October 2022

Examining variability in Spanish monolingual and bilingual phonotactics: A look at sC-clusters

Katerina A. Tetzloff
University of Massachusetts Amherst

Recommended Citation

Tetzloff, Katerina A., "Examining variability in Spanish monolingual and bilingual phonotactics: A look at sC-clusters" (2022). *Doctoral Dissertations*. 2674.
https://doi.org/10.7275/30944557 https://scholarworks.umass.edu/dissertations_2/2674

EXAMINING VARIABILITY IN SPANISH MONOLINGUAL AND BILINGUAL PHONOTACTICS: A LOOK AT SC-CLUSTERS

A Dissertation Presented

by

KATERINA A. TETZLOFF

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2022

Hispanic Literatures and Linguistics

Dept. of Languages, Literatures, and Cultures

EXAMINING VARIABILITY IN SPANISH MONOLINGUAL AND BILINGUAL PHONOTACTICS: A LOOK AT SC-CLUSTERS

A Dissertation Presented

by

KATERINA A. TETZLOFF

Approved as to style and content by:

Danielle Thomas, Chair

Luiz Amaral, Member

Joe Pater, Member

Barbara Zecchi, Department Head
Spanish and Portuguese Studies

Maria Soledad Barbón, Department Chair
Dept. of Languages, Literatures, and Cultures

DEDICATION

To my family, who continuously supports me in everything that I do.

ACKNOWLEDGMENTS

I would first like to acknowledge the National Science Foundation Graduate Research Fellowship Program (fellowship #2017240584) that provided me with financial support for three years during my doctoral studies. This grant was such a blessing to me, as it allowed me to devote more time to my research and to staying sane along the way.

This dissertation would not have been possible without the guidance and support of many people. First of all, I am forever grateful to my advisor, Dr. Danielle Thomas, who took me under her wing nearly ten years ago. If you had not encouraged and supported the eager undergrad me in doing research, I would not be where I am today. Thank you for guiding me through so many important steps in my career both at York University and at UMass. I love that you always have the most creative questions when reading papers and have taught me to do the same, which has made me a better researcher. I am also very thankful for your support and open-mindedness about my career and letting me explore the paths that are best for me, even if they aren't the expected ones. Without you, I would not have made my way to UMass, and even if I had, I don't think that I would have been able to finish. On a personal note, I will never forget our meetings in your office with Kristen DonPaul, or getting lost walking on the highway in France after our birthday celebration. I am so thankful to have you not only as an advisor, but also as a lifelong friend.

I would also like to thank Dr. Joe Pater for helping me through every milestone of this degree. I am still in awe of your brilliance and expertise in phonology and am so lucky to have learned the vast majority of my knowledge in this field from you. I am thankful for how you always supported me in your advising and teaching and for how you always made me feel like I am more than just my work. I hope to be able to have the same impact on my future students that you have had on me.

Another "thank you" goes to my third committee member, Dr. Luiz Amaral who has also supported me in this dissertation.

I am grateful to my other professors at UMass, in both the Spanish and Linguistics departments, who have helped me substantially during my graduate studies, and in particular Dr. Gaja Jarosz and Dr. Kristine Yu. I also thank my classmates from both departments who have struggled and celebrated with me on this journey!

Many more people have contributed to this dissertation, whom I am also grateful for, including Daniela Narvaez, Sara Perez, and Oriana Vile who assisted in data collection. I am also grateful to my medical doctors who kept me healthy through three surgeries during my graduate studies and to my psychiatrist, Dr. DeWitt, for helping me navigate all of the stress involved in getting a Ph.D. and in life in general.

An especially big thanks goes to my family for all of their support. To my husband, Jonathan: thank you so much for supporting me every day since we met and for understanding that sometimes the dishes couldn't get done because I needed to finish a chapter or run another analysis. I can't imagine having finished this dissertation without you, because you kept me calm when I wanted to panic more times than I can count. I love you! To my kitty, Taco: I could not have finished this degree without your help – thank you for sitting on my lap (or my keyboard) as I pushed through this dissertation and for always helping calm me down when I got overwhelmed. To my parents: thank you for supporting me throughout my whole life, and particularly during these last six years. Even though you didn't have to get up at 5am to drive me to the cold ice rink every day during my graduate studies, you still did more than your fair share of helping me through it all. You have both set such a good example of how to work hard, which taught me the skills I needed to get this degree. To my sister, Stefanie: thank you for tolerating me enough to live in our spare bedroom and letting me vent to you and bother you whenever I needed to. You're not as bad as I thought you were when you were first born. And to my Oma: I miss you everyday, but I am forever grateful for the love and support you continuously showed me while working on this dissertation. It makes me happy to know how proud you would be of me for finishing this degree.

Above all, I am most thankful to God for giving me the opportunity and guidance to achieve my academic goals and write this dissertation, and for bringing all of these wonderful people into my life who have made this journey so much better.

ABSTRACT

EXAMINING VARIABILITY IN SPANISH MONOLINGUAL AND BILINGUAL PHONOTACTICS: A LOOK AT SC-CLUSTERS

SEPTEMBER 2022

KATERINA A. TETZLOFF
B.A., YORK UNIVERSITY
M.A., YORK UNIVERSITY
M.A., UNIVERSITY OF MASSACHUSETTS AMHERST
Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Danielle Thomas

Current models of generative phonology have failed to address the variability that is observed in bilingual language patterns patterns. This dissertation addresses exactly that issue by examining the perception of Spanish sC-clusters in Spanish monolinguals and English-Spanish bilinguals.

Surface sC-clusters in onset position are prohibited in Spanish and are repaired by inserting a prothetic /e/ (sC → esC). English differs in that it allows sC-cluster onsets, and the structure of the sC-cluster has been shown to differ based on the sonority profile (i.e., s+stop clusters are bisyllabic, s+liquid clusters are tautosyllabic). A batch version of a Harmonic Grammar Gradual Learning Algorithm (HG-GLA) was given Spanish input and predicted that Spanish sC-clusters may be syllabified differently based on the sonority of the sC-cluster. It predicted that s+stop clusters are more likely to instantiate /e/ prothesis than s+liqud clusters, but that s+liquid clusters are most likely to be syllabified as a true branching onset like in English. This led to the hypothesis that s+stop and s+liquid clusters may show observable differences in perception in Spanish.

Furthermore, studies in bilingualism have shown strong evidence for bilingual variability, or non-monolingual-like language behavior, particularly in areas where there is non-identical structural overlap, as is the case with sC-clusters in Spanish and English. The perception of s+stop and s+liquid clusters was thus also analyzed with respect to the following language-external variables that affect bilingual variability: language profile (monolingual versus bilingual), age of exposure to bilingualism, and bilingual dominance.

To test these hypotheses, two experiments were performed. The first was a replication of an AX task that has been shown to exhibit variability in Spanish sC-cluster perception in past studies. In this task, native Spanish speakers (monolingual and bilingual) listened to stimuli pairs that differed in the duration and quality of the initial vowel preceding the sC-cluster and were asked to respond if they were the same or different. The second was a nonce word judgment task where participants were presented with Spanish-like nonce words beginning with sC-clusters and had to give them acceptability ratings of how 'Spanish-like' they sounded.

The results did not show evidence of a language-internal effect. s+stop and s+liquid clusters were treated the same in perception by Spanish native speakers, contrary to the predictions of the HG-GLA. Regarding the language-external variables, there was a strong effect of language profile on perception of sC-clusters in Spanish: monolinguals showed a strong dis-preference for sC-initial words, whereas bilinguals were more accepting of such clusters. However, the bilingual variability observed was not affected by age of exposure to bilingualism or by language dominance.

Finally, a sketch of a proposal is made for how generative theories of phonology, like Harmonic Grammar, could potentially be adapted to accommodate the observed differences between the phonotactics of monolinguals and bilinguals, particularly for the case of sC-clusters in English-Spanish bilinguals.

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The goal of phonology is to characterize speakers' knowledge of sound alternations and possible sound patterns that can be used to make up a word within a language's grammar. Phonotactics is a sub-field of phonology that addresses the latter goal: understanding possible and impossible sound sequences of a language. More specifically, phonotactics is concerned with which sequences of sounds are permissible together and also which sounds or sequences of sounds are allowed to occupy a specific slot in the syllable (i.e., onset and coda positions). Different languages have different phonotactic restrictions. For example, many languages allow sequences of two or more sounds to appear at the beginning of a syllable in an onset cluster. English allows onset clusters like /pl/ 'place', /dr/ 'draw', and /sm/ 'smell', among others, yet it disallows onset clusters like /bd/ and /lb/. Because of this, English speakers can make judgments on what is possible or impossible in their language. For example, 'blick' is not a word of English, but a native speaker intuitively knows that this could be or "sounds like" a possible word of English; 'bnick' and 'lbick' are not words of English either, yet, unlike 'blick', a native speaker has similar intuitions that these could never be or "don't sound like" possible words of English. Languages differ in this respect. Some languages, like Hawaiian, never allow more than one consonant to appear in syllable-initial position, while other languages, such as Polish, are much less restrictive in the phonotactic rules that govern onset clusters, allowing sequences such as /bd/, /nd/, and /rd/.

