language learners approach learning new-language sounds differently than monolinguals. In their 2007 chapter, Best and Tyler state that "the relative ease or difficulty of a given contrast varies according to the listener's native language" (p. 16), but make no mention of bilingual listeners. It is not clear whether, if given the option, speakers could be guided by both their native and subsequently learned languages when asked to categorize new language sounds. Very little research exists testing whether this is the case. Wrembel et al. (2019) examined the categorization of L3 German vowels by 10 multilingual L1 Polish, L2 English, L3 German speakers, given L1 and L2 categories from which they could choose. They found that the speakers assimilated L3 sounds to both their L1 and L2, with a preference for L2 sounds, suggesting that subsequently learned languages might guide subsequent language perception in addition to the native language. Onishi (2016) examined the impact of L2 experience on L3 contrast perception ability in speakers of L1 Korean, L2 English and L3 Japanese and found that better experience with specific L2 contrasts correlated to better perception of those same contrasts in the L3 contrasts, but only in some cases. These studies provide evidence that speakers may draw upon more than simply their native language when accounting for new language sounds. However, since the Wrembel et al. (2019) and Onishi (2016) both examined the perceptual assimilation of L3 sounds to L1 and L2 sounds, it is more difficult to disentangle whether their assimilation choices might be a result of L3 acquisition or previous language influence. For example, the target like contrast performance by the L3 Japanese speakers, while correlated to L2 performance contrast, does not provide evidence that these speakers began their development with an advantage due to their L2 knowledge.

The present chapter adds to the scarce work done in L3 perception in at least two meaningful ways. First, naive bilingual speakers were recruited, rather than actual multilingual speakers. This approach offers valuable information about the absolute starting point in L3 phonological acquisition, while controlling for (in the L3) the wide variation that one would expect in variables that have been shown to impact language learning outcomes, such as language proficiency, use, and learning context. That is, if these speakers assimilate new language sounds to both their L1 and L2 it would provide evidence that multilinguals not only seem to have access to their L1 and L2 later in development, but that they also begin with this access. This distinction is important in the context of L3 models, which debate whether speakers maintain access to both their known languages throughout development.

3.1.2 Overview of experiments

The goal of the experiments is to examine how participants categorize and discriminate vowel sounds in a new language. For this reason, two experiments were created. The first experiment, designed to measure categorization, was a vowel categorization task. During the vowel categorization task, participants heard a single new-language vowel sound and their task was to match the played sound to the most similar vowel sound. The methodology resembles studies carried out in L2 phonology to test the predictions of the PAM-L2. Differently from the L2 studies, the present study included both L1 and L2 categories and tested bilinguals rather than functional monolinguals.

The second experiment was an AX discrimination task. It examined how all groups discriminated English vowel contrasts when they are in new language words. Specifically, a seven step vowel quality continuum ranging from /i/ to /ɪ/ was created.

During a trial, participants heard two steps on the continuum back to back and had to determine whether the sounds they heard were the same or different. This task allowed for the examination of perception in a purely phonetic, rather than phonemic, context, since categorization was not explicitly involved in this task.

3.2 Experiment 1 - L3 Perceptual Assimilation Task

Experiment one examined the initial state of categorization of the vowels of two unknown languages, French and German, by Spanish-English bilinguals, Spanish monolinguals and English monolinguals. Each participant completed a background questionnaire and a perceptual assimilation task.

3.2.1 Participants

A total of 199 participants took part in the present study, after the removal of the data of 10 total participants who self-reported being bilingual yet not feeling comfortable speaking or understanding their second language. The present sample includes L1 Mexican Spanish - L2 English speakers (n = 59) and L1 English - L2 Spanish speakers (n = 55). The monolingual participants included 29 Mexican Spanish monolinguals and 56 American English speakers. All participants were recruited on prolific and were pre-screened according to criteria detailed below. In addition to filters in place from Prolific.co, the participants were screened further using an adapted version of the Bilingual Language Profile (Birdsong et al., 2012). All participants who answered 'no' to the question "Do you speak a language other than English and Spanish" were permitted to continue the experiment.

3.2.1.1 Bilinguals

Screening data and experiment-initial questionnaires were used to find bilingual participants who began learning their L2 later in life and reported not having learned a language aside from English and Spanish. The English L1 group came from all over the United States while Spanish L1 group came from Mexico. Each groups mean age, L2 use, self-reported oral and perceptive proficiency are seen in figure 3.1. As can be seen from the figures, The English L1-Spanish L2 group began L2 learning later on average, while they also felt comfortable in their L2 at a later age than the Spanish L1-English L2 group. The participants also rated their L2 proficiency. They were given a 0-6 Likert-type scale in which they answered the questions "How well do you speak [their L2]?" and "How well do you understand [their L2]?". "0" corresponded to "not very well at all", where "6" corresponded to "very well" (Figure 3.2)

3.2.2 Materials

3.2.2.1 Target phrases/conditions

The participants heard and categorized a total of 4 vowel conditions per language given 7 total carrier words to choose from. The 4 vowel conditions included the /i/, /y/, /o/, and /ʌ/. The four sounds were embedded in both a fricative /fVf/ and bilabial /pVp/ or /pVf/ frames and played a total of 5 times each. Thus, each participant categorized 40 tokens per language (5 repetitions x 2 frames x 4 vowel conditions), given 7 carrier words from which they chose. The 7 word choices included 3 English carrier word choices intended to represent the phonemes /i/, schwa/wedge, /u/ and /a/ (*feel*, *fun*, *fool* and *thought*). The remaining 3 choices were Spanish carrier words intended to represent the phonemes /i/, /u/ and /o/ (*fin*, *su*, *son*). The screen that the participants saw 3.3 shows an instance of the selection trail.

The vowel sounds included in both experiments were intended to bring about four distinct cross-linguistic situations. First, the L3 phoneme /i/ was included to create a conflict in which both source languages, Spanish and English, have a similar sound /i/. The phoneme /i/ was given in the Spanish word *fin* and the English word *feel.* Next, the L3 phoneme, referred to as either the wedge or (schwa depending upon lexical stress) was given in an attempt to bias the selection of English. This condition was intended to be assimilated to the English choice *fun.* Third, the phoneme /o/ was included to bias the same Spanish category, where a rounded /o/ does not exist in American English. The intended choice in this case was the Spanish word *son*, but

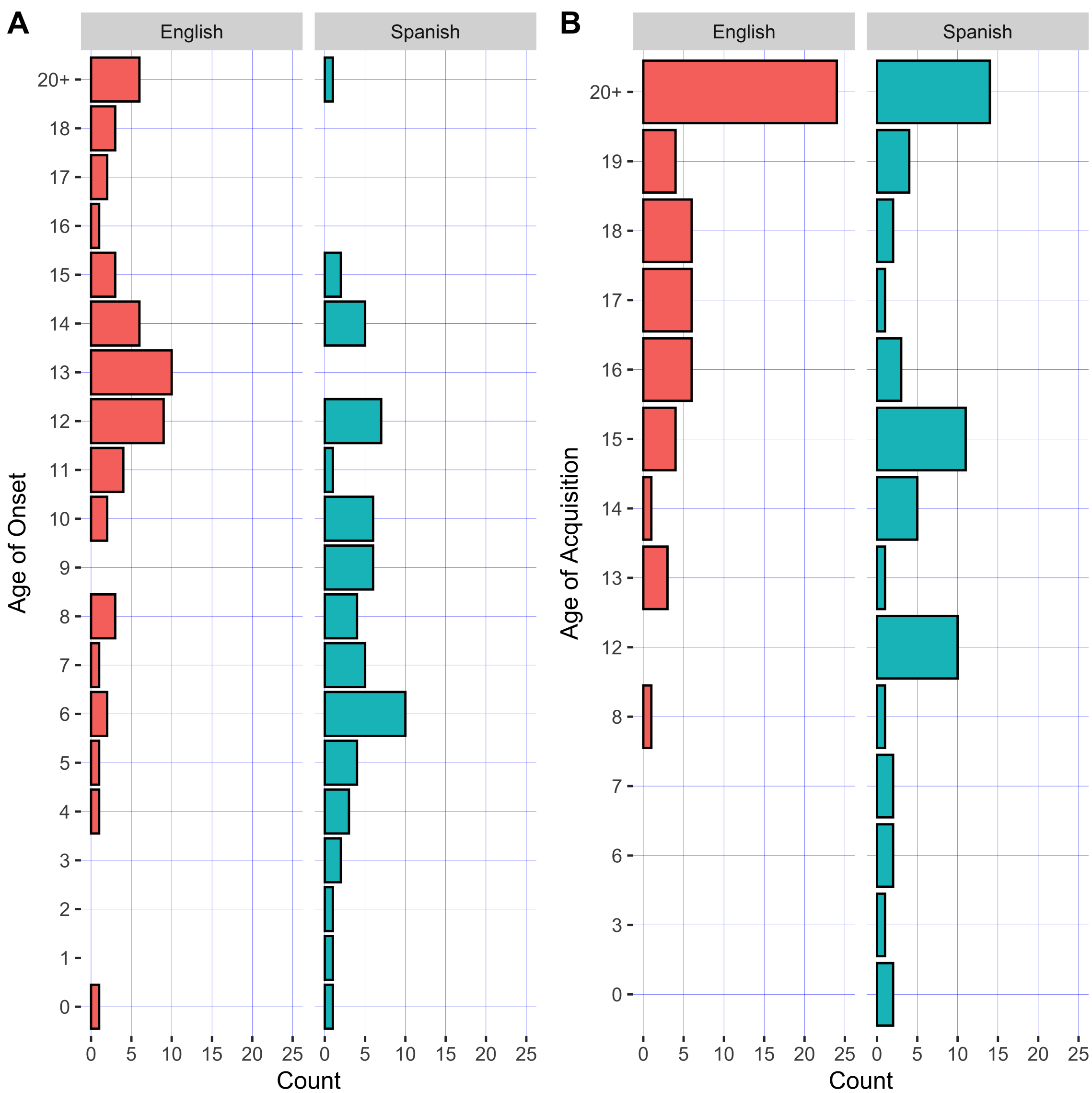


Figure 3.1: Age of Onset and Age of Acquisition in each bilingual group

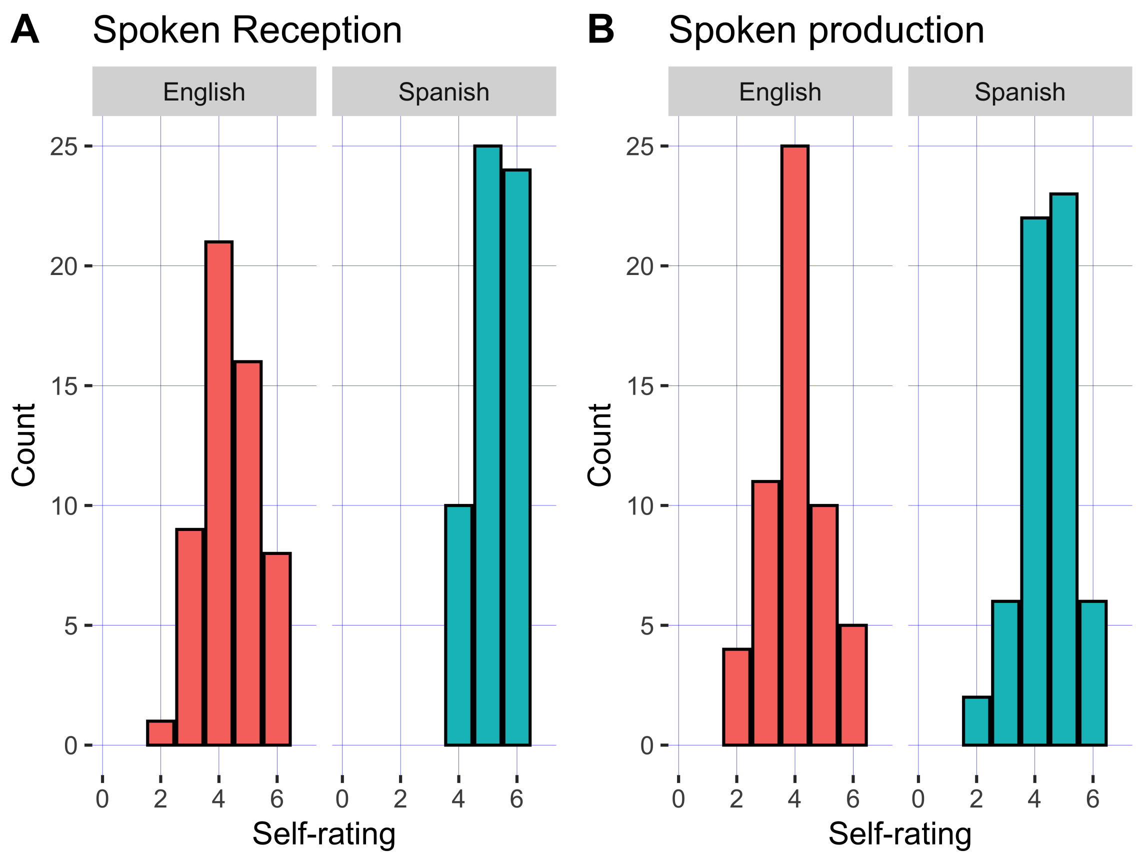


Figure 3.2: Self-rated 1-6 proficiency in production and perception

the English word *thought* was also provided as an alternative. Finally, the phoneme /y/ was added to explore how a sound that is not present in either language will be categorized. Given that /y/ is a high-front vowel, it is possible that it could be assimilated to other high vowels, either /i/ as in *feel* or *fin*, or /u/, as provided in *fool* or *su*. Additionally, after making each selection, the participants then rated their pick for goodness of fit by clicking a 1-5 continuous Likert scale (3.4.

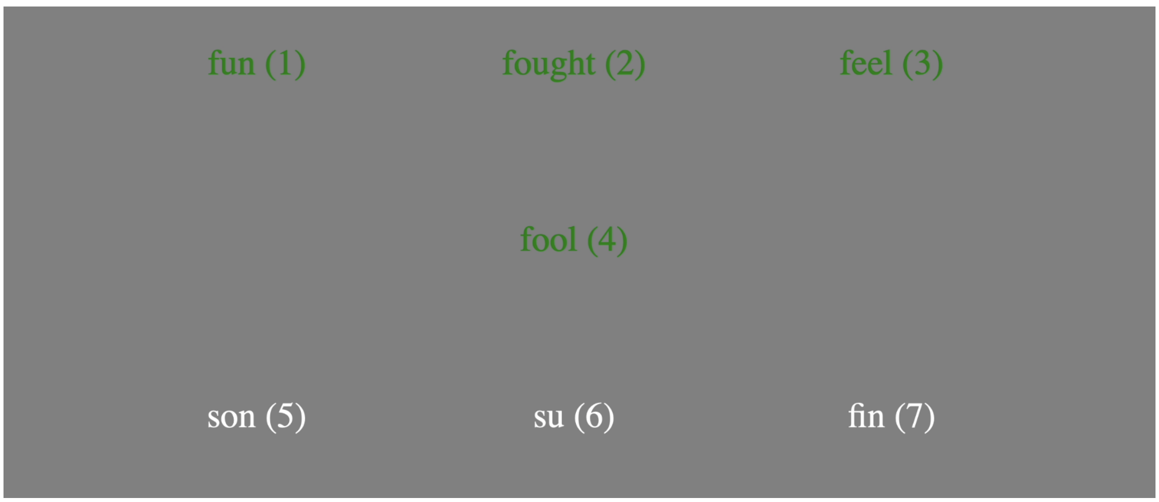


Figure 3.3: Example screen of the vowel categorization task

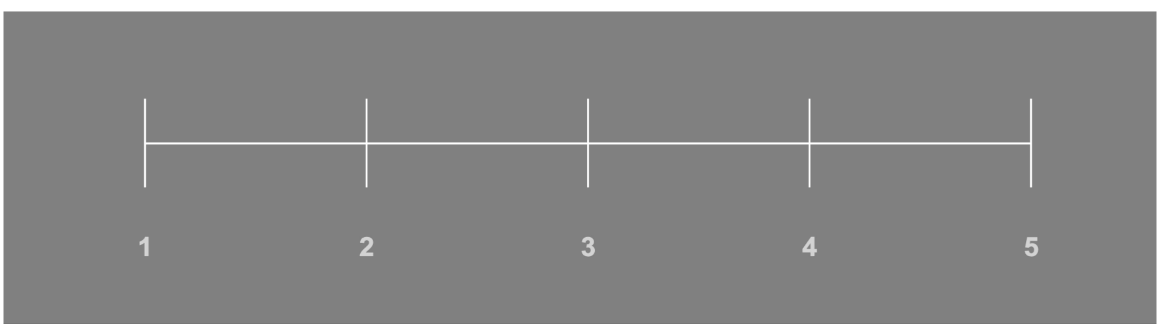


Figure 3.4: Example of the likert style rating after each selection

3.2.2.2 Stimuli

The stimuli were recorded by adult, female L1 speakers of French and German respectively and was also collected online. The speakers were given each vowel in a word or non-word in both a fricative and bilabial frame. In the event a non-word was provided, a real word containing that vowel sound was included to aid the informant in producing the intended pronunciation of the vowel. Once the stimuli were recorded, one of the two tokens provided by the speaker for each vowel was selected and re-synthesized adding the appropriate onset and coda. In total, 8 stimuli were created per language. Figure 3.5 shows the formant values of the included stimuli in German and French in comparison to similar sounds in English and Spanish. For the purpose of this Figure, an adult female speaker of Madrid Spanish and an adult female American English speaker provided vowel tokens of the answer choices in the present study by producing the carrier words while being recorded in PRAAT (*son*, *su* and *fin* in Spanish and *fought*, *feel*, *fool*, and *fun* in English).