A significant amount of attention has been given to characterizing the phonotactics of onset clusters across languages in the field of linguistics, both through theoretical proposals and experimental investigations. However, most of these studies only aim to account for monolingual speakers, despite the fact that the vast majority of the world's population speaks multiple languages. Generative theories of grammar aim to capture how it is that speakers of different languages uniformly acquire such complex linguistic systems that they can generalize to novel forms, giving them the knowledge of what is and is not permissible in their language, and all from imperfect and incomplete input. This is also the case in onset cluster phonotactics: speakers of any language are able to acquire the systematic knowledge of what sounds can appear together at the beginning of words or syllables, as well as the knowledge of what sounds cannot pattern together in this position. Although there is a substantial amount of work examining speakers' phonological competence, including their phonotactic grammars, much less work has investigated if and how this grammatical knowledge differs in bilingual speakers.

Non-monolingual-like language behavior in bilinguals has been referred to in a variety of ways in previous studies, using terms such as *language transfer*, *cross-language influence*, *cross-language interference*, *cross-language interaction*, etc. Some of these terms are systematically determined depending on the type of non-monolingual-like behavior observed, but in other cases they are used as descriptive terms that are used to refer to any type of non-monolingual-like behavior. In the current study, the term '*bilingual variability*' will be used as the descriptive term that indexes this non-monolingual-like language behavior.

In the domain of morphosyntax, it is widely accepted that bilinguals are not two monolinguals in one, meaning that they do not have the same language trajectories and outcomes as monolinguals in each of their languages (Deuchar & Quay, 1998; Müller, 1998; Hulk & Müller, 2000; Cenoz & Genesee, 2001; Serratrice, Sorace, & Paoli, 2004; Bialystok, Luk, Peets, & Yang, 2010; y Cabo, 2015; among many others). Studies have shown that the variability, or grammatical optionality, of bilinguals differs from that of monolinguals. However, this bilingual variability is not random: it does not occur in all grammatical domains and it does not occur to the same degree for all grammatical domains across different bilingual language pairings (i.e., English-Spanish bilingual variability is distinct from both English monolingual and Spanish monolingual variability, as well as English-Japanese bilingual variability, for example). Some of these studies in morphosyntax have sought to examine the impact of language internal factors, including the type of grammatical structure or characteristics of the structures, on this variability (Liceras, 1985; Pérez-Leroux & Glass, 1999; Sorace & Serratrice, 2009), while others have focused more on the effect that language external factors, such as age of first exposure to bilingualism or relative language proficiency, have on this bilingual variability (Montrul, 2006, 2010a; Sorace & Serratrice, 2009; Meisel, 2010).

The nature of bilingual variability in the phonological domain is much less studied than in the morphosyntactic domain. Within phonology, there is evidence that bilingual language acquisition and outcomes differ from monolinguals' language patterns (Paradis, 2001; Sebastián-Gallés & Bosch, 2002, 2005; Sebastian-Galles & Kroll, 2003; Frisch & Brea-Spahn, 2010; Freeman, Blumenfeld, & Marian, 2016; Carlson, Goldrick, Blasingame, & Fink, 2016; Carlson, 2019), but the nature of the interaction of the two phonological systems is not well understood. There is some evidence that language internal factors such as the internal sound structure have some effect on this variability, at least in second language (L2) learners or late bilinguals (Brown, 2000; Flege, 2007), as do various language external factors (Flege, Munro, & MacKay, 1995; Flege, Yeni-Komshian, & Liu, 1999; Carlson, 2019).

The present dissertation aims to contribute both to areas of generative phonological theory and to models of bilingual variability in phonology. The main goal is to investigate the behavior of native Spanish speakers, both monolingual and English-Spanish bilingual, with respect to what is canonically referred to as onset cluster phonotactics by providing experimental data from discrimination and judgment tasks that can shed light on the nature of measurable variability in the phonological domain. Furthermore, the present study aims to test the role that both language-internal and language-external factors have on bilingual variability in the phonological domain. These results can then be used as groundwork to refine generative models of phonology to accommodate bilingual language behavior and contribute to a better understanding of the nature of selective variability in bilingualism that has been more widely studied for other linguistic domains (Bullock & Toribio, 2004; Kim, 2011; Sorace, 2011; Hartsuiker, Pickering, & Veltkamp, 2004).

English-Spanish bilinguals are an ideal test population for this study because the two languages have different onset cluster phonotactics. Furthermore, in North America there is a unique situation of widespread English-Spanish bilingualism with a robust population of speakers of who have different ages of exposure to both languages and proficiency levels. As such, looking at the phonotactic behaviors of English-Spanish bilinguals is contextually relevant in this society because these bilinguals represent around 20% of the population, and there are significant clinical and educational concerns to which the results of this study could be relevant.

CHAPTER 2

PHONOTACTICS-BASED PREDICTIONS OF SPANISH SC-CLUSTER STRUCTURE

This chapter provides an overview of current generative models of phonology and a summary of sC-cluster phonotactics in Spanish and English. Then a phonological learning algorithm is used to make predictions about the structure of sC-clusters in Spanish and where variability might be present.

2.1 Language universals and generative grammar

The central goal of linguistic theory is to understand what is common to all languages: typology and acquisition. Typology refers to linguistic universals that constrain all languages, which reveals what is and what is not possible for human languages. For example, every language allows syllables of the shape consonant-vowel (CV), but not all languages permit CVC syllables. Language universals and markedness have also been researched through the lens of language acquisition. Children across a variety of languages show the same developmental patterns for most linguistic domains regardless of the language being acquired. One example is the preference for unmarked forms across languages and in language acquisition. In phonology, children across languages tend to simplify complex syllables to less-marked syllables during the acquisition process (e.g., CCV → CV or CVC → CV).

Generative theories of grammar account for these similarities across languages and in acquisition by assuming the existence of Universal Grammar (UG). UG is understood to be the cognitive starting point for the representation and development of language processes that constrains the grammatical properties (Chomsky, 1965). Because of this, the analysis of any particular language should shed light on the nature of UG and how it constrains not only that language but also others.

2.2 Optimality Theory and Harmonic Grammar as a generative model of phonology and phonotactics

Early forms of generative grammar within phonology relied on processes of transformaitonal rules and their specific ordering (Chomsky, 1965; Chomsky & Halle, 1968). These ordered rules were applied to underlying representations (URs) to yield surface representations (SRs), which is the observed form. For example, in English vowels preceding nasal coda consonants become nasalized; this would be represented as /V/ → [Ṽ] / __ C[+nasal]$ (vowels become [+nasal] in the context of preceding a coda that is a nasal consonant).

Within rule-based systems there are not strict limits on what rules are possible or impossible. In other words, rule-based systems do not allow one to straightforwardly make predictions about what is possible or impossible across languages, and it is difficult to explain why some language patterns and tendencies are extremely common across languages and while others do not exist (i.e., they do not capture linguistic markedness) beyond what is more natural phonetically. Furthermore, rule-based systems make weak predictions about phonological acquisition paths.

Such language universals and trends in acquisition can be straightforwardly captured in an Optimality Theory (OT) framework (A. Prince & Smolensky, 1993/2004). OT is a generative linguistic theory of universal constraints and their interaction; crucially, these constraints are violable. OT has been traditionally used as a model of phonology but has also been applied to other linguistic domains (Bresnan, 2000; Blutner, 2000; Legendre, 2001; Blutner, Bezuidenhout, Breheny, Glucksberg, & Happé, 2003; among others).

2.2.1 Basic architecture of OT

In OT, for any given input, an infinite number of output candidates are generated by GEN. These output candidates are subject to the violable, universal constraint set (CON), which are ranked in a language specific way. All of the output candidates created by GEN are subject to constraint set and its ranking by the function H-EVAL, which evaluates the harmony of, or violations incurred by, each candidate, and ultimately selects the winning output candidate. H-EVAL selects the winning candidate by starting with the highest ranked constraint and selects only the subset of output candidates that satisfy this constraint. This is repeated with each constraint in the order that it is ranked until there is only one output left, which is then optimal candidate. If, for example, all outputs in a subset violate a given constraint, none of them are extracted from the subset and all are then subject to the next highest ranked constraint. The overall structure of OT can be seen in Figure 2.1, adapted from Smolensky (1995).

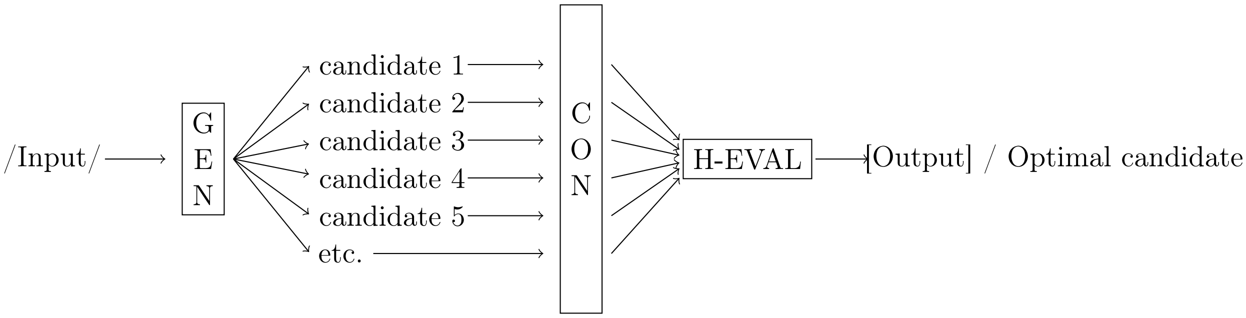


Figure 2.1: Basic architecture of OT. From any given input, GEN creates an infinite number of output forms, which are subject to the language-specific ranking of the set of universal constraints. Each output candidate is evaluated by a function, H-EVAL, based on which constraints they satisfy and violate. The candidate that least violates the highest ranked constraints is the output, or the optimal, winning candidate.

2.2.1.1 Universal markedness constraints

As previously discussed, previous rule-based models of generative phonology had no way of incorporating universal markedness into the structure of the grammar. However, universal markedness is built into the basic architecture of OT through the set of universal constraints. These constraints are satisfied by candidates that have unmarked structures and are violated by candidates that have marked structures. For example, in the Tableau in (1), there is a universal markedness constraint (C) that imposes the restriction of complex onsets (i.e., disprefers consonant clusters). Candidate (a) has the shape CV and thus satisfies this constraint, while candidate (b) has the shape CCV and violates it. Because candidate b violates this constraint, it is marked with an asterisk to denote this violation; the exclamation point signifies that the candidate fatally violates the constraint so that it is no longer a possible output candidate. As candidate a does not incur this fatal violation, it is selected as the optimal output candidate. As this tableau is oversimplified, it should be noted that if it were the case that both candidates in (1) violated the constraint, but candidate b violated it more than candidate a, candidate a would remain the winner. It is in this way that the universal markedness constraints given in an OT framework promote universally unmarked structures in languages.

(1)

	C
☞ a. CV	
b. CCV	∗!

Furthermore, the notion of universal markedness is implicational. Jarosz (2010) defines implicational markedness as follows. If there are two surface structures A and B, A is more marked than B if and only if the following two statements are true: 1) all languages that allow A also allow B, and 2) no language exists that allows B but not A. Understanding markedness as implicational then lends itself naturally to the notion of cross-linguistic frequency, as more languages will permit A than B because B is dependent on the presence of A but the reverse is not true. For example, some languages allow both CCV and CV syllables. Because CCV is more marked than CV, any language that permits CCV syllables must also permit CV syllables, but the reverse is not true. Because of this implicational markedness, no language can allow CCV but not CV syllables.

2.2.1.2 Faithfulness constraints

The universal constraint set also includes faithfulness constraints, which are violated whenever the output candidate forms differ from the input form. Faithfulness constraints interact with markedness constraints, which allows languages to have unmarked patterns. This is demonstrated in the tableaux in (2)-(3); M represents the markedness constraint, and F, the faithfulness constraint. In the tableau in (2) represents a language that does not allow onset clusters. The input is CCV, and the two competing output candidates are (a) CCV, which is the candidate that is faithful to the input representation and (b) CV, which is the unmarked candidate. Because the markedness constraint outranks the faithfulness constraint, candidate (b) is the winner, even though it differs from the input form, violating the faithfulness constraint. However, these constraints interact differently in the tableau in (3), which represents a language that does allow onset clusters. Here, the faithfulness constraint dominates the markedness constraint, such that the penalty for violating the markedness constraint (candidate a) is less than that of violating the faithfulness constraint (candidate b), so the optimal candidate is candidate (a). These examples show how different rankings of the constraints can result in different language patterns, even though the universal markedness pressures are always present and active.

(2)

/CCV/	M	F
a. CCV	*!	
☞ b. CV		*

(3)

/CCV/	F	M
☞ a. CCV		*
b. CV	*!	

2.2.2 Harmonic Grammar

Harmonic Grammar (henceforth, HG) (Legendre, Miyata, & Smolensky, 1990/2006; Legendre, Sorace, Smolensky, et al., 2006) is a similar phonological framework of constraint interaction but instead of ranking the constraints, the system assigns language-specific numerical weights to each constraint according to the language-specific input. The weights are used to scale the importance of constraints: for example, a constraint with a weight of 10 is “more important” to the grammar than a constraint with a weight of two. Output candidates are then evaluated based on these weighted constraints and given a harmony score. The harmony scores are calculated by multiplying the number of constraint violations of a given candidate for any given constraint by the weight assigned to constraint, and then the number of violations times the constraint weights are summed across all constraints to yield the harmony score. The candidate with the highest harmony score (i.e., closest to zero) is the most harmonic, or most optimal candidate.

An example HG tableau can be seen in (4). The generic markedness constraint has a weight of 1.5, while the generic faithfulness constraint has a weight of 1. Candidate (a) violates the faithfulness constraint once because the output candidate differs from the input in one way (a consonant is deleted). This yields a harmony score of -1. Candidate (b) violates the markedness constraint once because it does not have the unmarked syllable shape of CV but it satisfies the faithfulness constraint as the output is identical to the input, resulting in a harmony score of -1.5. Finally, candidate (c) is the least harmonic because it violates both the markedness and faithfulness constraints: the output candidate does not conform to the unmarked syllable shape of CV nor is it faithful to the input syllable shape. The violations of each constraint are multiplied by the constraint weight and summed for a harmony score of -2.5. Thus, candidate (a) is the most harmonic because its harmony score is the highest, making it the optimal candidate from this set.

(4)

	M	F	H
/CCV/	1.5	1	
☞ a. CV		-1	-1
b. CCV	-1		-1.5
c. CVC	-1	-1	-2.5

In some cases, HG is preferred as a model of grammar over traditional OT because harmony scores lend themselves to gradience in well-formedness (Coetzee, 2008). For example, in the tableau above int (4), although candidate (b) is not the optimal candidate, it is more harmonic than candidate (c), and can thus be viewed as less unacceptable than its counterpart with a lower harmony score.

2.3 Phonological learning algorithms for OT and HG grammars

An important component to phonology is its learnability. The learnability problem in language acquisition has been the focus of much research for decades (Wexler & Manzini, 1987; Tesar & Smolensky, 1998, 2000; Crain & Thornton, 2000; Hayes & Wilson, 2008; Pinker, 2009; Archibald, 2014; among many others). The learnability problem is derived from the fact that the vast majority of language input that children acquiring language receive is positive evidence. In other words, children are generally only exposed to what is grammatical or correct in their language and are not explicitly told what is incorrect (i.e., they do not receive negative evidence). This problem is accounted for by the notion of Chomsky's Universal Grammar (UG), which refers to the innate set of structural rules and parameters for human language (Chomsky, 1965/2014; Chomsky & Halle, 1965). As previously discussed, OT-type grammars incorporate UG principles as universal markedness constraints. When children acquiring language make errors, or produce language outputs that are not the same as that of the target language, they tend to rely on unmarked structures. For example, because CV syllables are the most unmarked, often times CCV or CVC syllables are reduced to the less marked CV shape. Because OT incorporates universal markedness into the structure of the grammar, it is able to capture these acquisition patterns better than a rule-based system could, and so it has been used to model the acquisition process of how a child arrives at a fully specified adult grammar over time.

In the phonological domain, this issue has been addressed via phonological learning algorithms. The goal of these computational models is to use data from a target language to find a grammar that is consistent with the data it has been provided. With enough data, learning algorithms should converge and form a grammar that is able to predict both what forms are and are not permitted in the target language.

2.3.1 OT learning algorithms

As described, OT grammars consist of input-output pairs that are evaluated by a set of ranked constraints. A learning algorithm for an OT grammar would thus be supplied with these input-output pairs and constraints and would use this information to find an appropriate constraint ranking that is compatible with the target grammar. The original OT learner is the an algorithm that uses Constraint Demotion (CD) (Tesar, 1995; Tesar & Smolensky, 1998, 2000). Simply put, when a CD algorithm is presented with an input-output pair, if the current constraint ranking does not predict the correct output, the constraint(s) that favor the losing output candidates are demoted in the ranking.