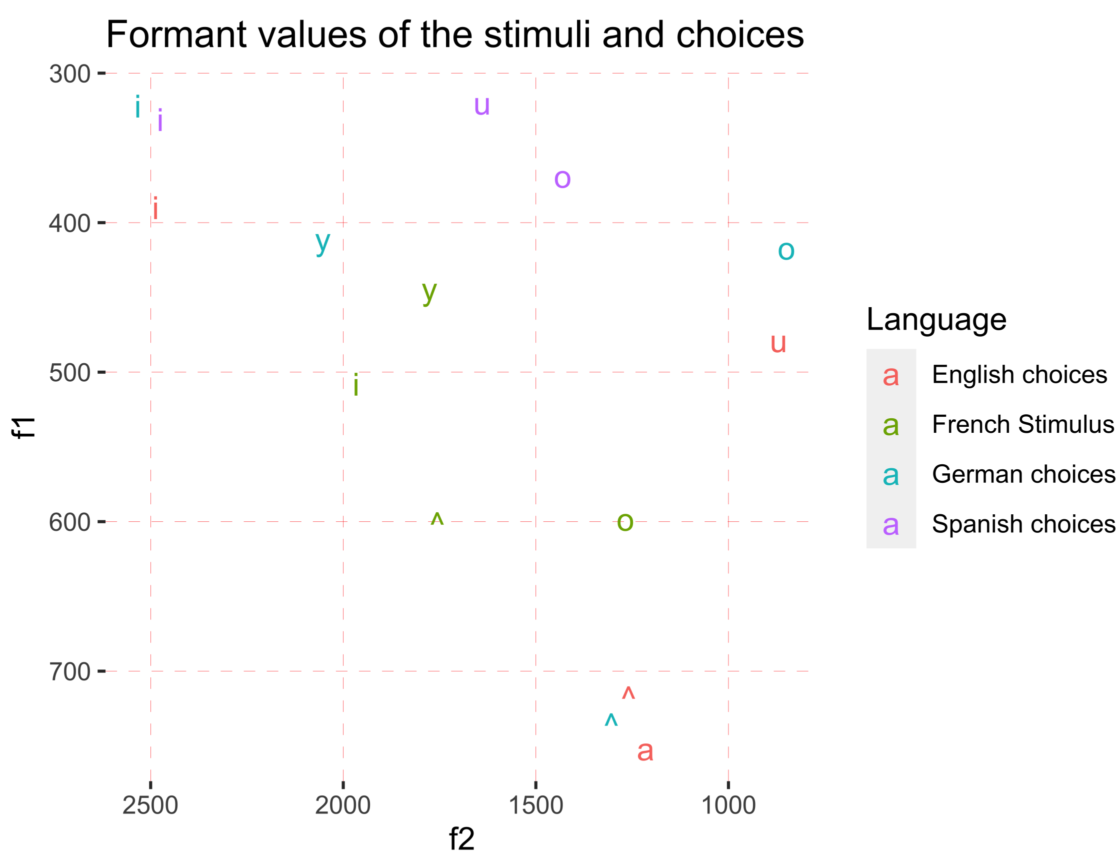


Figure 3.5: Formant values of the model speakers

Table 3.1: Written stimuli used to elicit the auditory stimuli from the native German and French speakers.

Written stimulus	Language	Intended Phoneme
Pief, Fief	German	/i/
Pof, Fof	German	/o/
Püf, Füf	German	/y/
Puff, Fuff	German	/ʌ/
Pif, Fif	French	/i/
Pof, Fof	French	/o/
Puf, Fuf	French	/y/
Puff, Fuff	French	/ʌ/

3.2.3 Procedure

All participants first completed the adapted Bilingual Language Profile (Birdsong et al., 2012) online. An English and Spanish version of the questionnaire was adapted and given to the participants based on their L1. All participants who answered "no" to the question "Besides English and Spanish, do you speak a third language?" were invited to take part in the experimental task. The vowel categorization task was given to participants online. An English and Spanish version of this task was also created, in which all instructions were given in either English or Spanish. During the task, all participants heard French first, followed by a brief pause, and then heard German sounds. The order of the stimuli was counterbalanced and the two tasks were given in a single session with a brief pause between them. The experiments were programmed in Psychopy (Peirce et al., 2019) and made available online via Pavlovia.

3.2.4 Statistical Analysis

For the vowel categorization tasks, the data were analyzed using a series of Bayesian multilevel multinomial logistic regression model in R @[R-base]. The models were fit using the R package `brms` (Bürkner, 2017). A model was run for each of 4 groups: L1 Spanish, L1 English, monolingual English and monolingual Spanish. In each model, the outcome variable was word choice. In the bilingual groups, this consisted of 7 total options (3 Spanish words: *fin*, *su*, *son* and 4 English words: *fun*, *fought*, *feel*, and *fool*.) Thus, outcome of the bilingual models estimates the log odds of choosing one of the seven choices, and would sum to 1 when converted to probability. The fixed effect predictors of the bilingual models were phoneme (/i/, /ʌ/, /y/ and /o/), stimulus language (French or German) and Lextale score (continuous and transformed to a z-score) and all higher order interactions. Random effects included a random intercept for participant to take into account the nested structure of the data.

The monolingual models modeled word choice as a function of phoneme and stimulus language, again with a random intercept for participant to take into account the nested structure of the data. In this case, language choice was more limited in each group limited, with the Spanish monolingual group only having 3 options: *fin*, *su*, *son*, while the English group had 4 word choices: *fun*, *fought*, *feel*, and *fool*. The model included regularizing, weakly informative priors (Gelman, Simpson, & Betancourt, 2017), which were normally distributed and centered at 0 with a standard deviation of 8 for all population-level parameters. The region of practical equivalence (ROPE) was set to 0.18, as the outcome variable was in log-odds (see Kruschke (2018)). All models were fit with 2000 iterations (1000 warm-up). Markov-chain Monte-Carlo sampling was carried out with 6 chains distributed between 8 processing cores.

3.2.5 Results of the L3 Perceptual Assimilation Task

Table 3.2 shows the overall percentage of each word choice (out of 4 possible in English), given each of the 4 phonemes in both French and Spanish by the English monolingual group. The bold numbers are cases in which a word received at least 33 percent of choices.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{3.1}$$

Figure 3.6 shows the probability of each word choice given a language and phoneme in the monolingual group and compares it to the conditional probability of the same

Table 3.2: The percentage of categorizations of phonemes in the English monolinguals group..

Stimulus Language	choice	i	o	ʌ	y
German	feel	**0.93**	0.01	0.01	0.06
German	fool	0.02	**0.45**	-	**0.78**
German	fought	0.05	**0.52**	**0.42**	0.11
German	fun	-	0.02	**0.57**	0.05
French	feel	**0.90**	-	0.14	0.04
French	fool	0.04	0.07	0.03	**0.69**
French	fought	0.04	**0.38**	0.31	0.21
French	fun	0.02	**0.54**	**0.52**	0.06

conditions in the bilingual group. These estimates were derived from the parameter estimates of the Bayesian Multinomial model, in which log odds were converted to probability. The bilinguals' conditional probability was calculated by dividing the probability from the full model by all categories of that language and shows the probability of a particular selection given that its language has been chosen. This calculation allows for the direct comparability between the monolingual and bilingual groups, since the bilinguals had both English and Spanish categories to choose from. In other words, the calculation of conditional probability tells us whether bilinguals behave similarly to monolinguals when they choose their L1 to categorize L3 sounds. The equation 3.1 illustrates the formula used to calculate the conditional probability of a word choice given that its language has been chosen.

The triangular shaped points represent the monolingual group and the circular points represent the conditional probability of the bilingual group in the event they chose and English category. The color of each point represents each word choice. In each panel of the figure, it is of interest whether the same color points of different shapes are close to one another, as this would indicate that the bilingual and monolingual groups are assimilating new language sounds to English similarly.

Tables 3.3, 3.4, 3.5 and 3.6 show the most probable responses of each group given a

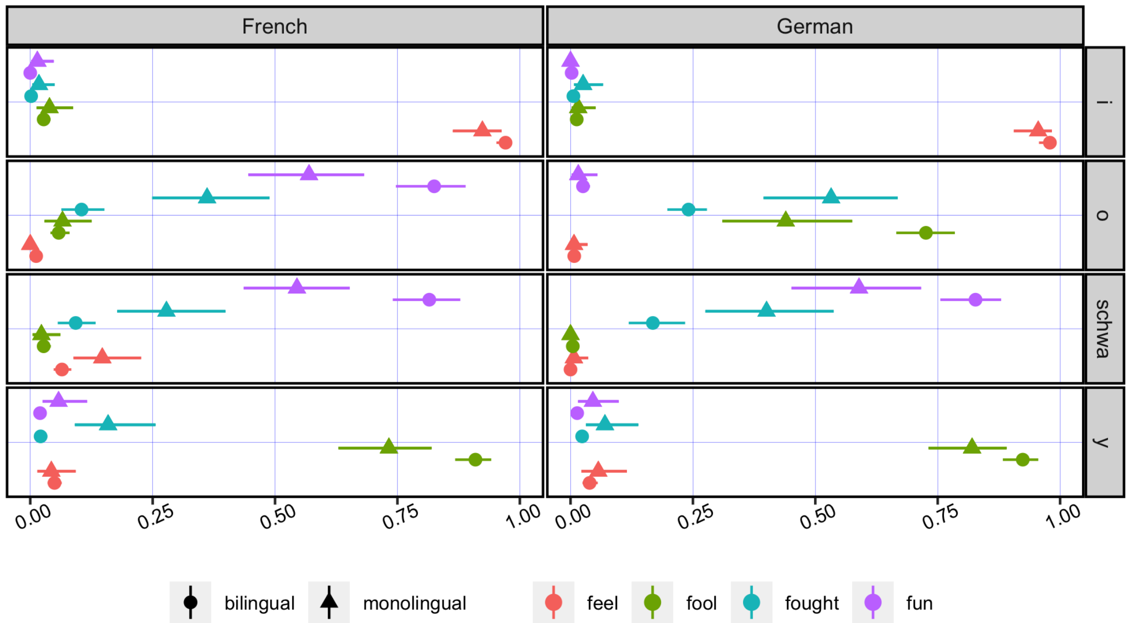


Figure 3.6: A comparison of the probability and conditional probability of each word choice per stimulus per language in the English monolingual group and English L1 bilingual group.

particular phoneme and language in both the bilingual and monolingual groups. The 2.5% and 97.5% columns represent the extremes of the 95% Highest Density Interval (HDI), in which 95% of the most probable parameter estimates fell. There were some differences between the monolingual and bilingual groups, which other categorizations were rather consistent. As inspection of Table 3.3 indicates that the phoneme /i/ was categorized similarly as *feel* by monolinguals and bilinguals when they chose an English category, regardless of L3. On the other hand, Table 3.4 shows that responses to /o/ saw some variation between groups. In particular, both *fought* and *fool* were probable responses by the monolingual group in German, while *fun* and *fought* were probable responses in French.

The bilingual group, however, preferred *fool* in German, and *fun* in French. Responses to /ʌ/ sound also varied. In German, the monolingual group preferred the choice *fun*, but *fought* was also a probable response. In French, the monolingual group preferred *fun*. The bilingual group preferred *fun* in both German and French. Finally, categorizations of /y/ were, like /i/, consistent between the groups and languages. Both groups assimilated /y/ to *fool* in both French and German.

Table 3.3: Most probable responses given /i/ in French and German by L1 English monolinguals and L1 English bilinguals.

Phoneme	Choice	Stimulus Language	Group	Probability	2.5%	97.5%
i	feel	German	monolingual	0.96	0.91	0.98
i	feel	French	monolingual	0.92	0.86	0.96
i	feel	German	bilingual	0.98	0.96	0.99
i	feel	French	bilingual	0.97	0.95	0.98

Table 3.4: Most probable responses given /o/ in French and German by L1 English monolinguals and L1 English bilinguals.

Phoneme	Choice	Stimulus Language	Group	Probability	2.5%	97.5%
o	fought	German	monolingual	0.53	0.39	0.67
o	fool	German	monolingual	0.44	0.31	0.58
o	fun	French	monolingual	0.57	0.45	0.68
o	fought	French	monolingual	0.36	0.25	0.49
o	fool	German	bilingual	0.73	0.67	0.79
o	fun	French	bilingual	0.83	0.75	0.89

3.2.5.1 Monolingual Spanish assimilations

Unlike the English groups, the Spanish monolingual group and Spanish bilingual group showed evidence of similar categorization patterns overall when the Spanish bilingual group picked a Spanish category to categorize an L3 sound. Table 3.14 shows the percentage of choices given a phoneme in the Spanish monolingual group in both French and German. In each of the four cases, the bilinguals and monolingual groups chose the same word the highest percentage of the time. In particular, given /i/, both groups chose *fin* the most, while both /o/ and the /ʌ/ were matched to *son*, and /y/ was assimilated to *su* by both groups in both languages. Figure 3.7 visualizes the

Table 3.5: Most probable responses given /ʌ/ in French and German by L1 English monolinguals and L1 English bilinguals.

Phoneme	Choice	Stimulus Language	Group	Probability	2.5%	97.5%
/ʌ/	fun	German	monolingual	0.59	0.45	0.72
/ʌ/	fought	German	monolingual	0.40	0.28	0.54
/ʌ/	fun	French	monolingual	0.54	0.44	0.65
/ʌ/	fun	German	bilingual	0.83	0.76	0.88
/ʌ/	fun	French	bilingual	0.82	0.74	0.88

Table 3.6: Most probable responses given /y/ in French and German by L1 English monolinguals and L1 English bilinguals.

Phoneme	Choice	Stimulus Language	Group	Probability	2.5%	97.5%
y	fool	German	monolingual	0.82	0.73	0.89
y	fool	French	monolingual	0.73	0.63	0.82
y	fool	German	bilingual	0.92	0.88	0.96
y	fool	French	bilingual	0.91	0.87	0.94

probabilities generated by the Bayesian Multinomial regression model. In all cases, the bilingual and monolingual participants appear to categorize French and German sounds similarly.

3.2.5.2 English L1 group

Figure 3.8 shows the categorization data of the English L1 group of each phoneme in both languages. The shaded bars represent the rating for goodness of fit, where a lighter shade represents a higher average rating. Tables 3.12 and 3.13 show the numerical values of the French and German categorization in the English L1 group respectively, where the choice with the highest percentage per for each phoneme is in bold. In the event that there were two choices that were above 33%, they are both in bold.

Table 3.7: The percentage of categorizations of phonemes in the Spanish monolinguals group.

Stimulus Language	choice	i	o	/ʌ/	y
German	fin	**0.96**	0.09	0.28	0.16
German	son	0.01	**0.59**	**0.56**	0.08
German	su	0.03	0.32	0.17	**0.76**
French	fin	**0.83**	0.13	0.25	0.20
French	son	0.03	**0.61**	**0.47**	0.06
French	su	0.13	0.26	0.29	**0.74**

Table 3.8: Most probable responses given /i/ in French and German by L1 Spanish monolinguals and L1 Spanish bilinguals

Phoneme	Choice	Stimulus Language	Group	Probability	2.5%	97.5%
i	fin	German	monolingual	0.98	0.95	0.99
i	fin	French	monolingual	0.87	0.80	0.92
i	fin	German	bilingual	0.99	.98	1.00
i	fin	French	bilingual	0.99	.98	1.00

As can be seen in Tables 3.12 and 3.13, the English L1 group, given the L3 French phoneme /i/, chose both their English category /i/ provided in the choice *feel* and their Spanish category /i/ provided in fin *fin*. German /i/ was categorized similarly. For /o/ and, this group chose *fun* the most often in both cases in French, while the most chosen word in German for /o/ was *fool*. The wedge in both French in German was most often assimilated to the intended category *fun* Finally, both the L3 French and German phonemes /y/ resulted in choices of *fool* and *su*, the English and Spanish categories for /u/.

3.2.5.3 Spanish L1 group

Figure 3.9 shows the percentage that each answer choice was chosen given a particular phoneme in each language. Table 3.14 and Table 3.15 similarly report the percentage of each answer choice given a particular phoneme in each language in the Spanish L1

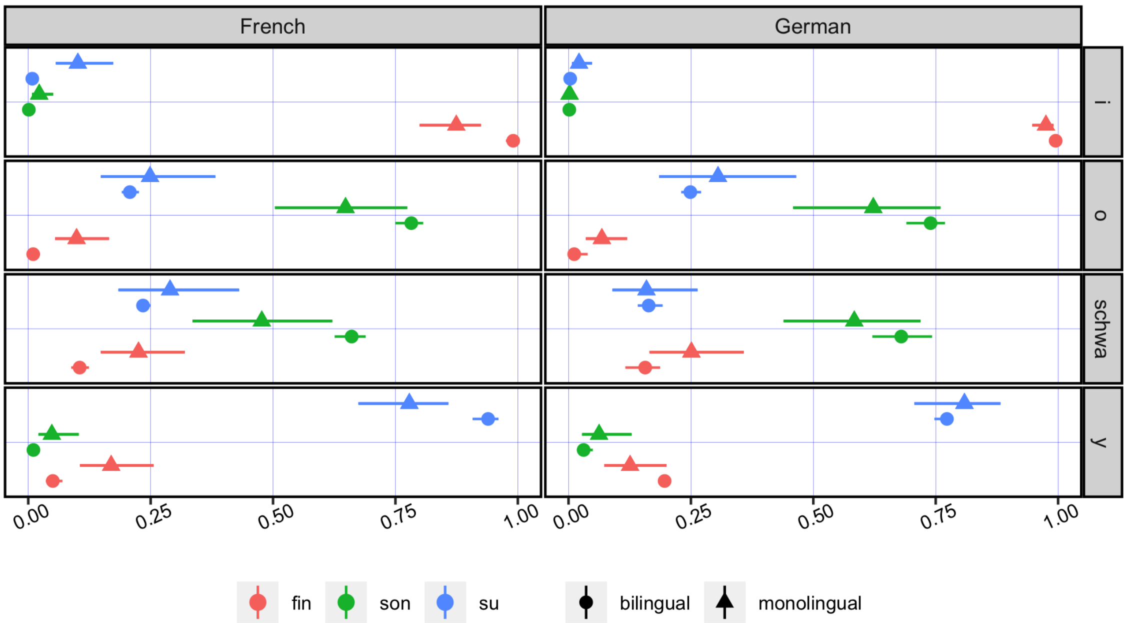


Figure 3.7: A comparison of the probability and conditional probability of each word choice per stimulus per language in the Spanish monolingual group and Spanish L1 bilingual group.

group. The Spanish L1 group had a similar preference to the English L1 group in their categorization of /i/, where both English and Spanish categories were chosen. However, the Spanish L1 group chose the English category given a German stimulus more often than the Spanish category. In both French and German, when both /o/ and the wedge were played, the most chosen answer was *fought*. Finally, given the phoneme /y/, the Spanish L1 group chose they English *fool* most often in German, and the Spanish *su* most often in French.

3.2.5.4 Results of the models

The Bayesian multinomial regression models provided an inferential framework which allows for the quantification of uncertainty and avoids the pitfalls of the over-reliance on so-called "statistical significance". The output of each model was converted from log-odds to probability using a combination of the `conditional_effects` and `make_conditions` functions in the R package `brms` (Bürkner, 2017). Thus, for each

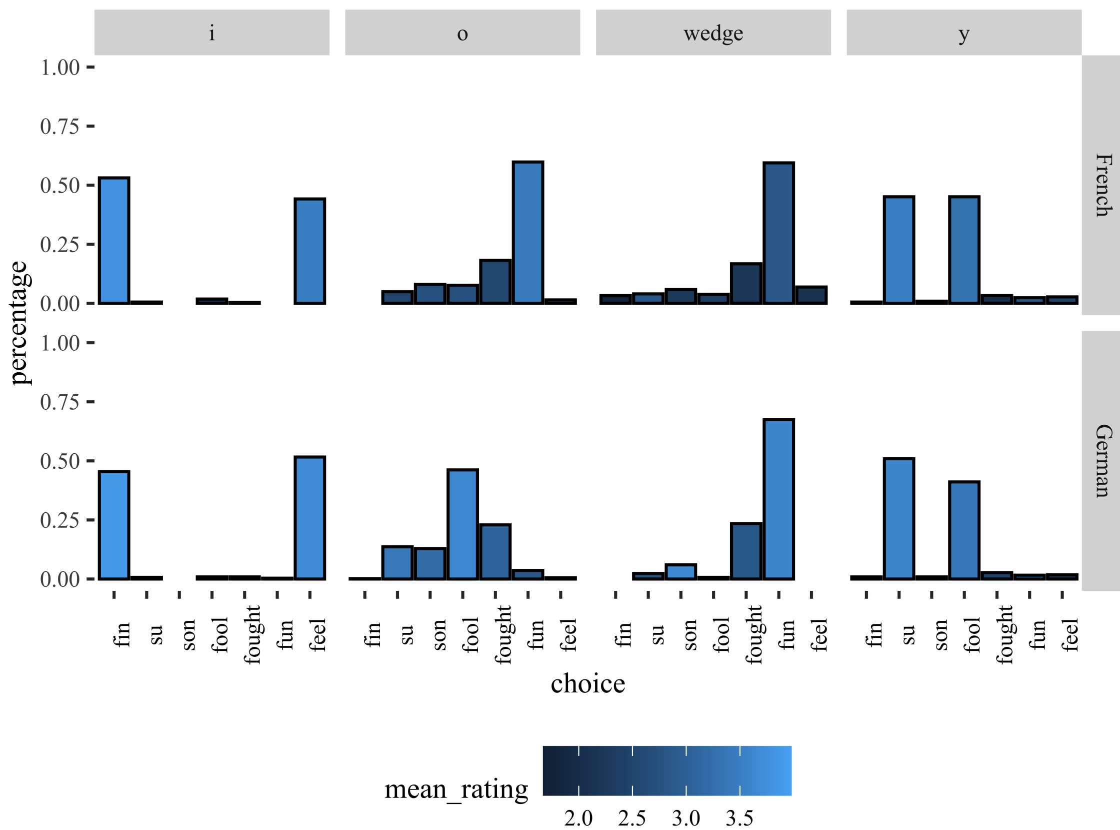


Figure 3.8: Percentage of each word choice given a phoneme in French and German in the L1 English-L2 Spanish group and Self-reported spoken production

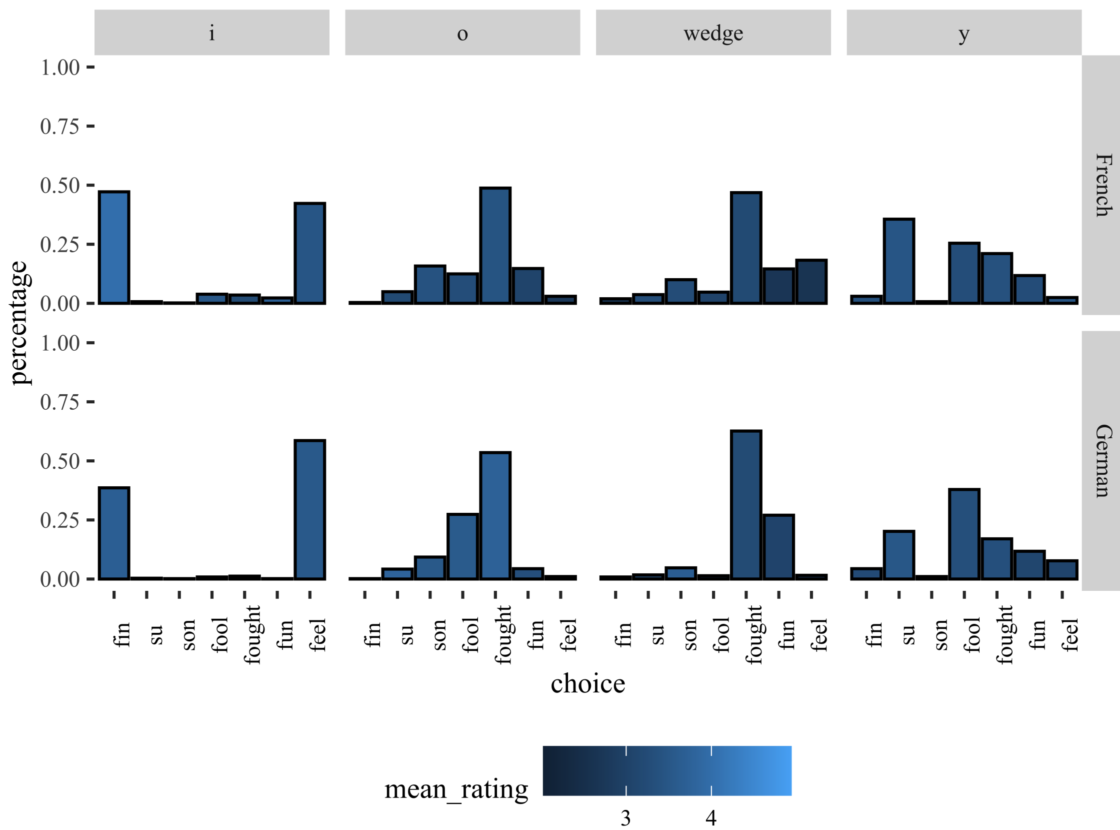


Figure 3.9: Percentage of each word choice given a phoneme in French and German in the L1 Spanish-English L2 group.

Table 3.9: Most probable responses given /o/ in French and German by L1 Spanish monolinguals and L1 Spanish bilinguals

Phoneme	Choice	Stimulus Language	Group	Probability	2.5%	97.5%
o	son	German	monolingual	0.62	0.46	0.76
o	son	French	monolingual	0.65	0.50	0.77
o	son	German	bilingual	0.74	0.69	0.77
o	son	French	bilingual	0.78	0.75	0.81

Table 3.10: Most probable responses given /ʌ/ in French and German by L1 Spanish monolinguals and L1 Spanish bilinguals

Phoneme	Choice	Stimulus Language	Group	Probability	2.5%	97.5%
ʌ	son	German	monolingual	0.58	0.44	0.72
ʌ	son	French	monolingual	0.48	0.34	0.62
ʌ	son	German	bilingual	0.68	0.62	0.74
ʌ	son	French	bilingual	0.66	0.63	0.69

phoneme in each language, the sum of the probability of choosing the categories combined is 1. Figure 3.10 shows the probability of each choice in French and German by both groups when /i/ was the phoneme. Figure 3.11, shows the same set of probabilities when /o/ is played. Figures 3.12 and 3.13 represent the probabilities of each response when the wedge and /y/ are played respectively.

Figure 3.10 corroborates the trends observed in the descriptive statistics when the participants in both groups were asked to categorize the phoneme /i/ in French and German. In particular, for the English L1 group in both French and German and the Spanish L1 group in French, portions of the credible parameter estimates from the posterior distribution overlap with a probability of .5, suggesting that compelling evidence is not present for a preference for the English or Spanish category. Differently, the Spanish L1 group did show evidence in German of a preference for the English category *feel*, in which the probability of picking *feel* when the stimulus was German

Table 3.11: Most probable responses given /y/ in French and German by L1 Spanish monolinguals and L1 Spanish bilinguals

Phoneme	Choice	Stimulus Language	Group	Probability	2.5%	97.5%
y	su	German	monolingual	0.81	0.71	0.88
y	su	French	monolingual	0.78	0.67	0.86
y	su	German	bilingual	0.77	0.75	0.78
y	su	French	bilingual	0.94	0.91	0.96

Table 3.12: The percentage of categorizations of French phonemes in the L1 English group.

choice	i	o	/ʌ/	y
feel	**0.44**	0.01	0.07	0.03
fin	**0.53**	-	0.03	0.01
fool	0.02	0.08	0.04	**0.45**
fought	-	0.18	0.17	0.03
fun	-	**0.60**	**0.59**	0.02
son	-	0.08	0.06	0.01
su	0.01	0.05	0.04	**0.45**

was 0.61 (HDI = 0.53 - 0.68). On the other hand, when phoneme was /i/ in French the probability of choosing *feel* was 0.46 (HDI = 0.37 - 0.55).

After hearing the stimulus /o/, the English L1 group preferred the choice *fun* in French (0.78 (HDI = 0.67 - 0.87), and *fool* (0.57 (HDI = 0.48 - 0.64) in German, and the Spanish L1 group preferred *fought* in both languages. /ʌ/ was categorized as the choice *fun* in both French (0.78 [HDI = 0.67 - 0.87] and German (0.79 [HDI = 0.67 - 0.87]. The Spanish L1 group preferred *fought* in both languages. (French: 0.52 [HDI = 0.44 - 0.6]; German: 0.68 [HDI = 0.61 - 0.74].

The English L1 group assimilated /y/ in both French and German to their English and Spanish /u/ (choices *su* and *fool*), without a clear preference for either. Probability of the English L1 group's choice of *su* given the phoneme /y/ and the stimulus language

Table 3.13: The percentage of categorizations of German phonemes in the L1 English group.

choice	i	o	/ʌ/	y
feel	**0.52**	0.01	0.02	
fin	**0.45**	0.00	0.01	
fool	0.01	**0.46**	**0.41**	0.01
fought	0.01	0.23	0.03	0.23
fun	0.00	0.04	0.02	**0.67**
son	-	0.13	0.01	0.06
su	0.01	0.14	**0.51**	0.02

Table 3.14: The percentage of categorizations of French phonemes in the L1 Spanish group.

choice	i	o	/ʌ/	y
feel	**0.42**	0.03	0.18	0.02
fin	**0.47**	0.00	0.02	0.03
fool	0.04	0.12	0.05	0.25
fought	0.04	**0.49**	**0.47**	0.21
fun	0.02	0.15	0.15	0.12
son	0.00	0.16	0.10	0.01
su	0.01	0.05	0.04	**0.36**

German: 0.53 [HDI = 0.44 - 0.61]. Probability of the English L1 group's choice of *fool* given the phoneme /y/ and the stimulus language German: 0.42 [HDI = 0.34 - 0.51] Probability of the English L1 group's choice of *su* given the phoneme /y/ and the stimulus language French: 0.47 [HDI = 0.36 - 0.58]. Probability of the English L1 group's choice of *fool* given the phoneme /y/ and the stimulus language French: 0.48 [HDI = 0.37 - 0.58] The Spanish L1 group, on the other hand, assimilated French /y/ to the Spanish *su*, and the German /y/ to English *fool*. Probability of the Spanish L1 group's choice of *su* given the phoneme /y/ and the stimulus language French: 0.35 [HDI = 0.24 - 0.46]. Probability of the Spanish L1 group's choice of *fool* given the phoneme /y/ and the stimulus language German: 0.43 [HDI = 0.35 - 0.5]

Table 3.15: The percentage of categorizations of German phonemes in the L1 Spanish group.

choice	i	o	/ʌ/	y
feel	**0.59**	0.01	0.02	0.08
fin	**0.39**	0.00	0.01	0.04
fool	0.01	0.27	0.01	**0.38**
fought	0.01	**0.54**	**0.63**	0.17
fun	0.00	0.04	0.27	0.12
son	0.00	0.09	0.05	0.01
su	0.00	0.04	0.02	0.20

3.3 Experiment 2 - AX Discrimination Task

Experiment 2 examined whether English vowel contrasts could be discriminated when they are produced in the context of French and German words, by Spanish-English bilinguals and English monolinguals. Spanish monolinguals were not recruited for this experiment, since the focus was on English vowel contrasts, and this population would not have acquired English. Just as in experiment 1, each participant completed a background questionnaire and a perceptual assimilation task.

3.3.1 Participants

A total of 150 participants took part in experiment 2. 82 of these participants also completed experiment 1. This total sample was made up of three total groups: L1 Spanish-L2 English speakers (n = 64; henceforth the Spanish L1 group), L1 English-L2 Spanish group (n = 37; henceforth the English L1 group) and an English monolingual group (n = 49). All participants were recruited on prolific and were pre-screened according to the same criteria as experiment 1. In addition to filters in place from Prolific.co, the participants were screened further using an adapted version of the Bilingual Language Profile (Birdsong et al., 2012). All participants who answered ‘no’ to the question “Do you speak a language other than English and Spanish” were permitted to continue the experiment.