However, as Boersma and Pater (2008) point out, CD learners do not allow for any variation, which is clearly present in natural language. Variation is particularly important when it comes to bilingual language, but that will be discussed further on. A phonological learner like the Gradual Learning Algorithm (henceforth, GLA) is compatible with some variation because it assumes Stochastic OT, a version of OT with a mechanism that produces variation (Boersma et al., 1997; Boersma & Hayes, 2001; Boersma & Pater, 2008; Boersma, Levelt, et al., 2000; Boersma, 2000). Another benefit to the GLA compared to CD is that it converges on the grammar much faster and has a learning curve that more accurately reflects actual language acquisition (Boersma, 1998, 2000).

2.3.2 HG-GLA$_O$

Boersma and Pater (2008) developed a version of the GLA that is an extension of Boersma's (1998) and Jäger's (2007) phonological learning algorithms that were based on earlier proposals from machine learning (Rosenblatt, 1958, 1962). Boersma and Pater's (2008) model is compatible with HG, which is tasked with finding optimal constraint weights rather than the constraint ranking, which allows for the possibility of variation in the grammar. The HG-GLA$_O$ is an online (the $_O$ represents online), error-driven learner that works as follows. The learner is provided with the constraints and is given a single input with its associated output candidates. If the winning candidate of the learner is different from the correct output (i.e., target adult form), the weights are adjusted so that the error is less likely with future inputs.

When the learner's winner differs from the actual winner, this is shown with an error tableau like in (5). In this example, the current weights of the grammar predict an incorrect output candidate O_2 as the winner, marked with the ☞symbol. However, the candidate O_1 is the actual correct output, marked with the ✓symbol. The learner then updates the constraint weights by adding weight to C_2 and removing weight from C_1.

(5)

	input	C_1 1.5	C_2 1.0	H
✓	O_1	-1		-1.5
☞	O_2		-1	-1
	O_3	-1	-2	-3.5

HG-GLA$_O$ error tableau

The weight that is added to and subtracted from constraints during each update is not arbitrary. To calculate this, one must create an error vector that shows the difference between the constraints that the incorrect 'winner' satisfies versus those that the correct form satisfies. For each constraint, the number of violations of the incorrect winner is subtracted from the number of violations of the correct winner. This is shown in (6). Candidate O_2, the incorrect winner (Learner's grammar's selected output), violates constraint C_1 zero times and candidate O_1, the correct winner (from the training data), violates this same constraint once. For this constraint C_1, the score in the error vector is equal to -1 - 0 = -1. The second constraint C_2 is violated once by the learner's grammar's incorrect winner and zero times by the training data's correct output: this results in an error score of 0 - (-1) = +1.

(6)

Correct ∼ incorrect	C_1	C_2
Training data (winner) [O_1]	-1	
Learner's grammar (loser) [O_2]		-1
Difference (T-L)	-1	+1

HG-GLA$_O$ error vector

Then, the values in the error vector are multiplied by a constant, the learning rate, which is usually a small number. This product is then added to the constraint's weight. As Boersma and Pater (2008) point out, smaller learning rates result in longer acquisition periods, but the HG-GLA$_O$ will still eventually find the optimal constraint weights that are appropriate for the grammar. Here, the learning rate is set to 0.4, and the updated grammar is shown in (7). The updated grammar now makes the correct output prediction, as the learner's grammar and training data winners are the same. This process continues as more input is provided until the learner converges on a grammar given the data set supplied.

(7)

input	C_1 1.1	C_2 1.4	H
✓☞ O_1	-1		-1.1
O_2		-1	-1.4
O_3	-1	-2	-3.9

HG-GLA$_O$ updated grammar

2.3.2.1 Maximum entropy grammar and batch gradient descent HG-GLA

The HG-GLA$_O$ is an example of an online learner that receives one in put at a time, and so learning updates are dependent on the probability of each input being selected from the set of possible inputs. Pater and Staubs (2013) have since developed a gradual batch learner version of the HG-GLA$_O$. The batch HG-GLA (henceforth, HG-GLA) is an offline version of the HG-GLA$_O$. It is gradual in the sense that the weights change by a small amount during each iteration, but it is batch in the sense that each update is dependent by the probability distribution of the whole data set rather than by a single sample that is selected from the set of possible inputs (Jäger, 2007; Moreton, Pater, & Pertsova, 2017). The HG-GLA calculates an error vector similar to that of the HG-GLA$_O$ except that the values are weighted by the learner's expectations, which are probabilities of each output candidate as calculated by a Maximum Entropy (henceforth, MaxEnt) grammar.

MaxEnt is an extension of HG that uses a candidate's harmony score as the basis for assigning probability to each candidate (Smolensky, 1986; Goldwater & Johnson, 2003; Jäger, 2007; Hayes & Wilson, 2008). An output candidate's probability is computed in three steps. First, the eHarmony (e^H) is calculated, which is the mathematical constant e (i.e., the base of the natural logarithm) to the exponent of the given candidates harmony score (e^x where x equals the harmony score of that given candidate). After determining the eHarmony of each candidate, the eHarmony scores of all candidates are summed together, and then the probability (P) of any given candidate is found by dividing the eHarmony score of that candidate by the sum of all candidates' eHarmony scores. This is shown in the tableau in (8).

(8)

input	C_1 1.5	C_2 1	H	e^H	P
✓Training data (winner) O_1	-1		-1.5	0.22	0.38
☞Learner's grammar (loser) O_2		-1	-1	0.36	0.62

When the learner does not predict the correct output for each epoch, a violation vector is created by weighting (multiplying) the violations of each constraint by the probability of each candidate, which is shown in the bottom row of the tableau in (9).

(9)

input	C_1 1.5	C_2 1	H	e^H	P
Training data (winner) O_1	-1		-1.5	0.22	0.38
Learner's grammar (loser) O_2		-1	-1	0.36	0.62
	-.38	-.62			

As previously discussed, the HG-GLA$_O$ receives inputs one at a time, and the learner's expectations about the candidates' probabilities are derived from sampling the possible outputs: outputs with higher probability are seen more often. In contrast, because the batch HG-GLA is not an online learner, it requires information about how likely each output candidate is. For the present example, output candidate O_1 is always the correct output form, so its probability is 1, and O_2's probability is 0. However, if there were variability in output forms, these proportions would reflect the probability of each output. These probabilities are used as training data. The HG-GLA then uses the same learning update as the HG-GLA$_O$ but over the training data and the learner's expectation (i.e., the error vector). As shown in the table in (10), this update would result in the weight of C_1 being adjusted by -0.62 and the weight of C_2 by +0.62.

(10)

	C_1	C_2
Training data O_1	-1	0
Learner's grammar O_2	-.38	-.62
Weight adjustment (T-L)	-0.62	+0.62

2.3.3 Summary of OT, HG, and phonological learning algorithms

In sum, both OT and HG are constraint based models of grammar that rely on a set of universal markedness and faithfulness constraints that interact with one another to predict language outputs. Markedness constraints refer to constraints that impose universal tendencies, such as preferred syllable shapes. Faithfulness constraints promote outputs that are most similar to the input. These constraints interact in a language-specific way to produce a grammar that is able to capture the patterns of that language, as well as predict what other output forms are possible versus impossible in that language. In OT the constraints are ranked with the most important constraints outranking the less important constraints. HG, on the other hand, utilizes constraint weights, where the more important constraints are assigned a higher weight than less important constraints.

This type of grammatical model lends itself well to modeling language acquisition stages. Models such as the HG-GLA and the online version (HG-GLA$_O$) are able to find the optimal constraint weights for the grammar based only on the positive input they are given. If an output of the model differs from the correct output of the target language, the constraint weights are updated to make this error less likely in the future. These algorithms are also compatible with MaxEnt probabilities, meaning that the probability of each output candidate can be easily calculated. This allows for some variation in the language, a task that simple OT learners are not able to do.

In the current dissertation, the HG-GLA is used to make predictions about bilinguals' acceptability of forms that are ungrammatical in Spanish but grammatical in English in order to explain variability in phonology when the two grammars of bilinguals are at odds with each other.

2.4 Phonotactics

Part of being a competent speaker of a fully-acquired language is knowing the rules that characterize the words of a language. At the level of the word, much of this implicit knowledge is phonotactic knowledge. As briefly stated above, phonotactics is a branch of phonology that is concerned with which sounds can and cannot appear in various positions of a syllable, which include the onset, nucleus, and coda positions. Cross-linguistically, a syllable is minimally represented by a nucleus. All languages permit vowels to be syllable nuclei, but only some languages, allow consonants to hold this position as well. For example in English, a word like ‘bottle’ [bɑɾl̩] has a liquid /l/ as the nucleus of the second syllable, and the language of Berber allows any consonant, even obstruents, to fill this position (Dell & Elmedlaoui, 1985, 1988, 2012).