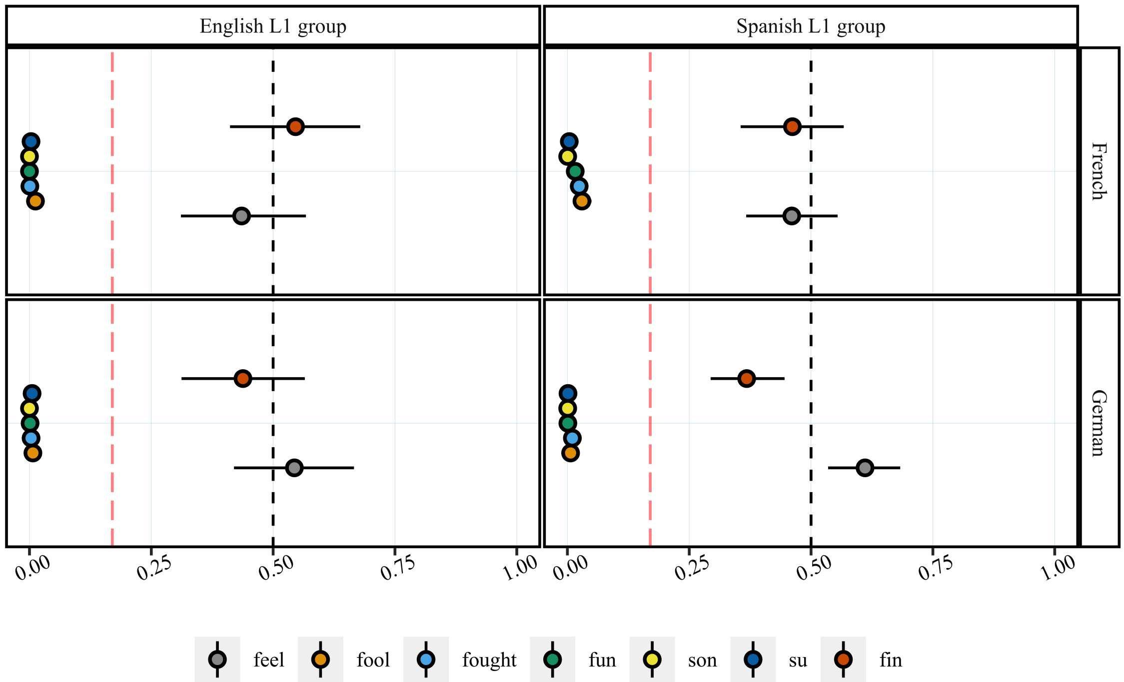


Figure 3.10: Probability of each choice given /i/

3.3.2 Materials

3.3.2.1 Acoustic continua

Several 7-step vowel continua were created for this experiment to measure two vowel contrasts in “French” and “German”. In each case, a German or French word-pair was produced by a native English speaker by producing the word on both ends of the continuum by producing, for example, the sheep-like vowel or the ship-like vowel in the given word. Following these productions, the vowel sounds were extracted and a single onset and offset was saved per word. Then, using a PRAAT script, the first 3 formants were manipulated to create the continuum, and the vowels were spliced with the saved onsets and offsets, to re-create whole words.

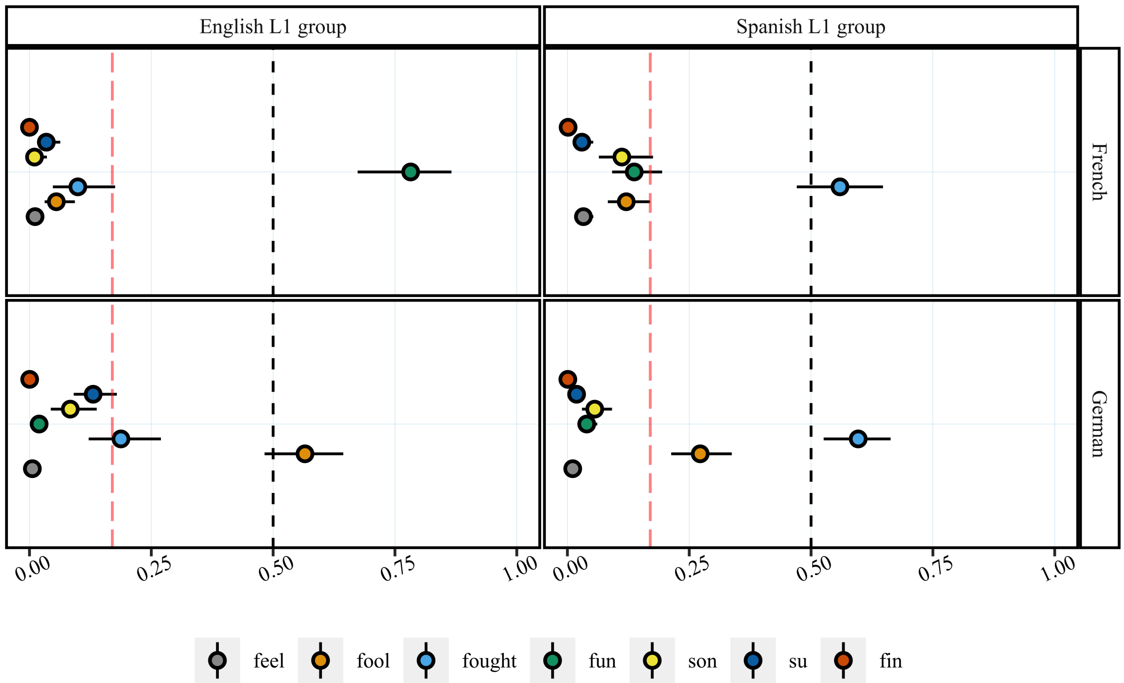


Figure 3.11: Probability of each choice given /o/

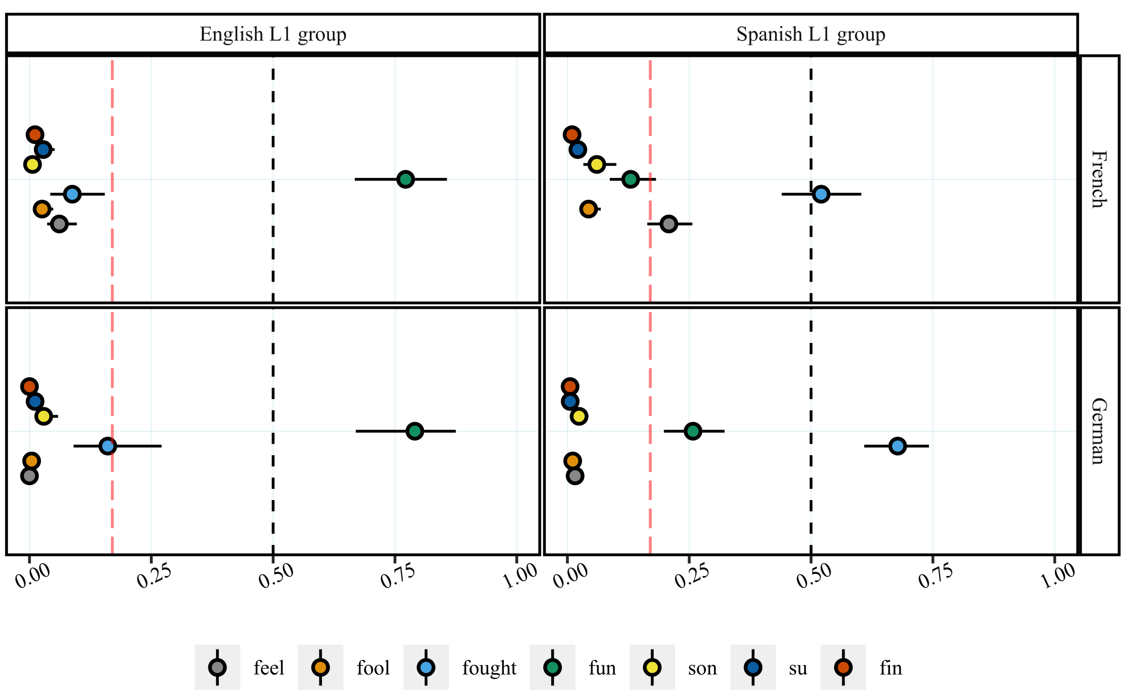


Figure 3.12: Probability of each choice given schwa

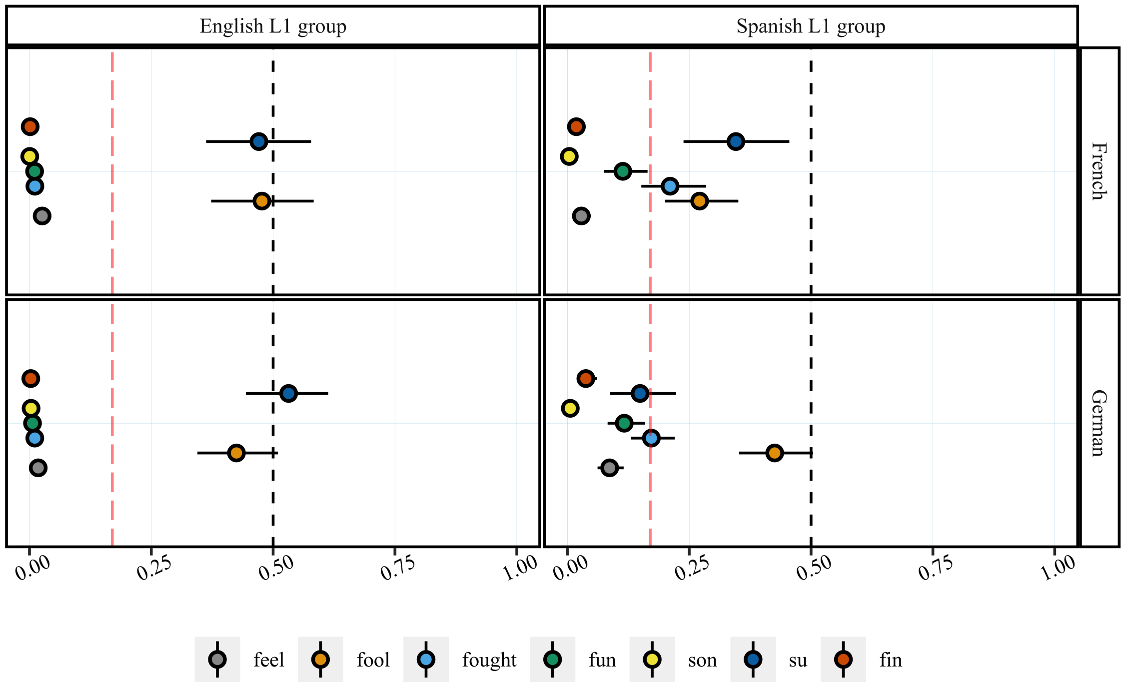


Figure 3.13: Probability of each choice given /y/

3.3.2.2 2-alternative Forced Choice Task

Different or the same? The same word is played twice, with either the same or distinct vowel sounds - participants must choose whether they hear a difference.

eight 7-step continua were created for use in an AX discrimination task, where four were used in French and four were used in German.

French fricative 1: fils - son

French fricative 2: fard - blush/make up

French bilabial 1: pic - peak

French bilabial 2: pas - not

German fricative 1: fisch - fish,

German fricative 2: faden - thread

German bilabial 1: pils - beer

German bilabial 2: paar - pair

Due to an error, the German continuum had three /i/ continua and one /a/ continuum. However, this does not appear to have impacted the results, since discrimination was very poor for every contrast.

3.3.3 Results

Overall, the discrimination of every sound contrast was very poor.

3.4 Discussion

The present study examined the categorization of vowel sounds in two unknown languages by Spanish-English bilinguals in both orders of acquisition. Although there was an overall preference for English, the results suggested that bilinguals do have access to both their L1 and L2 during L3 perception and categorization, a finding which single-language access models, such as the TPM and L2SF cannot account for. In several instances, L3 sounds had a similar probability of being assimilated to English or Spanish sound, such as in 3 out of 4 cases with the phoneme /i/. The results were mixed in terms of whether the L3 that the participants heard impacted their preference for English or Spanish categories. In total, there were three cases in which the L3 did result in the preference for an English or Spanish category out of 16 possible. In particular, the Spanish L1 group showed a slight preference for the English /i/ (the choice *feel*) when they heard /i/ in German, but not in French. Additionally, the Spanish L1 group also categorized the phoneme /y/ differently when they heard it in L3 French than when they heard it in L3 German. In particular, the L3 French /y/ was assimilated to Spanish /u/ (*su*) and the L3 German /y/ was assimilated to English /u/ (*fool*).

These results are best explained by the Linguistic Proximity Model (Westergaard, 2021; Westergaard et al., 2017), which suggests that L3 learners have access to categories in both their L1 and L2 during L3 acquisition. On the other hand, The Typological Primacy Model (Rothman, 2010, 2011, 2013, 2015) suggests that, while bilinguals have access to both their languages at first exposure, this access is lost at some point during what the model refers to as the "initial stages" of L3 acquisition. While the present data does not necessarily provide counter-evidence to the TPM, it provides a starting point for future research by providing a picture of the categorization patterns of L3 phonemes in two languages by Spanish-English bilinguals which could be examined in a longitudinal design. If the TPM is correct, a distinct change in categorization would be expected to reflect a bias for a single language's category. The present study provides a basis of comparison, and evidence that bilinguals do have access to both the L1 and L2 at first exposure. However, in 6 our of 8 cases, global typological similarity (English and German vs. French and Spanish), did not result in any obvious bias in categorization.

It is unclear why the Spanish L1 group seemed to be sensitive to which L3 they heard, where the English L1 group was not. It is worth noting that there are differences between the groups in self-rated proficiency and age of onset and acquisition; the Spanish L1 group rated themselves as more proficient on average in spoken and perceptual abilities and had an earlier age of onset and acquisition on average. The role of proficiency has been addressed in L3 models, but its precise role in influence is not yet clear. The TPM has stated that the L2 must be sufficiently proficient to be a source of influence. In the TPM studies to date, this has been taken to mean "advanced proficiency". Though the present study was not interested in the effect of L2 proficiency on L3 categorizations, it is worth noting that the group differences in proficiency might have impacted the results, although it is difficult to say exactly how. This notion of the impact of L2 proficiency CLI could be further investigated, in which the more traditional categorical grouping of proficiency (e.g., novice, intermediate, advanced) could also be examined in a continuous fashion and corroborated across measures. In other words, it's unclear for the purpose of the present study, and likely with many proficiency measures, where the TPM's cutoff for "advanced" should be.

The present study also had limitations. Firstly, online studies offers unique challenges and limitations. One potential issue is the lack of the ability to control the participant's environment outside of the instructions and experiment itself. As a result, language mode effects (Casillas & Simonet, 2018; Grosjean, 1998) cannot be completely ruled out. Additionally, headphone quality, speaker quality, background noise and volume level are all variables within the participant's control that could not be reasonably controlled for in the context of online data collection. Second, self-reported proficiency, while convenient and fast, is a subjective measure of language ability that would likely be improved by a more objective proficiency measure such as the LexTALE (Lemhöfer & Broersma, 2012). While the present study did not initially hypothesize that differences in proficiency impact categorization, a look at the data suggests that this is a possibility.

Chapter 4: The Production of the new language vowels by Spanish-English bilinguals

4.1 Context

More general paragraph.

4.2 What has been done

Research in L3 production has yielded mixed results. Global accent production - L2 then L1. L2 global also found by wrembel in accent judgments.

VOT - intermediate Llama 1, llama 2, Parrish

Vowels -

4.3 What has not been done

4.4 How this study fills the gap

New language pairing, and use of first exposure

Hammarberg study.

The results of this study corroborated the so-called ‘foreign language effect’ first suggested by Meisel (1983)

The informants rated the speaker as having English-accented Swedish. Further studies in L3 phonology yielded mixed results, with some studies reporting L2 status effects (Gut, 2010; Llama et al., 2010; Wrembel, 2010), while others concluded that there was the mixed influence of both the L1 and the L2 (Llama & Cardoso, 2018; Wrembel, 2014).

4.5 Methods

4.5.1 Participants

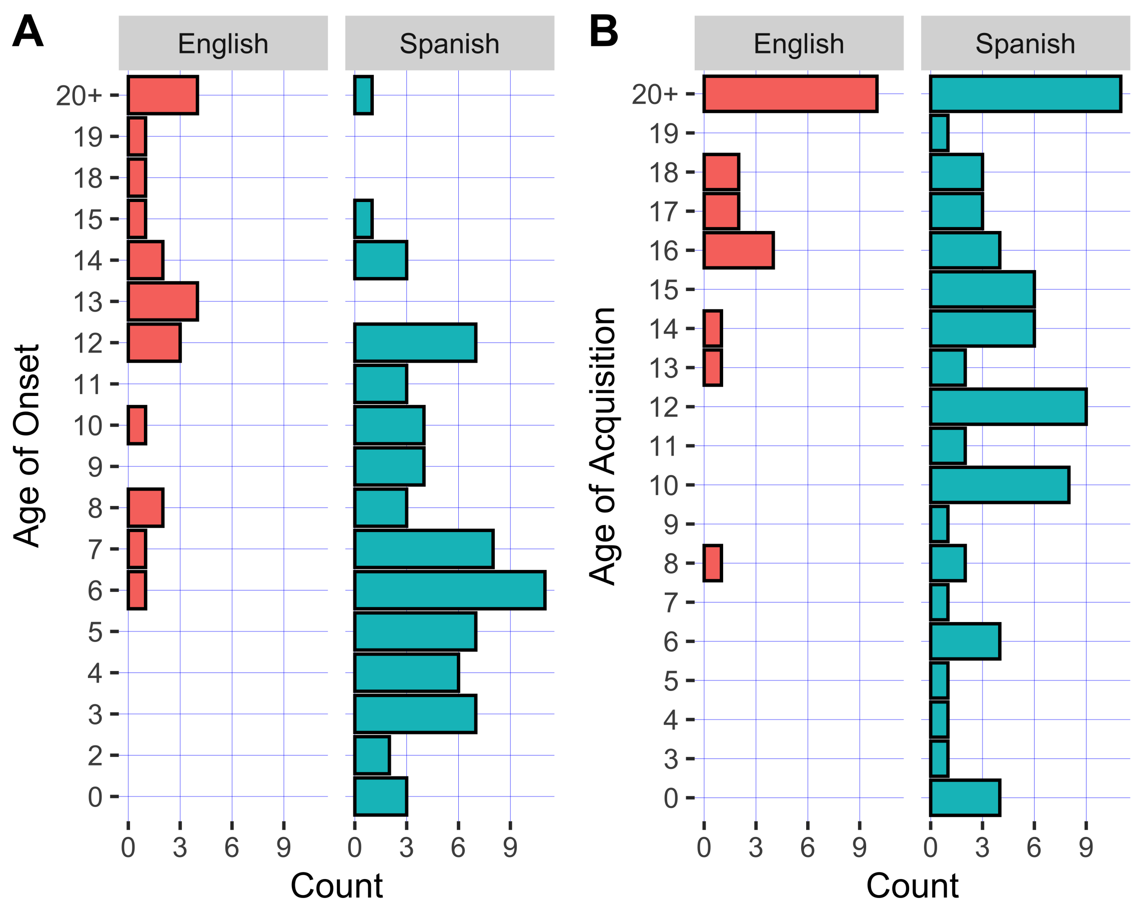
L1 English bilingual (n = 20)

L1 Spanish bilingual (n = 66)

L1 English monolingual (n = 58)

L1 Spanish monolingual (n = 28)

?? shows the age of Onset (A) and age of acquisition (B) by the L1 English and L1 Spanish bilingual groups.



4.5.2 Materials

4.5.3 Procedure

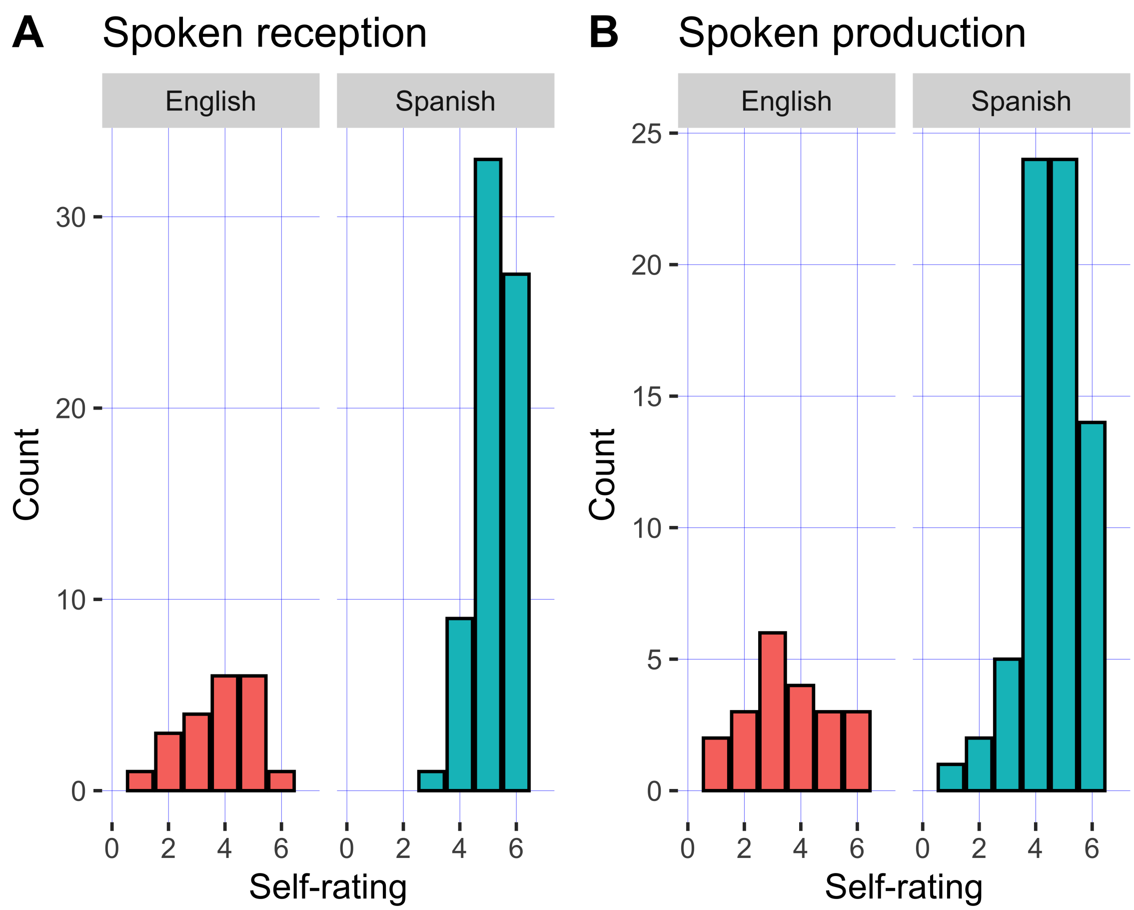


Figure 4.1: Self-rated spoken and eceptive proficiency by the L1 English and L1 Spanish bilingual groups

4.5.4 Statistical Analysis

A series of Bayesian multilevel regression models were run to determine whether there were differences in vowel height (F1) an and frontness (F2) between groups. In each model, the formant was analyzed as a function of language (4 levels for bilingual groups, 3 levels for monolingual groups), and phoneme (4 levels). Random intercepts for token and participant were also included to take into account the nested structure of the data. Model priors were the default in `brms`, a student's T distribution with 3 degree of freedom. All models were fit with 4000 iterations (1000 warm-up). Hamiltonian Monte-Carlo sampling was carried out with 6 chains distributed between 8 processing cores.

4.6 Results

4.6.1 Forced Alignment

A total of 15019 data points were force aligned in the initial analysis using Webmaus basic (*cite*). In order to estimate the accuracy of the forced-alignment, about 3 percent of the bilingual data was randomly subset and hand-corrected (380 tokens). Subsetting was done using an R script. Figure 4.2 shows the quantity of each phoneme in each language that was hand-corrected in the random subset.

Figure 4.3 shows the formant values of hand-corrected tokens and force-aligned tokens. An inspection of the figure suggests that /a/, wedge and /o/ were mostly accurately aligned. On the other hand, the tokens for /i/ and /u/ benefited from hand-correction.

Figure 4.4 shows the root mean square deviation between the hand-corrected values and force-aligned values. Root mean square deviation is an measurement of the difference between predicted and actual values. In the present study, these numbers represent Hertz for F1 and F2. Each numerical label in the figure represents the root mean square deviation, while the colors represent phonemes. The phonemes /i/ and /u/ have the largest root mean square deviation, showing, again, that they were the

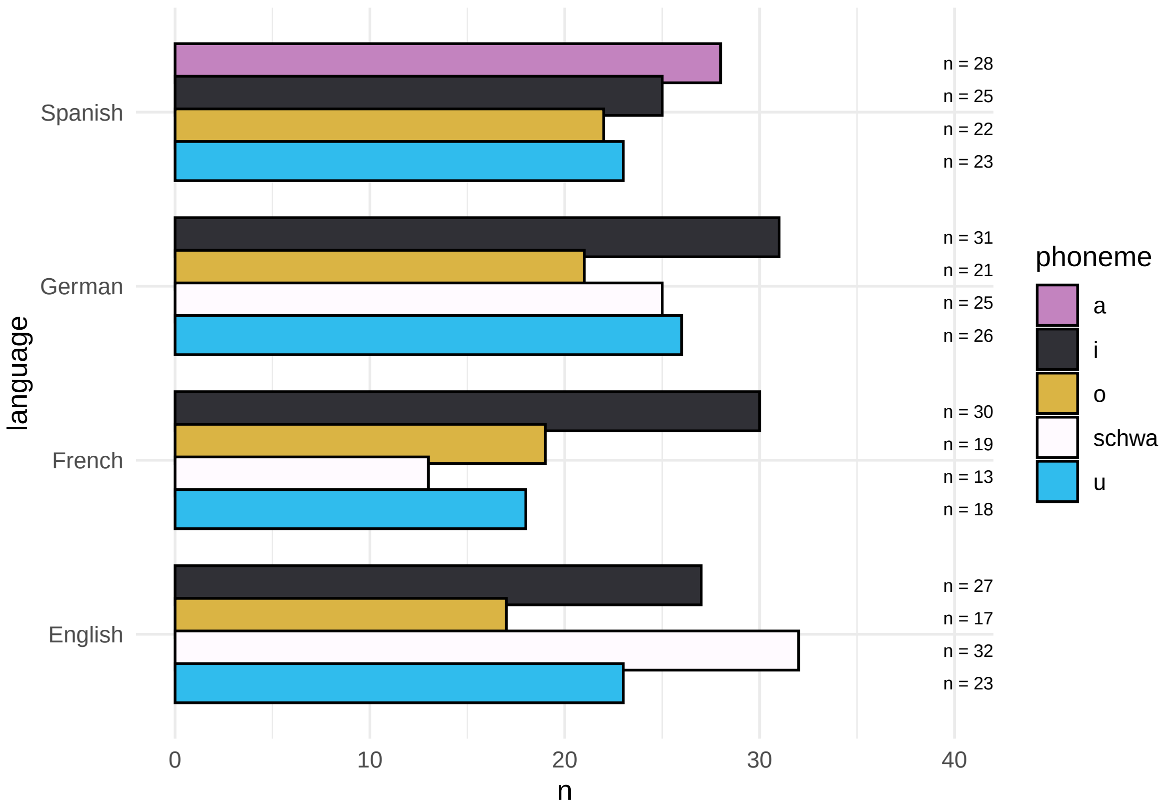


Figure 4.2: Quantity of Hand-correced tokens per phoneme per language

Figure 4.3: Hand Corrected versus automated formant values of a subset of the data

least accurately force-aligned.

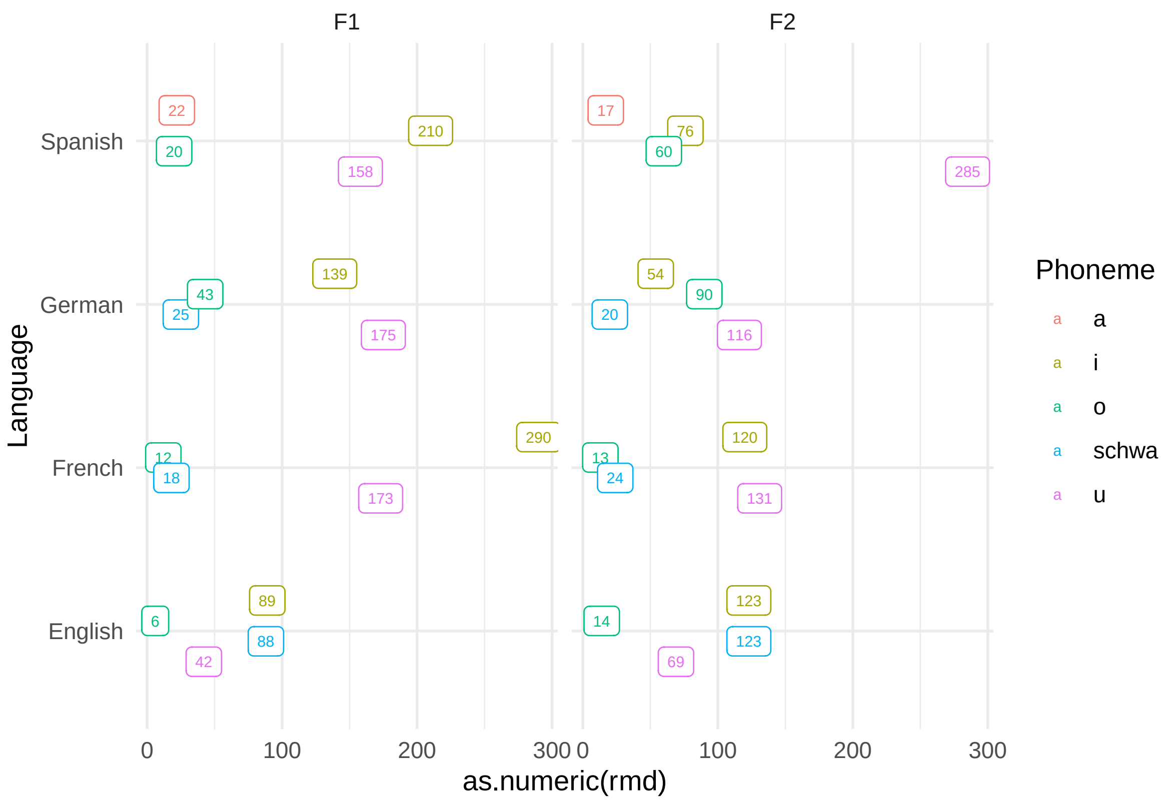


Figure 4.4: Root Mean Square Deviation for F1 and F2

A total of 2005 tokens were removed from the dataset. A token was considered an outlier when it was more than plus or minus two standard deviations of the mean of the hand-corrected subset. Table A.1 in the appendix shows the mean and standard deviations of each phoneme in each language used to identify outliers in the main dataset. Figure 4.5 shows the number of removed tokens compared to the root mean square deviation. The most removed phoneme was German /o/, in which 371 tokens were removed as outliers. Table 4.1 summarizes the total number of each phoneme removed. Figure 4.6 shows the the final number of tokens analyzed after removal of outliers (n = 13014) broken down by language and phoneme.

Table 4.1: Total Number of each removed phoneme

phoneme	n
a	54
i	496
o	676
schwa	278
u	501

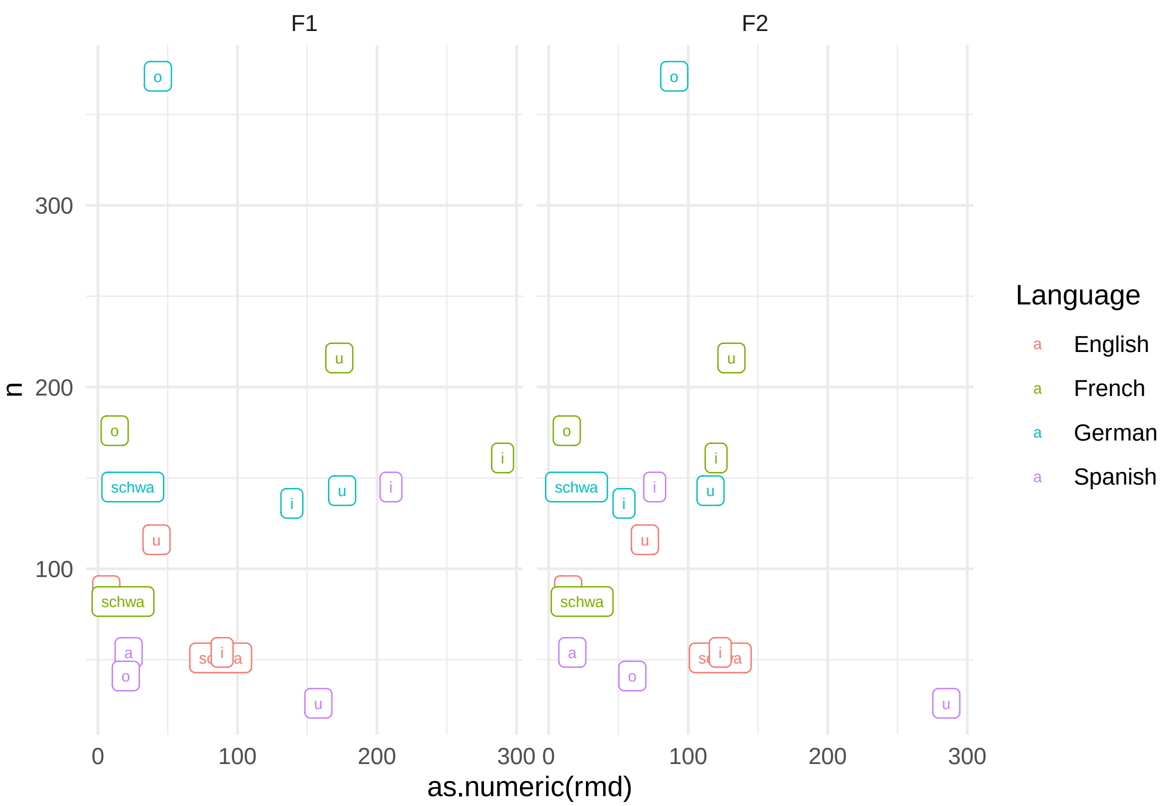


Figure 4.5: Root Mean Square Deviation for F1 and F2 versus total removed tokens

4.6.2 Analysis of the included data

Figure 4.7 shows the productions of each of the four groups for the phonemes /u/ (in Spanish and English) and /y/ (in French and German). This condition was intended

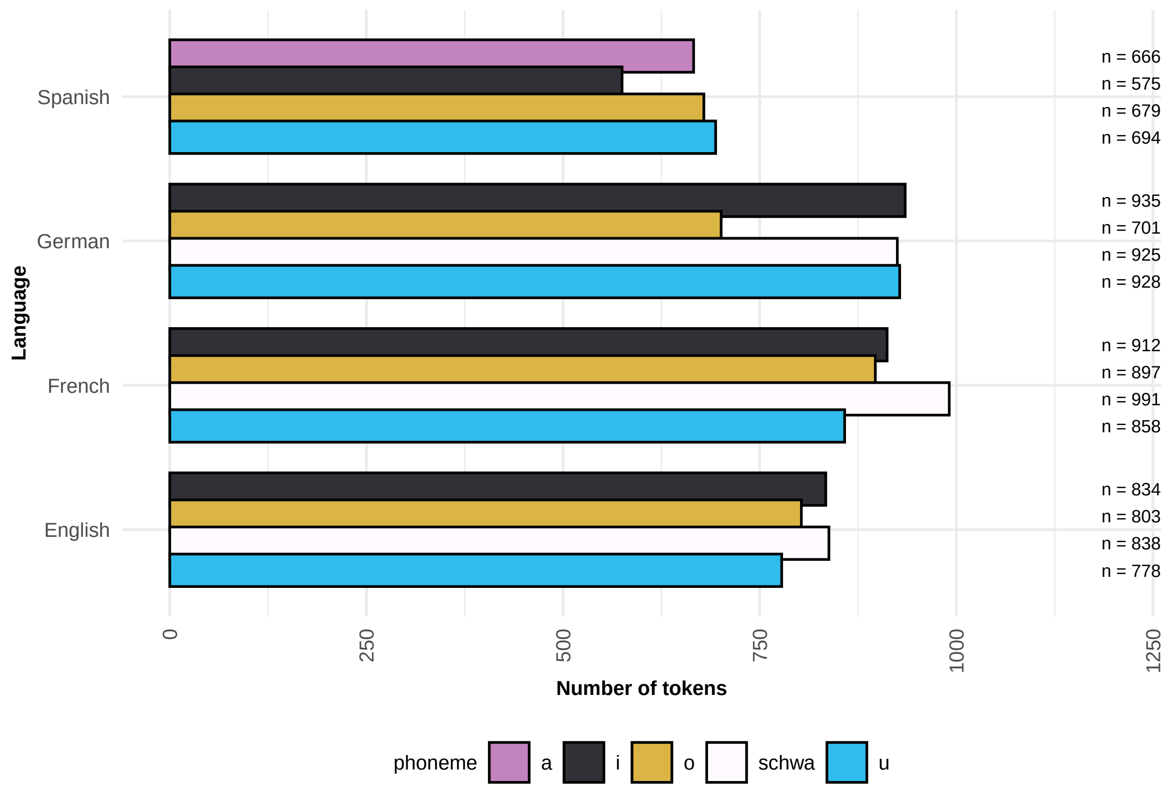


Figure 4.6: Total Number of Tokens per phoneme and language after removal of outliers

to explore how Spanish-English bilinguals would produce /y/, which is not present in either Spanish or English. An inspection of this figure suggests that /y/ was effectively imitated by all groups. In addition, it does not appear that there is much meaningful difference between the French and German imitations of the /y/ sound.