Languages further vary in the constraints that are applied to the syllable-initial (i.e., onset) and syllable final (i.e., coda) positions. As previously mentioned, the least marked syllable cross-linguistically has a CV structure, which is present in every language, but languages differ in what other syllable forms are allowed. These syllable variations are unidirectional though, in the sense that the presence of a complex coda, for example, implies the presence of singleton codas, but the reverse is not true. For example, if a word /pæst/ is present, then a word like /pæt/ must be well-formed because singleton codas are universally less-marked than complex codas, but the existence of /pæt/ does not mean that a word like /pæst/ is automatically grammatical.

2.5 Sonority Sequencing Principle

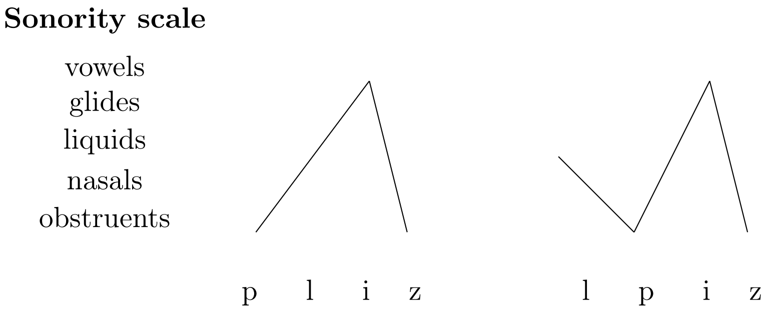
Constraints related to syllable structure, or phonotactics, are most commonly discussed in relation to sonority. Although languages differ in their phonotactics, there are clear universal tendencies in sonority and sonority patterns that have been described in terms of the Sonority Sequencing Principle (SSP). The SSP is a scale that characterizes cross-linguistic well-formedness in syllable structures (Selkirk, 1984; Clements, 1990; Blevins, 1995). It is a universal hierarchy that ranks sounds based on their sonority, which is often correlated with phonetic intensity or loudness (Parker, 2002) and articulatory stricture (Keating, 1983). Vowels are universally the most sonorant segments, sequentially followed by glides (e.g., /w, j/), liquids (e.g., /l, r/), nasals (e.g., /m, n/), and obstruents (e.g., /p, s/), as shown in (11); this broad classification of sounds has been shown to be overwhelmingly consistent between languages. Further distinctions within each broad category have been proposed, but they seem to differ between languages (Selkirk, 1984; Parker, 2002; Jany, Gordon, Nash, & Takara, 2007).

(11) **Least to most sonorant sound classes:**

obstruents < nasals < liquids < glides < vowels

In addition to ranking segments based on sonority, the SSP also requires that each syllable has exactly one sonority peak: the nucleus. The syllable margins (i.e., onsets and codas) that surround the nucleus must create a linear slope rising towards and falling from the nucleus, respectively. Thus, the nucleus should always be the most sonorous segment of a word. It is for this reason that a monosyllabic, English word like ‘please’ /pliz/ is well-formed: the slope of sonority progressively rises towards the nucleic vowel /i/ and falls following this vowel. In turn, a similar, monosyllabic nonce word like ‘lpease’ /lpiz/ is be ill-formed in English, because liquids are more sonorous than obstruents, creating a second sonority peak in the syllable and violating the sonority hierarchy. This contrast is demonstrated in (12).

(12)



Despite the robust tendency for languages to adhere to the SSP, some languages do allow syllables that violate this generalization that requires a sonority rise in the consonants in an onset cluster (Kawasaki, 1982). Some examples include Russian (Halle & Jones, 1971; Itô,

1982) and Polish (Cyran & Gussmann, 1999; Jarosz & Rysling, 2017). Even though these languages permit highly marked syllable structures, they all follow the universal implication that any language with reversed sonority onset clusters (e.g., /lb/) always allows less marked onset clusters as well, including those with sonority plateaus (e.g., /bd/) and sonority rises (e.g., /bl/). Thus, the restrictiveness is unidirectional: a highly restrictive language that only allows sonority rises will not allow sonority falls in onset clusters, such as in English and Spanish, but a less restrictive language that allows sonority falls in onset clusters must also allow sonority rises, such as in Russian and Polish.

In conjunction with the SSP, the notion of a Minimum Sonority Distance (MSD) between consonants in clusters is also used to characterize onset cluster phonotactics both within and across languages. The MSD is a general tendency where languages require a certain sonority distance between the first and second consonant in onset clusters (Steriade, 1982; Selkirk, 1984). Cross-linguistically, consonant clusters that are composed of segments that are closer together in sonority are more marked, whereas clusters whose segments are farther apart in sonority are less marked. For example, an onset cluster composed of /pn/ is more marked than one composed of /pl/ because there is a smaller sonority distance between the obstruent /p/ and the nasal /n/ than there is between this obstruent and the liquid /l/. This is why most languages that allow onset clusters will permit sequences like /pl/ and /pr/ but will not allow /pn/ and /ps/, despite the fact that these clusters all satisfy the SSP's pressure to have a sonority rise.[1] In English, for example, 'please' is well-formed because there is a large sonority distance between /p/ and /l/, but a word like 'pmease' is ill-formed because, even though there is a sonority rise between /p/ and /m/, the distance is not great enough. The MSD thus prefers adjacent segments with larger sonority differences, and because of the implicational markedness in grammar, if a language permits an onset cluster with a smaller MSD, it will also permit larger sonority differences.

The SSP and MSD are examples of implicational markedness in language. For example, if a language's SSP parameters allow stop+nasal clusters, then stop+liquid clusters must also

---

[1] With respect to /ps/, it should be noted that it is generally accepted that fricatives are more sonorous than stops within the natural class of obstruents, although there are some exceptions to this (Selkirk, 1984; Parker, 2012; Jany et al., 2007).

be permitted because the sonority distance is greater. Similarly, if a language's MSD allows a sonority distance of zero between consonants in a cluster (e.g., stop+stop clusters), then it must also allow larger sonority distances in clusters as well. However, these implications are unidirectional: the presence of a smaller sonority distance implies the presence of a larger sonority distance, but a larger sonority difference does not imply the presence of a smaller sonority distance. As discussed in Section 2.2, the notion of implicational markedness is built into the architecture of OT-type grammars, and the SSP is no exception (De Lacy, 2006).

2.6 Spanish and English onset cluster phonotactics

Given that the present dissertation aims to investigate phonotactics in English-Spanish bilinguals, it is important to summarize the phonological inventories and onset cluster phonotactics of each language. Spanish and English are similar in that they both allow complex onsets, but they differ in what segments are present in the phonological inventories and what sonority profiles are allowed in onset cluster position.

Spanish and English phonotactics are similar in that they both allow onset clusters. Although both languages allow many of the same onset clusters, such as /pl/ (e.g., *playa* and 'play'), the biggest difference between the two phonotactic systems is that English permits sC-clusters to surface (e.g., [skul] 'school'), whereas Spanish does not (e.g., [*skwela], [es.kwela] *escuela*).

2.6.1 Spanish onset clusters

Spanish allows some types of onset clusters. All stop allophones ([p, t, k, b, d, g, β, ð, ɣ]) and the labial fricative [f] can occupy the initial spot of an onset cluster, followed only by a liquid ([ɾ, l]). On the other hand, nasals ([m, n, ŋ, ɲ]), liquids ([ɾ, l]), and glides ([w, j])

cannot be the first consonant of an onset cluster (Harris, 1983). As such, all Spanish onset clusters abide by the SSP. Examples are shown below in Table 2.1.[2, 3]

Table 2.1: Licit onset clusters in Spanish

cluster	transcription	translation	cluster	transcription	translation
pl	[pla.ta]	‘silver’	pɾ	[pɾe.sjo]	‘price’
			tɾ	[tɾa.xe]	‘suit’
kl	[kla.βo]	‘nail’	kɾ	[kɾe.a.sjon]	‘creation’
bl	[blu.sa]	‘blouse’	bɾ	[bɾu.xa]	‘witch’
			dɾ	[dɾo.ɣa]	‘drug’
ɡl	[ɡlo.sa]	‘gloss’	ɡɾ	[ɡɾan.xa]	‘farm’
βl	[la βlu.sa]	‘the blouse’	βɾ	[la βɾu.xa]	‘the witch’
			ðɾ	[la ðɾo.ɡa]	‘the drug’
ɣl	[la ɣlo.sa]	‘the gloss’	ɣɾ	[la ɣran.xa]	‘the farm’
fl	[flor]	‘flower’	fɾ	[fɾu.ta]	‘fruit’