The models further supported the effective imitation of French and German /y/ by both bilingual groups. Figure 4.8 shows the conditional effects of the F1 and F2 models for /u/ in English and Spanish, and /y/ in French and German. Overall, the results of the models suggest that both bilingual groups performed similarly and could effectively imitate the French and German /y/. The models provided evidence that both bilingual groups had more fronted production of French and German /y/ relative to their Spanish and English /u/ (panel b). Additionally, although the English /u/ was slightly more fronted than Spanish /u/ by both groups, there was compelling evidence that productions English and Spanish /u/ German and French /y/ were distinct in terms of frontness, since there was no overlap in the posterior distributions.

Figure 4.9 shows the productions of /o/ by all groups in Spanish, French and German, while the English sound was /ɑ/. Although this condition was intended to bias the use of Spanish, this visualization does not imply that this was successful, nor that there were many meaningful differences between groups. It appears that all groups distinguish and similarly imitate the French and German /o/ sounds, regardless of whether they are bilingual or monolingual, or whether they differently produce the Spanish /o/ from the English /ɑ/. Figure 4.10 again shows the conditional effects of the Bayesian models for F1 and F2. The models shows that both bilingual groups produced English /ɑ/ and French /o/ as more fronted than Spanish and German /o/ (panel b), while height varied by both groups as a function of language (panel a).

Figure 4.11 shows the productions of /i/ in all languages and all groups. This condition was intended to examine how Spanish-English bilinguals produced novel language sounds that were present in both their L1 and L2. The figure does suggest that any of the groups produced /i/ distinctly as a function of language. The Bayesian

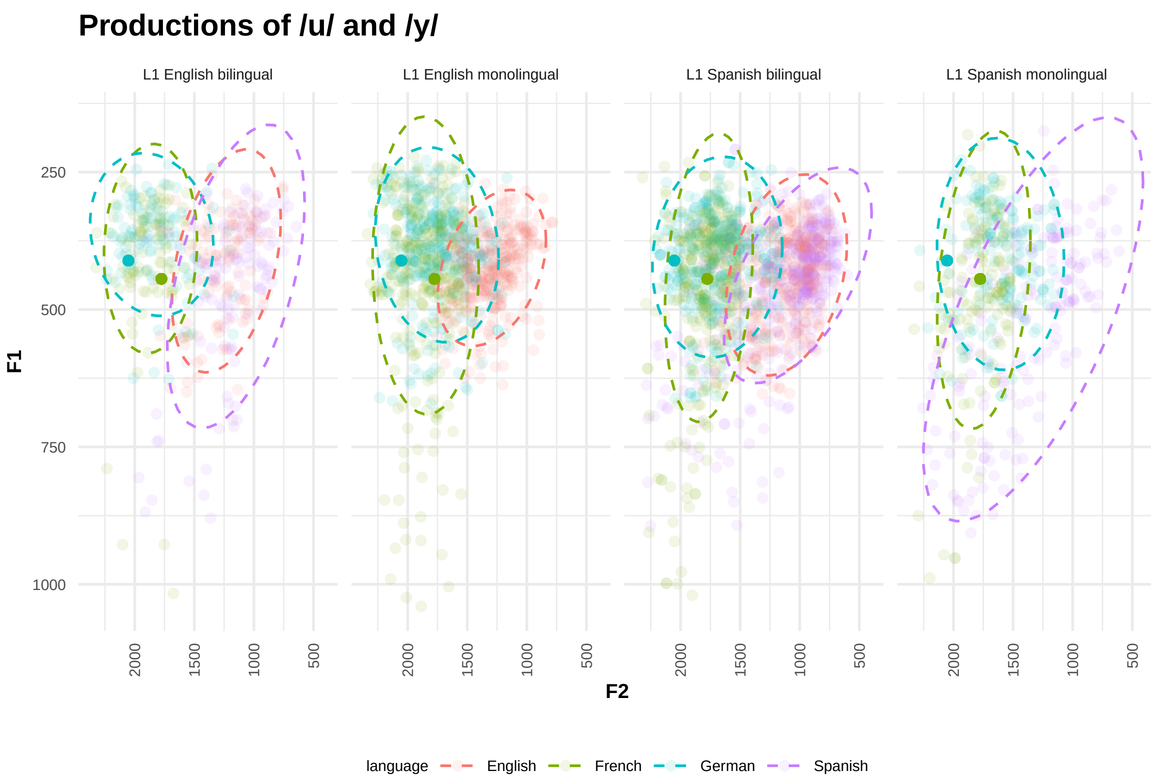


Figure 4.7: Productions of /u/ in each language by each group

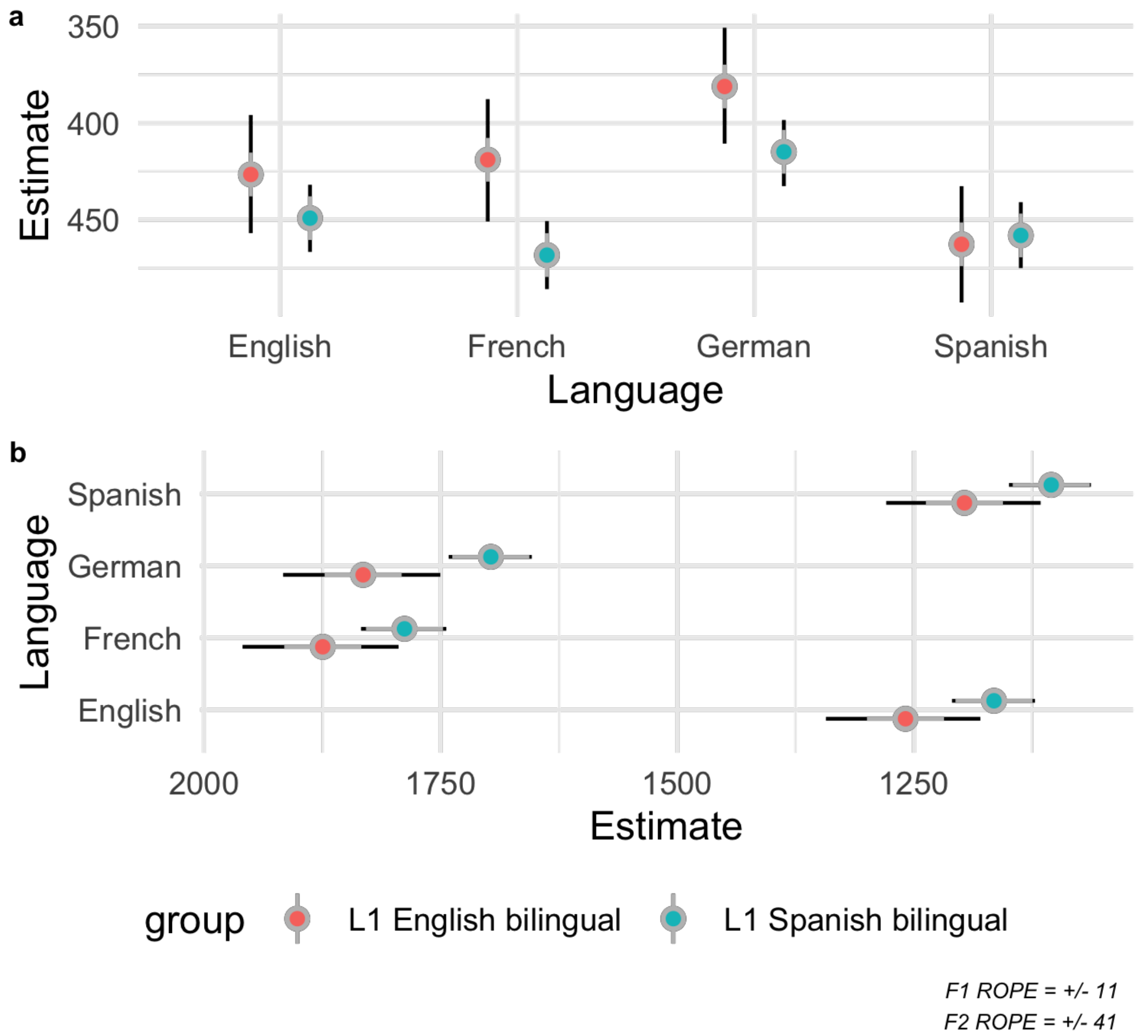


Figure 4.8: Conditional effects plot of /u/ for F1 (a) and F2 (b) in the bilingual groups

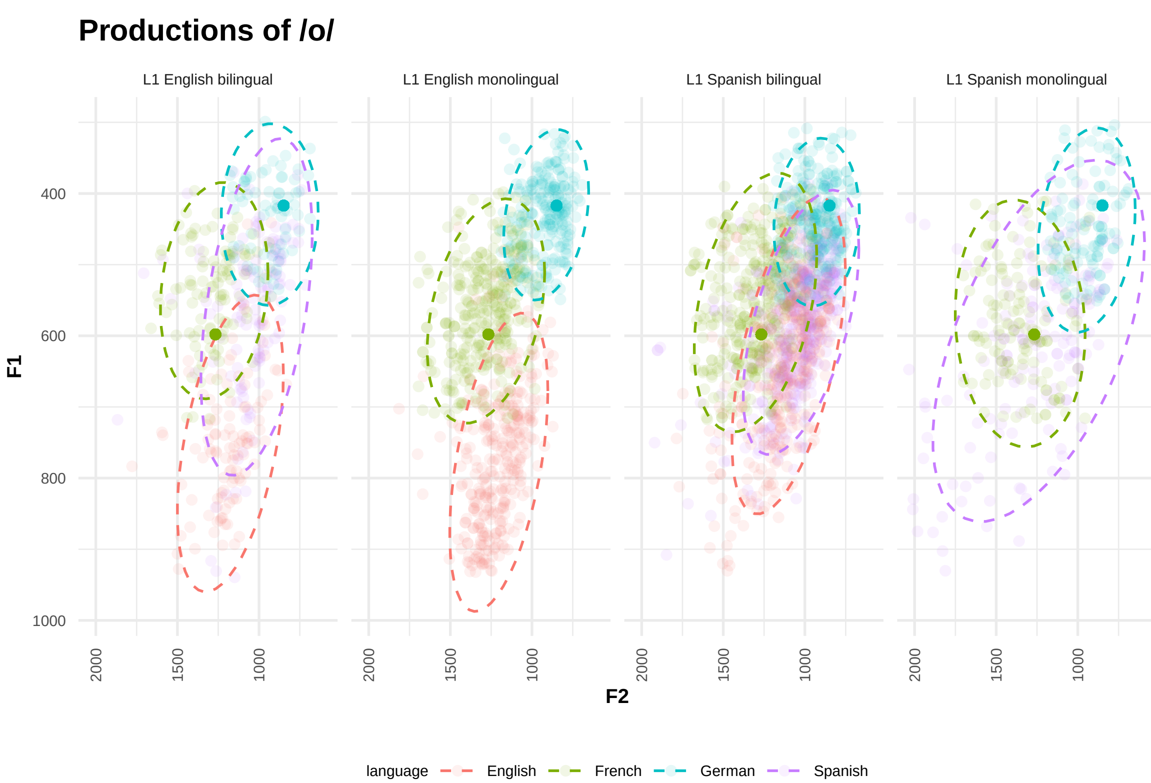


Figure 4.9: Productions of /o/ in each language by each group

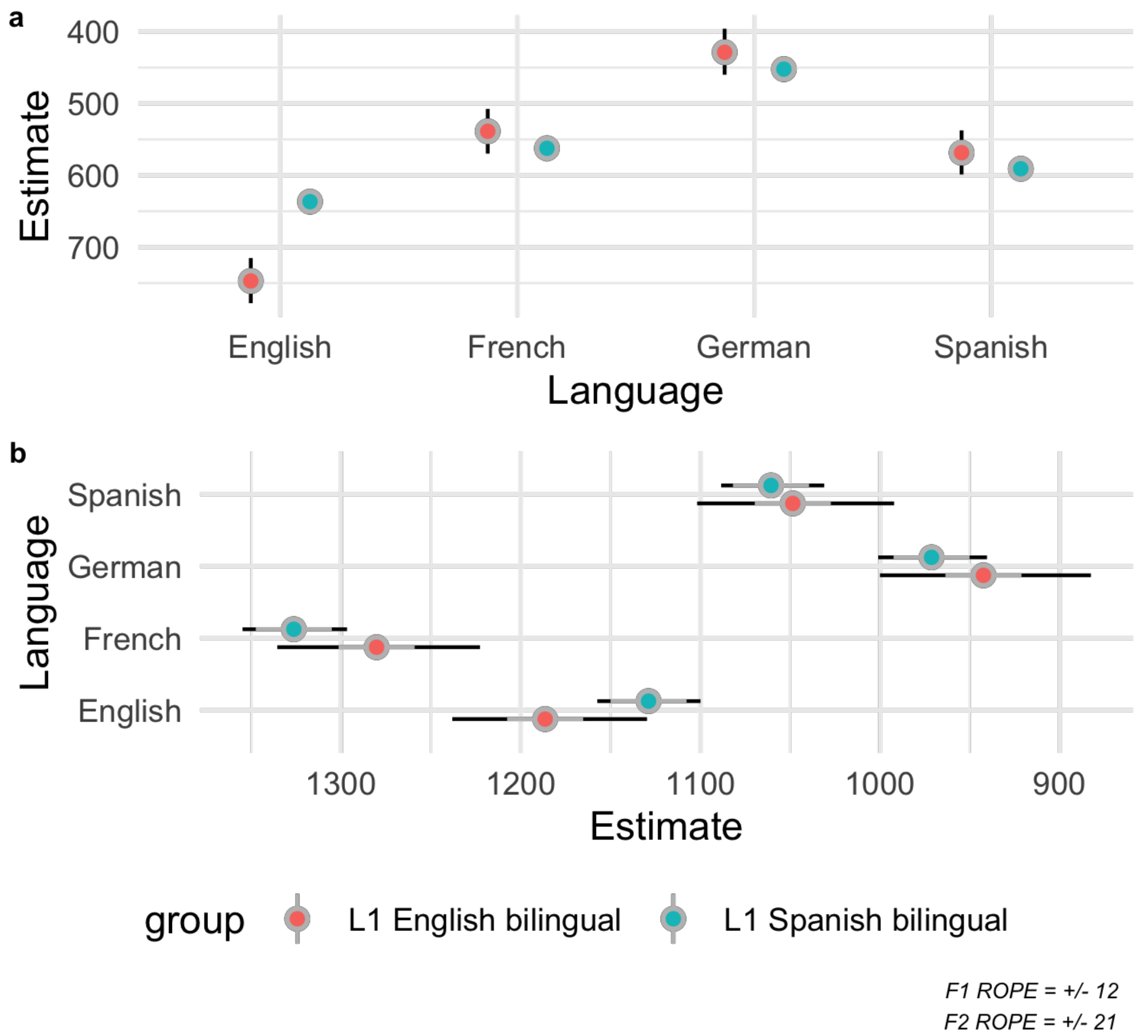


Figure 4.10: Conditional effects plot of /o/ for F1 (a) and F2 (b) in the bilingual groups

analysis, however, revealed that the English L1 group had a more fronted /i/ in German and English than in Spanish and French (4.12, panel a). The Spanish L1 group showed less variation in /i/ fronting as a function of language (4.12, panel b). 4.12 also shows that both /i/ was slightly higher in German and Spanish by both groups.