In Spanish, /s/ cannot form the first part of an onset cluster, even if it conforms to the SSP (e.g., [*slaβo]). Although phonotactically illegal clusters are often repaired with an intrusive, epenthetic vowel (e.g., /lbo/ → [lə.bo]), sC-clusters behave differently cross-linguistically and tend to be repaired with a prothetic vowel (e.g., /slo/ → [es.lo]) (Broselow, 1992; Fleischhacker, 2001a, 2001b; Gouskova, 2001). Prothetic vowels are like epenthetic vowels in that they are inserted between consonants to break up illegal clusters; the difference is just that prothetic vowels are inserted before the illegal consonant cluster so that the first consonant becomes a coda of a new syllable and the second consonant becomes a singleton onset of the following syllable (i.e., CC → VC.C), whereas an epenthetic vowel is inserted

---

[2] [dl] and [ðl] clusters are exceptions to this generalization in all Spanish dialects, and [tl] clusters are as well in some dialects. However, this is likely due to constraints unrelated to syllable structure but rather constraints that are perceptually-motivated and/or articulatory in nature (i.e., place-related constraints) (Bradley, 2006; Hallé & Best, 2007; Breen, Kingston, & Sanders, 2013; Flemming, 2007, 2013; Parker, 2012). These clusters are prohibited not only in Spanish, but also in other languages such as English and Italian, among others (Kawasaki, 1982). Because these exceptions are distinct for reasons that are not related to sonority and syllable structure, further discussion of them will be omitted from this paper, but an analysis of this restriction can be seen in Bradley (2006).

[3] Words beginning with the sound sequences of [pw, bw, dw], etc. are not included here because there is debate as to whether these are actual onset clusters or singleton onsets with a diphthong vowel in Spanish. This topic is outside the scope of this paper, so these clusters will not be discussed. However, because onset cluster phonotactics are typically characterized in terms of sonority sequencing, and these onset clusters with a glide have the highest possible sonority difference, any such generalization would not exclude these clusters.

between the two consonants (i.e., CC → CVC). Spanish is no exception to this, inserting a prothetic /e/ to repair sC-clusters, as shown in Table 2.2.[4] It is important to mention that the prothetic /e/ is distinct from a lexical vowel. For example, the /e/ in 'escuela' is, at least historically, not a lexical vowel and is present to prevent the consonants [sk] from forming an onset cluster. This differs from the /a/ in a word like 'astuto', which is a lexical vowel and is not inserted solely to break up an illegal onset cluster. Although /e/ is not a lexical vowel in instances of sC-cluster prothesis, it is a contrastive phoneme in Spanish (e.g., *paso* 'step' versus *peso* 'weight').

Table 2.2: Illicit onset clusters in Spanish

cluster	transcription	translation
*sp	[es.pu.ma]	'foam, froth'
*st	[es.tu.fa]	'stove'
*sk	[es.kwe.la]	'school'
*sb	[ez.βel.to]	'slender, slim'
*sd	[ez.dru.ju.la]	'proparoxytone, word with stress on the antepenultimate syllable'
*sg	[ez.ɣri.ma]	'fencing, swordplay'
*sm	[ez.mal.te]	'enamel, nail polish'
*sl	[ez.lo.ɾa]	'length'
	[ez.lo.βa.ko]	'Slav'
*sɾ	[ez.ɾi.laŋ.kes]	'Sri Lankan'

To summarize, the onset cluster phonotactics of Spanish can be concisely described as allowing any non-sibilant obstruent followed only by a liquid to form onset clusters (Tetzloff, 2020). Any word that would instantiate an sC-cluster (historically or as a loan word) is repaired by inserting a prothetic /e/ to create a bisyllabic structure.

2.6.2 English onset clusters

Like Spanish, English also allows onset clusters, but compared to Spanish, English is seemingly more liberal in what segments can pattern together in this position. English

---

[4] In this cluster context, the /s/ assimilates in voicing to the following segment, meaning it is realized as [z] when preceding a voiced segment. This differs in Peninsular Spanish when /s/ is followed by /ɾ/, in which case the /s/ completely assimilates to the following rhotic, resulting in a trill [r]; for example, a word like 'Israel' would be pronounced as [i.ra.el] (Navarro Tomás, 1980: 80).

allows stops ([p, t, k, b, d, g]) and voiceless fricatives ([f, θ, s]) to form the first part of an onset cluster, immediately followed by a liquid ([l, ɹ]) or a glide ([w, j]). A notable difference between Spanish and English is the fact that English permits onset clusters with a sibilant consonant as the first consonant of the cluster (e.g., ’school’, ‘smell’, ‘slick’ etc.). Within these sC-clusters, some of them conform to the SSP, while others don’t, namely those involving stop consonants (e.g., /sp, st, sk, sm, sn/).[5,6]. Table 2.3 illustrates these possible onset clusters of English.

Table 2.3: Licit onset clusters in English

cluster	transcription	word	cluster	transcription	word
pl	[pʰleɪ]	‘play’	pɹ	[pʰɹɑp]	‘prop’
			tɹ	[tʰɹi]	‘tree’
kl	[kʰlaʊd]	‘cloud’	kɹ	[kʰɹaɪ]	‘cry’
bl	[blu]	‘blue’	bɹ	[bɹʌðɚ]	‘brother’
			dɹ	[dɹim]	‘dream’
gl	[glʌv]	‘glove’	gɹ	[gɹæs]	‘grass’
fl	[flæg]	‘flag’	fɹ	[fɹi]	‘free’
			θɹ	[θɹi]	‘three’
tw	[twɪŋ]	‘twin’	dw	[dwɛl]	‘dwell’
kw	[kwin]	‘queen’	gw	[gwɛn]	‘Gwen’
sp	[spid]	‘speed’	sm	[smɛl]	‘smell’
st	[stɪk]	‘stick’	sn	[snoʊ]	‘snow’
sk	[skul]	‘school’	sl	[slip]	‘sleep’
sf	[sfiɹ]	‘sphere’	sw	[swit]	‘sweet’
spl	[splɪt]	‘split’	spɹ	[spɹɪŋ]	‘spring’
			stɹ	[stɹɛŋkθ]	‘strength’
skl	[sklɚɹoʊsɪs]	‘sclerosis’	skɹ	[skɹu]	‘screw’

2.6.3 Summary of Spanish and English onset clusters

Descriptively, Spanish and English are similar in their onset cluster phonotactics in that they allow stop consonants followed by liquids to form complex onsets. English differs from Spanish, however, in that it also allows sC-clusters that abide by the SSP and sC-clusters

---

[5] English, like Spanish, prohibits [tl] and [dl] clusters, along with [sr] clusters for independent, perceptually-motivated reasons (Moreton, 2002; Bradley, 2006; Hallé, 2008; Breen et al., 2013).

[6] The onset clusters /pw/ and /bw/ do not occur in native English words. There is some debate as to whether these are absent due to OCP co-occurrence restrictions (Duanmu, 2002; Cardoso & Liakin, 2009) or due to an accidental gap (Albright, 2009; Daland et al., 2011; White & Chiu, 2017)

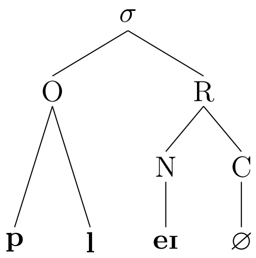
that appear to violate it. In Spanish, all clusters beginning with /s/ are illegal and are repaired by the prothesis of /e/ (e.g., /sp/ → [es.p]).

2.7 sC-clusters

sC-clusters have received a substantial amount of attention in the literature because they often pattern differently from other types of onset clusters across languages. This has led to the proposal that some sC-clusters have a different suprasegmental structure than other types of onset clusters.

Typical onset clusters, which in both English and Spanish must rise in sonority, have a branching onset. Branching onsets are universally left-headed (Kaye, Lowenstamm, & Vergnaud, 1990), meaning that onset clusters are right branching. This structure is shown in (13) with the word 'play' [pleɪ]. The sigma represents the syllable, *O* for onset, *R* for rhyme, *N* for nucleus, and *C* for coda. Because this word has a true complex onset, /p/ is the head of the onset cluster and both the /p/ and /l/ segments branch from the onset node of the syllable structure.

(13)



Branching onsets in a language that only permits sonority rises in onset clusters implies that no sonority plateaus or falls will appear at the beginning of a syllable. However, English and other languages like German, Dutch, Greek, and Italian, among others, permit words that seemingly violate this principle: sC-clusters. On the surface, some sC-clusters appear to violate the SSP restrictions of these languages, which is evidenced by the fact that other clusters of this sonority profile are unattested. For example, in English /st/ can appear at the beginning of words but /fp/ cannot despite the sonority profiles being virtually identical.

This has led to the proposal that sC-clusters have a different syllabic structure than other SSP-abiding onset clusters with branching onsets.

The distinction between branching onsets and sC-clusters leads to the typological prediction that there should exist languages that allow branching onsets but not sC-clusters, languages that allow sC-clusters but not branching onsets, and languages that allow both. This prediction is supported: Hindi has branching onsets but no sC-clusters, the Native American language of Acoma has sC-clusters but no branching onsets, and English has both. Because these two phenomena are argued to be separate, they should also be combinable, which they are in English. English allows words like 'stripe' and 'splash' that have true branching onsets preceded by the initial /s/; this is only possible because the /s/ is not part of the onset cluster, but part of a preceding syllable.

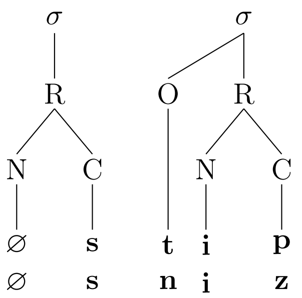
Furthermore, sC-clusters behave differently with regards to place restrictions in onset clusters. For example, English does not allow [*tl] clusters due to the Obligatory Contour Principle (OCP) which bans coronal-coronal clusters like /tl/ and /dl/ (McCarthy, 1986). However, sC-clusters are different in that the OCP does not seem to apply, as English allows neighboring coronals in this context (e.g., /st, sn/). Likewise, in Dutch obstruent+nasal clusters are well-formed but only for /n/, the coronal nasal (i.e., /kn, fn, ɣn/ but not /*km, *fm, *ɣm/); sC-clusters again pattern differently from the other onset clusters in that the second consonant does not need to be coronal (e.g., /sp, sm/). Similar patterns are attested in other languages as well, such as Modern Greek (Steriade, 1982). This evidence supports the analysis of sC-clusters being structurally distinct from other onset clusters, which have branching onsets.

2.7.1 Theories of sC-cluster structure

There have been multiple proposals for how to account for the difference in sC- versus other onset clusters. One proposal is that the initial /s/ is extraprosodic, meaning it is not part of the syllable but on the left periphery (Steriade, 1982; Goldsmith, 1990). Van der Hulst (1984) posited that /s/ is part of the syllable but not syllabified as an onset. Kaye

(1992) and Goad (2011, 2012), however, claim that the /s/ in sC-clusters is actually the coda of a preceding syllable that lacks a nucleus, as in (14). [7]

(14)



Goad (2012) presents data from European Portuguese allomorphy, shown in (15), which supports this analysis of sC-clusters of coda /s/ plus an onset. European Portuguese does not permit nasal consonants in coda position. Because the stem of (16a), *admissivel*, is vowel initial, the nasal is syllabified as the onset of the second syllable (e.g., /i.nad/). The stems in (16b-d) differ in that they are consonant or consonant cluster (16d) initial; as such, the nasal of the prefix /in/ cannot be syllabified as an onset nor as a coda, resulting in the nasalization of the vowel and the deletion of the /n/ (e.g., [ĩ.tra]). sC-clusters, however, do not pattern with the other consonant- or consonant cluster-inital stems, as in (16e). This can be accounted for if it is assumed that sC-clusters consist of a nucleus-less syllable with the sibilant, in this case [ʃ] rather than /s/, as the coda, shown in (16).[8]

---

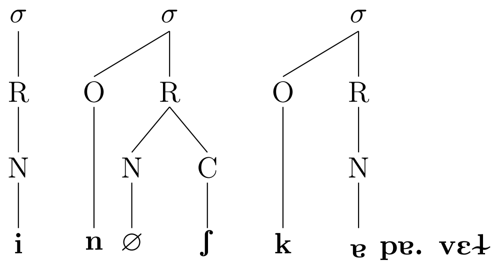
[7] It should also be noted that words like 'especially' and 'asteroid' share this structure but with a lexical vowel that occupies the nucleus of the first syllable, which is fully vocalized.

[8] This differs slightly from what is attested in English, because in English it is impossible to have a nucleus-less syllable that also has an onset. Nevertheless, the structure of the sC-cluster being a coda+onset is the same.

(15) European Portuguese Data

(a)	[in]*admissivel*	‘inadmissiable’
(b)	[ĩ]*pureza*	‘impurity’
(c)	[ĩ]*satisfeito*	‘dissatisfied’
(d)	[ĩ]*tratavel*	‘unsociable’
(e)	[inʃk]*apavel*	‘inescapable’

(16)



Child language acquisition provides further evidence in favor of this coda+onset structure of sC-clusters. In the early stages of English production (∼1;0-3;0), children often do not faithfully produce both consonants in onset clusters. They tend to reduce the cluster to a single consonant, preserving the less sonorous consonant (e.g., ‘blue’ /blu/ → [bu], or ‘flow’ /floʊ/ → [foʊ]) (Fikkert, 1994; Barlow & Gierut, 1999; Barlow, 2001; McLeod, Van Doorn, & Reed, 2001; among others). This pattern of cluster reduction is rather uniform for branching onsets, but sC-clusters appear to behave differently.

As mentioned, in early child production the less sonorous consonant is typically maintained and the more sonorous consonant is deleted in order to maximize the sonority distance between the onset and the nucleus vowel. With sC-clusters, on the other hand, it is not always the more sonorous consonant that is deleted in cluster reduction. Two of the children examined in Pater and Barlow’s (2003) study consistently deleted the /s/ in sC-clusters, even when this segment was less sonorous than its adjacent consonant. For example, ‘sneeze’

/sniz/ was reduced to [niz], ‘smell’ /smɛl/ to [mɛl], and ‘snow’ /snoʊ/ to [noʊ], among others. Children in Goad and Rose’s (2004) study showed similar patterns of cluster reduction for sC-clusters in both English and Dutch: in /s/+stop clusters, the /s/ was deleted while the following obstruent was maintained.

These data from allomorphy in European Portuguese and child acquisition strongly support the hypothesis that sC-clusters are distinct from branching onsets and are composed of a coda /s/ followed by an onset of the following syllable.

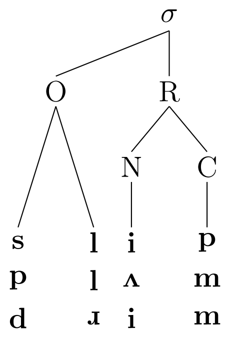
2.7.2 A dual model of sC-clusters

Although there is consensus that some sC-clusters are distinct from branching onsets when the sC-sequence violates the SSP and/or the MSD (e.g., /sp, st, sk, sm, sn/), there is some debate as to whether or not SSP-abiding sC-clusters (e.g., /sl, sw/) pattern with /s/+stop clusters or with branching onset clusters. Child acquisition data for branching onsets and sC-clusters that violate the sonority restrictions of a language are quite uniform in their reduction patterns. For example ‘stop’ /stɑp/ and ‘spot’ /spɑt/ will be consistently reduced to [tap] and [pɑt], respectively. However, there is variability in the cluster reduction patterns of sC-words that abide by sonority preferences (i.e., /sl, sw/). Some data show that children tend to follow the more unmarked pattern observed with branching onsets, in which the more sonorous consonant is deleted (e.g., ‘slow’ /sloʊ/ → [soʊ]) (Barlow, 1997; Pater & Barlow, 2003; Yavaş & Someillan, 2005; Yavaş & Core, 2006; Yavaş, Ben-David, Gerrits, Kristoffersen, & Simonsen, 2008; Yavaş, 2011). Yavaş and colleagues argue that this reduction from /sl/ → /s/, which follows the reduction patterns of branching onsets, is evidence that /sl/ and other sonority-obeying sC-clusters have a branching onset structure and are distinct from other sC-clusters. Pater and Barlow (2003) account for this same reduction pattern in child sC-productions by ranking a constraint that disprefers fricatives over a constraint that disprefers lateral onsets (i.e., *LATERAL-ONSET » *FRICATIVE). Although this analysis accounts for the child acquisition data, it does not inform the possible differences in the syllabification of the two types of sC-clusters. Nevertheless, these data support the notion of a dual model of sC-clusters: sC-clusters that do not violate the SSP are syllabified as branching onsets and follow the same reduction patterns as other non-sC onset clusters, while SSP-violating sC-clusters are syllabified differently as a coda /s/ followed by a singleton onset.

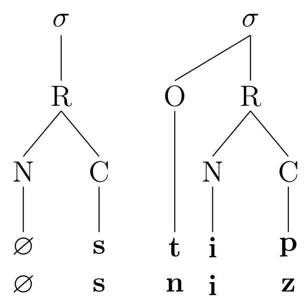
Other acquisition data has shown the opposite reduction pattern in these clusters, with the more sonorant segment being retained (e.g., ‘slow’ /sloʊ/ → [loʊ]) (Goad & Rose, 2004; Kristoffersen & Simonsen, 2006). The former pattern would suggest that SSP-abiding sC-clusters are structured as branching onsets (i.e., taughtosyllabic), whereas the latter pattern would suggest that these sC-clusters share the structure of other sC-structures that violate sonority preferences (i.e., bisyllabic). These authors acknowledge that there is variability in /sl/ cluster reductions, but they argue that this reflects different stages of sC-cluster acquisition and that the final grammar of all sC-clusters is uniform, with the non-sibilant being the head of the clusters and the /s/ an appendix or coda of a preceding syllable (Goad, 2011).