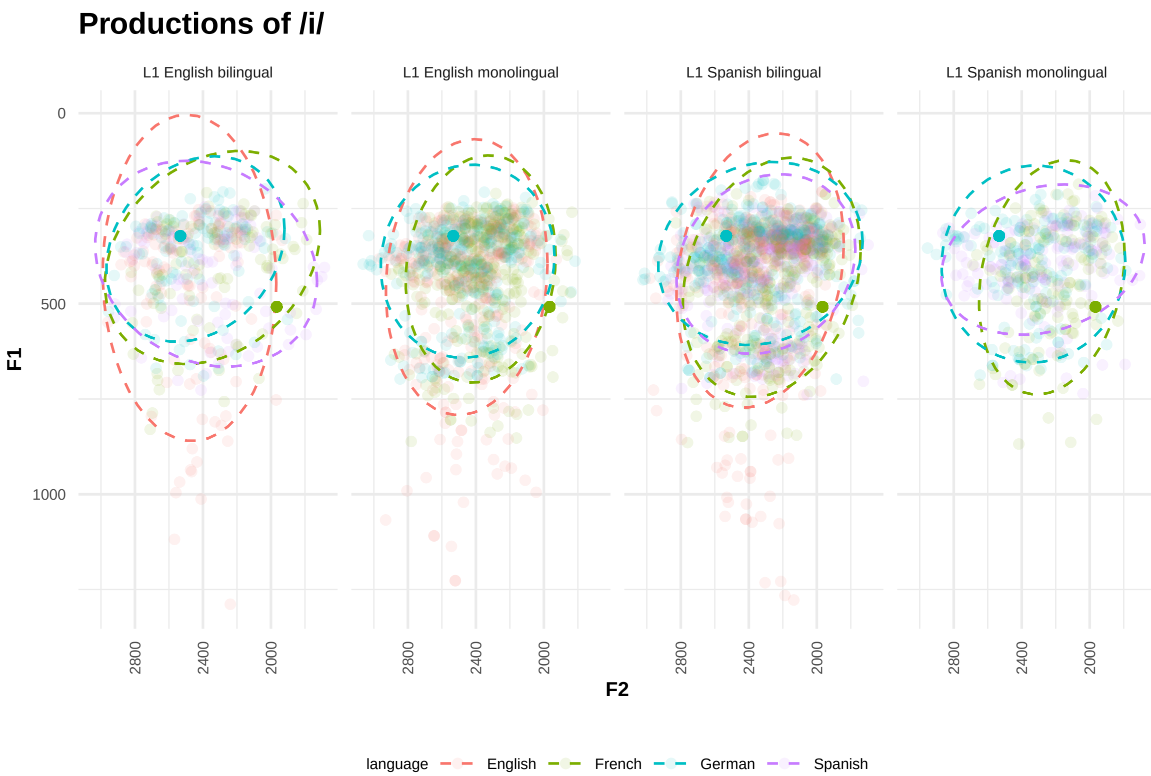


Figure 4.11: Productions of /i/ in each language by each group

However, figure 4.13 shows the production of the wedge in English, French and German and /a/ in Spanish. This condition was intended to bias the production of English-like L3 vowels, with the Spanish /a/ for comparison. The results in this case are not consistent across groups. The L1 English bilingual group appeared to produce the German wedge, the English wedge, and the Spanish /a/ similarly. The L1 Spanish group produced the German and English wedge similarly, but diverged slightly in their production of /a/. All groups produced the French wedge similarly to one another, and dissimilarly from their own productions of other languages. 4.14

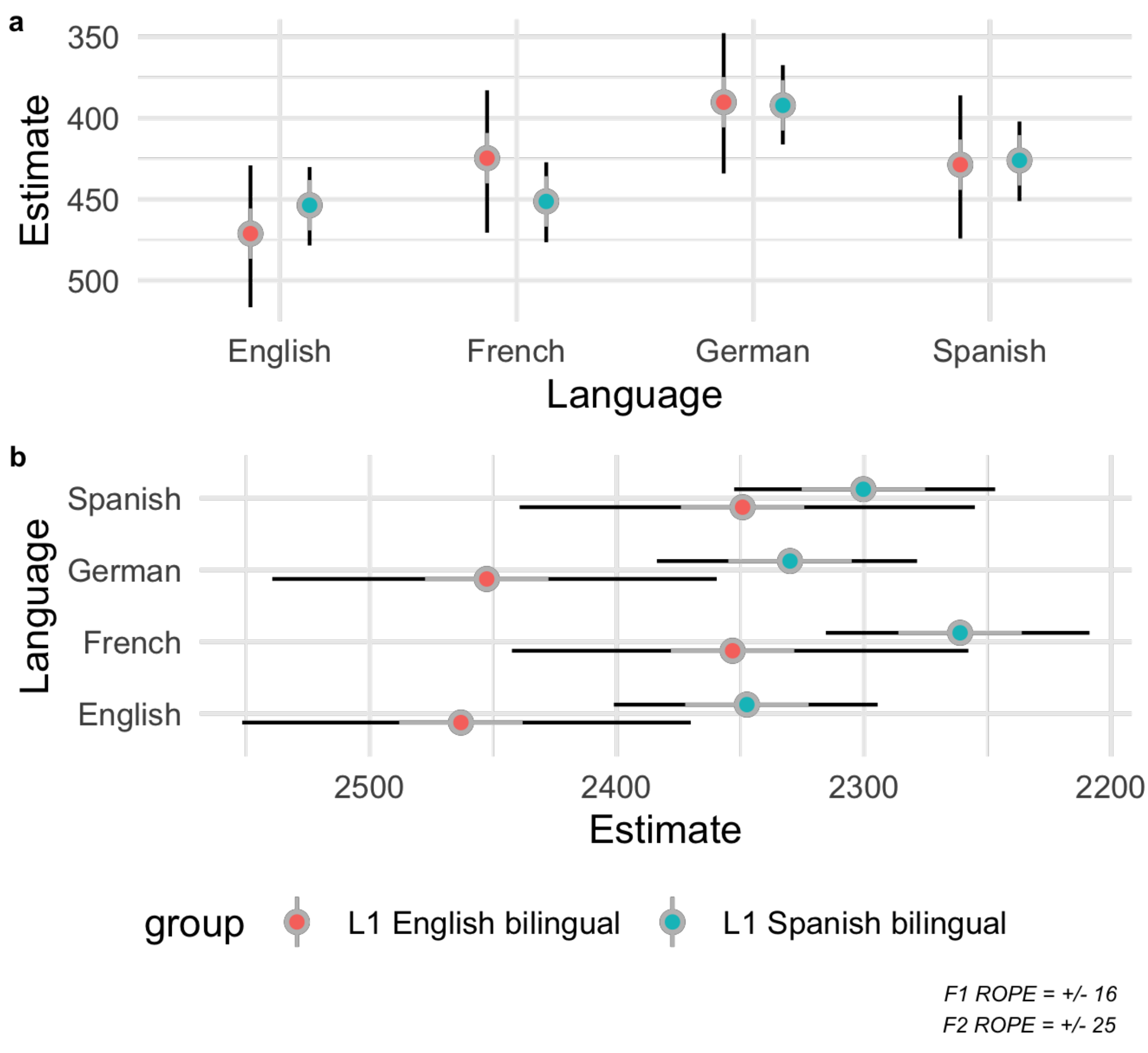


Figure 4.12: Conditional effects plot of /i/ for F1 (a) and F2 (b) in the bilingual groups

shows that both bilingual groups produced vowels similarly in a particular language when they produced the wedge or /a/ (in the case of Spanish).

Figure 4.13: Productions of /a/ and the wedge in each language by each group

4.6.3 Discussion

- When bilinguals encounter a new language sound that is not phonemic in their L1 or L2, at least with the German and French /y/, they appear to have no problems imitating this sound, and do not assimilate the sound to their /u/ or /i/ categories in production.

No clear bilingual advantage was found in the analysis of the data.

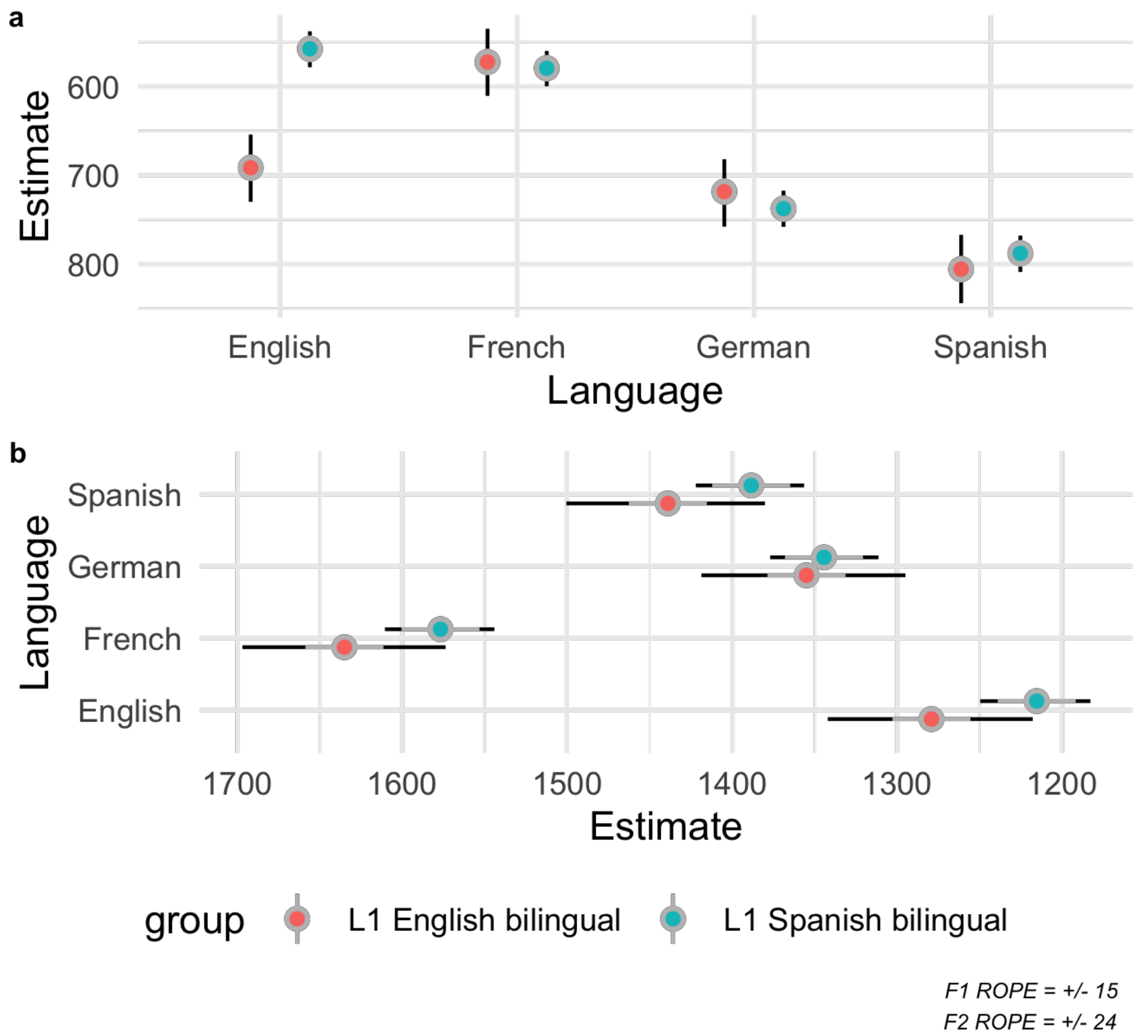


Figure 4.14: Conditional effects plot of the wedge for F1 (a) and F2 (b) in the bilingual groups

Chapter 5: Sourcing .R files

5.1 Import scripts

We can use the following command to import an r script:

```
library("knitr")
read_chunk('../includes/scripts/test.R')
```

Notice that the `read_chunk()` command takes **this** file as the reference for specifying the path (this is different with regard to inserting graphics).

5.2 Call chunks

We can directly call knitr chunks from the `test.R` script. First let's load the libraries we will need.

```
library("dplyr")
```

Now we will generate some data.

```
# Generate data
set.seed(1)
vot  <- rnorm(20, 15, 5)
vot  <- sort(vot)
phon <- c(0,1,0,0,0,0,0,1,0,1,0,1,0,1,1,1,1,1,1,1)
df   <- as.data.frame(cbind(vot, phon))
```

Let's fit a model.

```
# Fit model
glm <- glm(phon ~ vot, data = df, family = "binomial")
```

What is the phoneme boundary?

```
# Get crossover point by hand
co_point <- as.numeric(coef(glm)[1] / (coef(glm)[2] * -1))
co_point
```

```
## [1] 15.53595
```

It looks like the boundary is at 15.5359483. Good. Let's plot it to see what it looks like:

```
# Plot regression with crossover point
plot(df$vot, df$phon, xlab = "vot", ylab = "phon",
     pch = 16, col = rgb(0, 0, 204, 102, maxColorValue = 255))
curve(predict(glm, data.frame(vot = x), type = "resp"), add = TRUE)
points(vot, fitted(glm), pch = 20)
abline(v = co_point, lty = 2, lwd = 0.75)
abline(h = 0.5, v = 0)
```

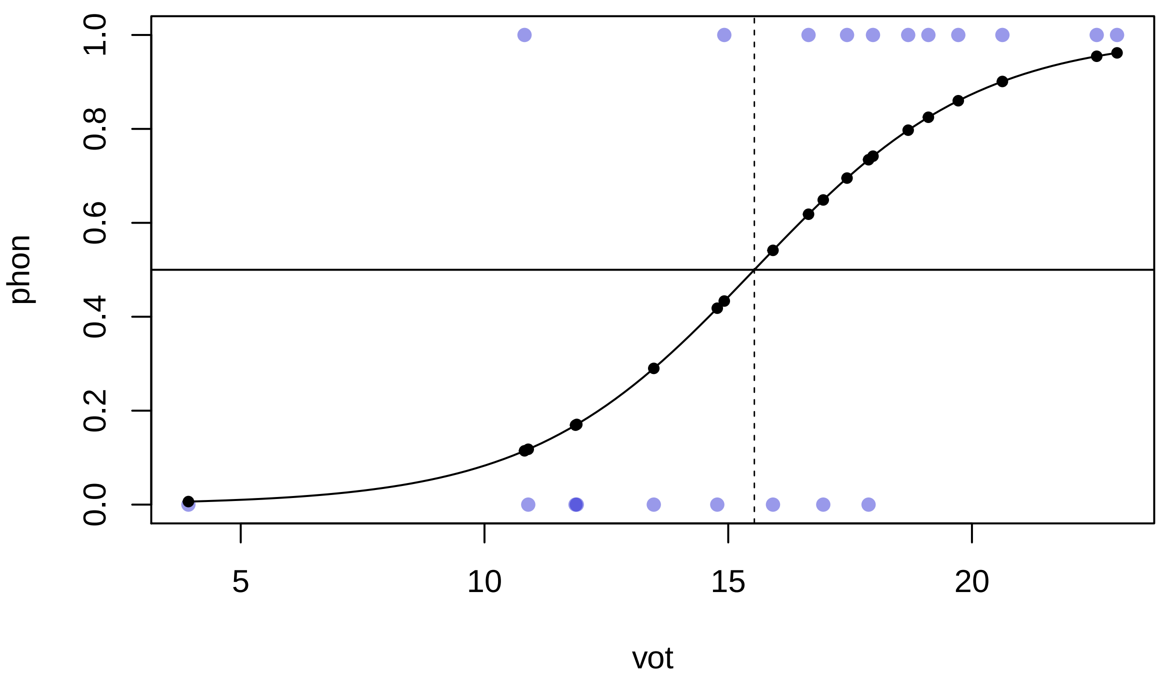


Figure 5.1: This is the caption

Appendix A: Sample Appendix

A.1 Bookdown style captions

A.1.1 Figure captions

Figure A.1 is an example of how we can use bookdown style figure captions.

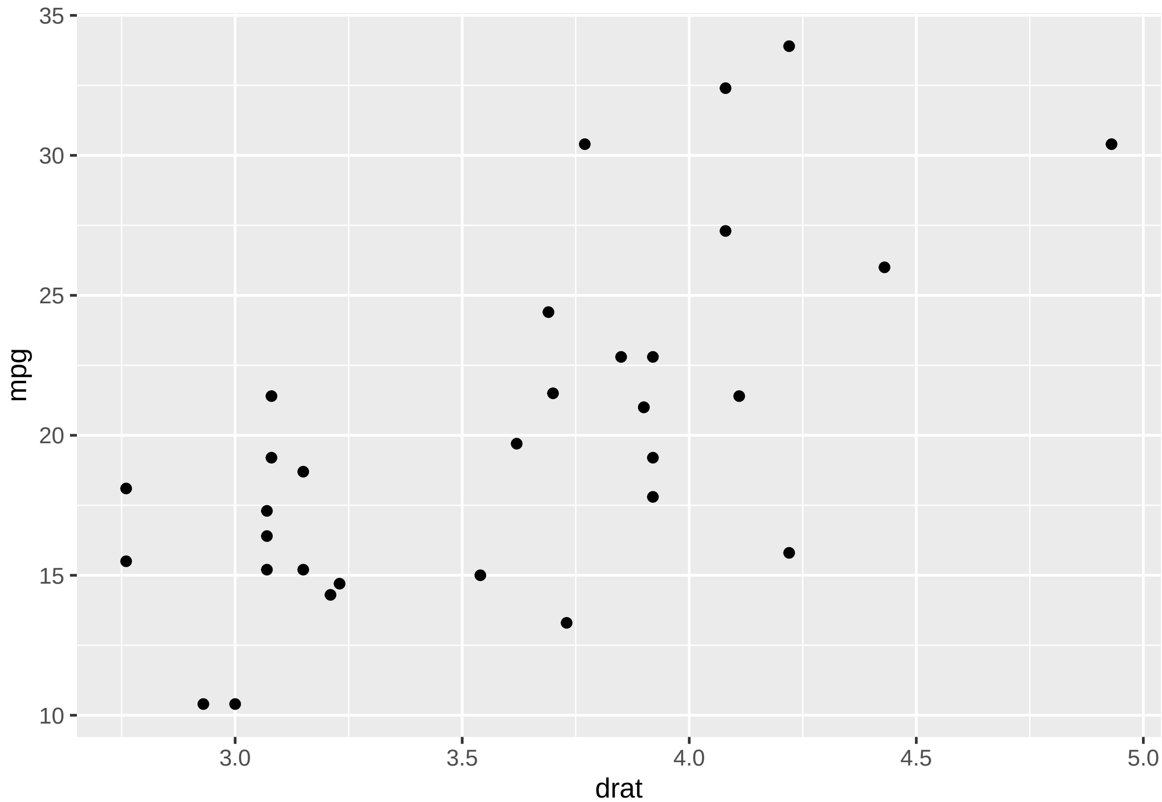


Figure A.1: This is an awesome figure caption.

A.1.2 Table captions

Table A.2 of the supplementary materials provides a numeric summary of the posterior distribution.

Table A.1: Mean and standard deviation of each phoneme in the hand-corrected subset

phoneme	language	mean_f1	sd_f1	mean_f2	sd_f2
a	Spanish	746	130	1487	293
i	English	476	409	2453	268
i	French	472	203	2343	279
i	German	421	140	2419	345
i	Spanish	428	158	2319	318
o	English	681	126	1206	307
o	French	553	83	1306	200
o	German	424	63	940	128
o	Spanish	658	144	1306	366
schwa	English	622	181	1301	286
schwa	French	630	113	1592	248
schwa	German	736	101	1307	164
u	English	437	109	1256	275
u	French	558	265	1920	209
u	German	404	135	1724	290
u	Spanish	542	189	1298	496

Table A.2: This is the table caption.

mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2

mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
18.1	6	225.0	105	2.76	3.460	20.22	1	0	3	1
14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4
24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2
22.8	4	140.8	95	3.92	3.150	22.90	1	0	4	2
19.2	6	167.6	123	3.92	3.440	18.30	1	0	4	4

Appendix B: Another Appendix

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.[1]

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.[2] Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud

---

[1] Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

[2] Check footnote spacing.

exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

REFERENCES

Antoniou, M., Best, C. T., & Tyler, M. D. (2013). Focusing the lens of language experience: Perception of Ma'di stops by Greek and English bilinguals and monolinguals. *The Journal of the Acoustical Society of America*, *133*(4), 2397–2411. http://doi.org/10.1121/1.4792358

Antoniou, M., Liang, E., Ettlinger, M., & Wong, P. C. M. (2015). The bilingual advantage in phonetic learning. *Bilingualism: Language and Cognition*, *18*(4), 683–695. http://doi.org/10.1017/S1366728914000777

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, *32*(2), 233–250. http://doi.org/10.1016/S0095-4470(03)00036-6

Balas, A. (2018). *Non-native vowel perception: The interplay of categories and features = Percepcja samogłosek w języku nienatywnym: Oddziaływanie kategorii i cech* (Wydanie I). Poznań: Wydawnictwo Naukowe UAM.

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

Bardel, C., & Falk, Y. (2021). L1, L2 and L3: Same or different? *Second Language Research*, *37*(3), 459–464. http://doi.org/10.1177/0267658320941033

Bayona, P. (2009). The Acquisition of Spanish Middle and Impersonal Passive Constructions from SLA and TLA Perspectives. In *Third language acquisition and universal grammar*. Bristol, UK; Buffalo: Multilingual Matters.

Berkes, É., & Flynn, S. (2012). Further Evidence in Support of the Cumulative-Enhancement Model: CP Structure Development. In *Third Language Acquisition in Adulthood*. Amsterdam; Philadelphia: John Benjamins.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn

& M. J. Munro (Eds.), *Language Learning & Language Teaching* (Vol. 17, pp. 13–34). Amsterdam: John Benjamins Publishing Company. http://doi.org/10.1075/lllt.17.07bes

Birdsong, D., Gertken, L. M., & Amengual, M. (2012). Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism. University of Texas at Austin: COERLL. Retrieved from %3Chttps://sites.la.utexas.edu/bilingual/%3E.

Blank, C. A., & Zimmer, M. C. (2009). A transferência fonético-fonológica L2 (francês) – L3 (inglês): Um estudo de caso. *Revista de Estudos Da Linguagem*, *17*(1), 207–233. http://doi.org/10.17851/2237-2083.17.1.207-233

Bohn, O.-S., & Flege, J. E. (1992). The Production of New and Similar Vowels by Adult German Learners of English. *Studies in Second Language Acquisition*, *14*(2), 131–158. http://doi.org/10.1017/S0272263100010792

Bohnacker, U. (2006). When Swedes begin to learn German: From V2 to V2. *Second Language Research*, *22*(4), 443–486. http://doi.org/10.1191/0267658306sr275oa

Bradlow, A. R. (1995). A comparative acoustic study of English and Spanish vowels. *The Journal of the Acoustical Society of America*, *97*(3), 1916–1924. http://doi.org/10.1121/1.412064

Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. *The Journal of the Acoustical Society of America*, *117*(6), 3890–3901. http://doi.org/10.1121/1.1906060

Brysbaert, M. (2020). Power considerations in bilingualism research: Time to step up our game. *Bilingualism: Language and Cognition*, 1–6. http://doi.org/10.1017/S1366728920000437

Bürkner, P.-C. (2017). **Brms** : An *r* Package for Bayesian Multilevel Models Using *stan*. *Journal of Statistical Software*, *80*(1). http://doi.org/10.18637/jss.v080.i01

Casillas, J. V., & Simonet, M. (2018). Perceptual categorization and bilingual language modes: Assessing the double phonemic boundary in early and late bilinguals. *Journal of Phonetics*, *71*, 51–64.

http://doi.org/10.1016/j.wocn.2018.07.002

Curtin, S., Goad, H., & Pater, J. V. (1998). Phonological transfer and levels of representation: The perceptual acquisition of Thai voice and aspiration by English and French speakers. *Second Language Research*, *14*(4), 389–405. http://doi.org/10.1191/026765898674095369

Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics*, *34*(2), 269–284. http://doi.org/10.1016/j.wocn.2005.06.002

Escudero, P., & Boersma, P. (2004). BRIDGING THE GAP BETWEEN L2 SPEECH PERCEPTION RESEARCH AND PHONOLOGICAL THEORY. *Studies in Second Language Acquisition*, *26*(04). http://doi.org/10.1017/S0272263104040021

Escudero, P., & Chládková, K. (2010). Spanish listeners' perception of American and Southern British English vowels. *The Journal of the Acoustical Society of America*, *128*(5), EL254–EL260. http://doi.org/10.1121/1.3488794

Escudero, P., Sisinni, B., & Grimaldi, M. (2014). The effect of vowel inventory and acoustic properties in Salento Italian learners of Southern British English vowels. *The Journal of the Acoustical Society of America*, *135*(3), 1577–1584. http://doi.org/10.1121/1.4864477

Escudero, P., & Vasiliev, P. (2011). Cross-language acoustic similarity predicts perceptual assimilation of Canadian English and Canadian French vowels. *The Journal of the Acoustical Society of America*, *130*(5), EL277–EL283. http://doi.org/10.1121/1.3632043

Escudero, P., & Williams, D. (2011). Perceptual assimilation of Dutch vowels by Peruvian Spanish listeners. *The Journal of the Acoustical Society of America*, *129*(1), EL1–EL7. http://doi.org/10.1121/1.3525042

Falk, Y., & Bardel, C. (2011). Object pronouns in German L3 syntax: Evidence for the L2 status factor. *Second Language Research*, *27*(1), 59–82. http://doi.org/10.1177/0267658310386647

Falk, Y., Lindqvist, C., & Bardel, C. (2015). The role of L1 explicit metalinguistic knowledge in L3 oral production at the initial state. *Bilingualism: Language and Cognition*, *18*(2), 227–235. http://doi.org/10.1017/S1366728913000552

Flege, J. (1995). Second language speech learning: Theory, findings, and problems. In *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.

Flege, James Emil. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *The Journal of the Acoustical Society of America*, *89*(1), 395–411. http://doi.org/10.1121/1.400473

Flege, James Emil, Aoyama, K., & Bohn, O.-S. (2021). The Revised Speech Learning Model (SLM-r) Applied. In R. Wayland (Ed.), *Second Language Speech Learning* (1st ed., pp. 84–118). Cambridge University Press. http://doi.org/10.1017/9781108886901.003

Flege, James Emil, Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of english vowels. *Journal of Phonetics*, *25*(4), 437–470.

Flege, James Emil, & Eefting, W. (1988). Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *The Journal of the Acoustical Society of America*, *83*(2), 729–740. http://doi.org/10.1121/1.396115

Flege, J. E., Frieda, E. M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics*, *25*(2), 169–186. http://doi.org/10.1006/jpho.1996.0040

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

García Mayo, M. del P., & Slabakova, R. (2015). Object drop in L3 acquisition. *International Journal of Bilingualism*, *19*(5), 483–498.

http://doi.org/10.1177/1367006914524643

Gelman, A., Simpson, D., & Betancourt, M. (2017). The Prior Can Often Only Be Understood in the Context of the Likelihood. *Entropy*, *19*(10), 555. http://doi.org/10.3390/e19100555

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

Grosjean, F. (1998). Transfer and language mode. *Bilingualism: Language and Cognition*, *1*(3), 175–176.

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

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *The Journal of the Acoustical Society of America*, *118*(5), 3267–3278. http://doi.org/10.1121/1.2062307

Izura, C., Cuetos, F., & Brysbaert, M. (2016). *Lexical Test for Advanced Learners of Spanish*. American Psychological Association. Retrieved from http://doi.apa.org/getdoi.cfm?doi=10.1037/t47086-000

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

Jongman, A., Fourakis, M., & Sereno, J. A. (1989). The Acoustic Vowel Space of Modern Greek and German. *Language and Speech*, *32*(3), 221–248. http://doi.org/10.1177/002383098903200303

Kamiyama, T. (2007). Acquisition of French vowels by Japanese-speaking learners: Close and close-mid rounded vowels. Freiburg, Germany.

Kellerman, E. (1983). Now you see it, now you don't. In *Language Transfer* *in Language Learning. Issues in Second Language Research.* (pp. 112–134). Rowley, MA: Newbury House Publishers.

Kruschke, J. (2018). Rejecting or Accepting Parameter Values in Bayesian Estimation. *Advances in Methods and Practices in Psychological Sciences*, *1*(2), 270–280. http://doi.org/https://doi.org/10.1177/2515245918771304

Lakens, D. (2017). Equivalence Tests: A Practical Primer for *t* Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science*, *8*(4), 355–362. http://doi.org/10.1177/1948550617697177

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343. http://doi.org/10.3758/s13428-011-0146-0

Lemhöfer, K., & Broersma, M. (2016). *Lexical Test for Advanced Learners of English*. American Psychological Association. Retrieved from http://doi.apa.org/getdoi.cfm?doi=10.1037/t47085-000

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

Leussen, J.-W. van, & Escudero, P. (2015). Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in Psychology*, *6*. http://doi.org/10.3389/fpsyg.2015.01000

Lisker, L., & Abramson, A. S. (1964). A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. WORD, *20*(3), 384–422. http://doi.org/10.1080/00437956.1964.11659830

Liu, J., & Lin, J. (2021). A Cross-Linguistic Study of L3 Phonological Acquisition of Stop Contrasts. *SAGE Open*, *11*(1), 215824402098551. http://doi.org/10.1177/2158244020985510

Liu, Z., Gorba, C., & Cebrian, J. (2019). EFFECTS OF LEARNING AN ADDITIONAL LANGUAGE ON VOT PERCEPTION. In *19th International Congress of Phonetic Sciences*. Melbourne, Australia: Australasian Speech

Science; Technology Association Inc.

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

Mitrofanova, N., & Westergaard, M. (2018). Wholesale vs. Property-by-property transfer: Acquisition of morphological case in an artificial L3. Braunschweig.

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

Onishi, H. (2016). The effects of L2 experience on L3 perception. *International Journal of Multilingualism*, *13*(4), 459–475. http://doi.org/10.1080/14790718.2016.1217604

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

Parrish, K. (2021). The relative VOT productions of L3 French by Spanish-English bilinguals at first exposure. In *Multiple language acquisition, processing, and development*. Poznań, Poland: Adam Mickiewicz University.

Patihis, L., Oh, J. S., & Mogilner, T. (2015). Phoneme discrimination of an unrelated language: Evidence for a narrow transfer but not a broad-based bilingual advantage. *International Journal of Bilingualism*, *19*(1), 3–16.

http://doi.org/10.1177/1367006913476768

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. http://doi.org/10.3758/s13428-018-01193-y

Plonsky, L. (Ed.). (2015). *Advancing quantitative methods in second language research*. New York: Routledge.

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

Slabakova, R. (2017). The scalpel model of third language acquisition. *International Journal of Bilingualism*, *21*(6), 651–665. http://doi.org/10.1177/1367006916655413

Slabakova, R., & García Mayo, M. D. P. (2015). The L3 syntax–discourse interface. *Bilingualism: Language and Cognition*, *18*(2), 208–226. http://doi.org/10.1017/S1366728913000369

Stadt, R., Hulk, A., & Sleeman, P. (2016). The influence of L2 English and immersion education on L3 French in the Netherlands. *Linguistics in the Netherlands*, *33*, 152–165. http://doi.org/10.1075/avt.33.11sta

Stadt, R., Hulk, A., & Sleeman, P. (2018a). The influence of L1 Dutch and L2 English on L3 French: A longitudinal study. *Journal of the European Second Language Association*, *2*(1), 63. http://doi.org/10.22599/jesla.42

Stadt, R., Hulk, A., & Sleeman, P. (2018b). The role of L2 exposure in L3A: A comparative study between third and fourth year secondary school students in the Netherlands. In J. Berns, H. Jacobs, & D. Nouveau (Eds.), *Romance Languages and Linguistic Theory* (Vol. 13, pp. 279–296). Amsterdam: John Benjamins Publishing Company. http://doi.org/10.1075/rllt.13.17sta

Stadt, R., Hulk, A., & Sleeman, P. (2020). L2 Influence in L3 Acquisition: The Role of the L3. In A. Trotzke & T. Kupisch (Eds.), *Formal Linguistics and Language Education* (Vol. 43, pp. 203–221). Cham: Springer International Publishing. http://doi.org/10.1007/978-3-030-39257-4_11

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*(1), 1–25. http://doi.org/10.1016/S0749-596X(03)00105-0

Werker, J. F. (1986). The effect of multilingualism on phonetic perceptual flexibility. *Applied Psycholinguistics*, *7*(2), 141–155. http://doi.org/10.1017/S0142716400007360

Westergaard, M. (2021). Microvariation in multilingual situations: The importance of property-by-property acquisition. *Second Language Research*, *37*(3), 379–407. http://doi.org/10.1177/0267658319884116

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. (2012). Foreign accentedness in third language acquisition: The case of L3 English. In J. Cabrelli Amaro, S. Flynn, & J. Rothman (Eds.), *Studies in Bilingualism* (Vol. 46, pp. 281–310). Amsterdam: John Benjamins Publishing Company. http://doi.org/10.1075/sibil.46.16wre

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. (2021). Multilingual acquisition property by property: A view from a wider perspective. *Second Language Research*, *37*(3), 441–447. http://doi.org/10.1177/0267658320935167

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

Wrembel, M., Marecka, M., & Kopečková, R. (2019). Extending perceptual assimilation model to L3 phonological acquisition. *International Journal of Multilingualism*, *16*(4), 513–533. http://doi.org/10.1080/14790718.2019.1583233

Wunder, E.-M. (2010). Phonological cross-linguistic influence in third or additional language acquisition. In *Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech* (pp. 566–571). Poznań, Poland: Adam Mickiewicz University.