This dissertation assumes the former proposal, where sC-clusters that do not violate the SSP parameters of the language are structurally distinct from sC-clusters that do: /sl/ clusters have a branching onset, while /sp, st, sk, sm, sn/ clusters are syllabified across two syllables with the /s/ being the coda of a nucleus-less initial syllable and the following consonant(s) forming the onset of the second syllable. These schema are shown in (17) and (18). This distinction in the syllabification of s+stop versus s+liquid clusters can be achieved by ranking the constraints below as in the Hasse diagram in (19).

(17)



(18)



Markedness constraints

SSP: Assign one violation for every output candidate that has two adjacent consonants within a syllable whereby the first has a higher sonority rank than the second. Adjacent consonants within a syllable abide by the SSP and rise in sonority (Selkirk, 1984; Clements, 1990; Kager, 1999; Gnanadesikan, 1995/2004).

SylCon: Assign one violation for every pair of adjacent consonants separated by a syllable boundary in which the first segment is more sonorous than the second. The final segment of a syllable should be more sonorous than the initial segment of the following syllable (Gouskova, 2004).

***sC**: Assign one violation for every onset cluster that consists of an /s/ + another consonant. /s/ should never the the initial segment in an onset cluster (Tetzloff, 2020).

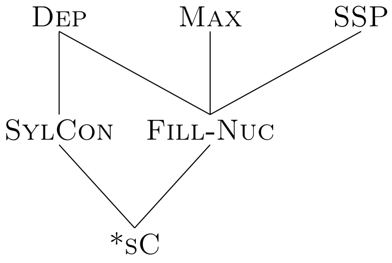
Fill-Nuc: Assign one violation for every syllable that lacks a nucleus. All syllables should have a nucleus (A. Prince & Smolensky, 1993/2004).

Faithfulness constraints

Dep-V: Assign one violation for every vowel in the output that does not have a corresponding vowel in the input. Do not insert vowels (McCarthy & Prince, 1995).

Max-V: Assign one violation for every vowel in the input that does not have a corresponding vowel in the output. Do not delete vowels (McCarthy & Prince, 1995).

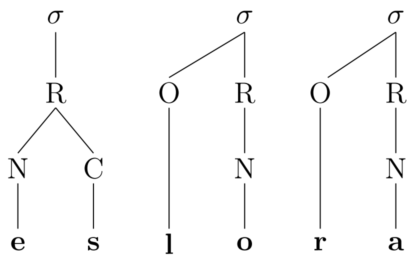
(19)



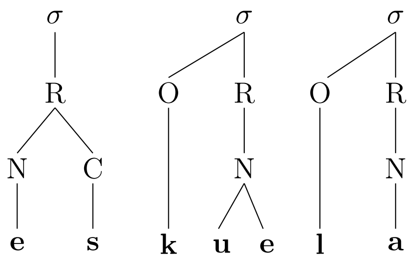
2.7.3 The structure of sC-clusters in Spanish and English

While the dual model sC-clusters outlined in the previous section applies to languages like English and Italian that permit word-initial sC-clusters, other languages like Spanish do not, regardless of the sC-cluster's sonority profile. As previously described, for words that historically would have had sC-clusters (e.g., *sloerie* > *eslora* 'length', *scuola* > *escuela* 'school') Spanish inserts and vocalizes a prothetic vowel before the cluster, which then resyllabifies the word and splits apart the cluster, as shown in (20) and (21).[9] So although Spanish and English have identical syllabic structures for non-sibilant onset clusters, the languages differ in two significant ways when it comes to applying phonotactic constraints regarding words that begin with sC-clusters. First, the English sC-clusters that obey the SSP (e.g., /sl/) have a branching onset structure like in (17), while the sC-clusters that violate English's sonority restrictions (e.g., /sp, st, sk, sm, sn/) have a bisyllabic structure, illustrated in (18). Unlike English, however, it has previously been assumed that all Spanish words that are loanwords with initial sC-clusters have a bisyllabic structure, regardless of sonority (Yavaş & Core, 2006; Escartín Ortiz, 2005; Coffey, 2009).

(20)



(21)



Second, in the case of the bisyllabic structure of Spanish sC-clusters, English instantiates a null, or non-vocalized, nucleus for that initial syllable, while Spanish requires the vocalization of the prothetic vowel in that position.

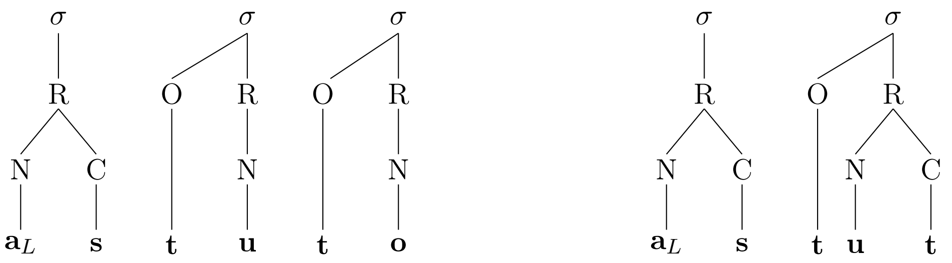
---

[9] The fact that the inserted vowel is prothetic and not epenthetic, as is the case with other illicit onset clusters, provides additional support for the model that /s/ is actually outside of the cluster containing the surface onset cluster, because illicit branching or 'true' onset clusters are separated with an epenthetic vowel between the two consonants (Gouskova, 2001).

It should be further noted that, aside from sC-clusters, both Spanish and English maintain a set of words with this same structure as (20) and (21), but where there is a lexical vowel in the nucleus of the first syllable. Some examples include *aspecto* [as.pek.to] and *ostensible* [os.ten.si.ble] in Spanish and ‘aspect’ [æs.pɛkt] and ‘ostensible’ [ɑs.tɛn.si.bl̩] in English.

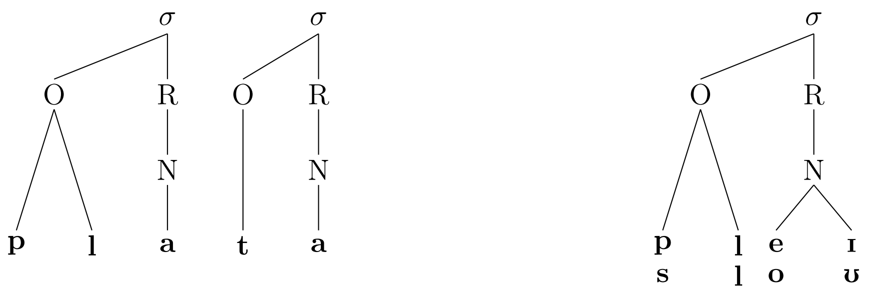
To summarize, the phonological grammar of Spanish overlaps here with English but in a non-identical way. The first overlapping structure is shown in (22), in which the two languages show *full overlap.* The words have fully parallel or identical structures, containing a lexical vowel as the nucleus of the first syllables; the /s/ is the coda of this initial syllable in both languages.

(22) **Full overlap:** Spanish syllabification of *astuto* (and all words with initial lexical vowels) and English syllabification of ‘astute’ (and all words with initial lexical vowels)



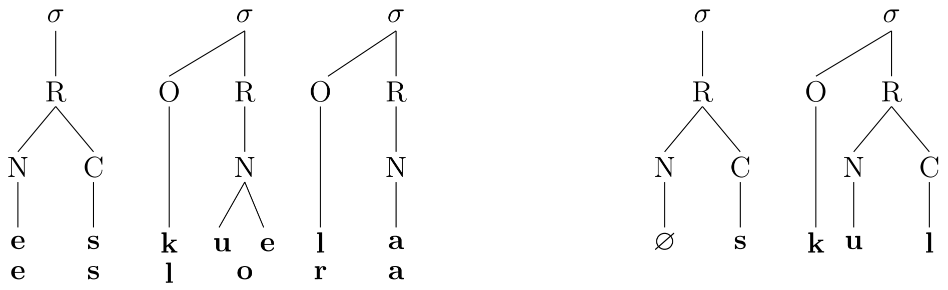
The second overlapping structure is shown in (23), where there is a *partial overlap* between English and Spanish. Both languages use the structure in (23) for all onset clusters that are not sC-clusters, but only English and not Spanish uses this structure for sC-clusters that have a large enough sonority distance and abide by the phonotactic rules of English.

(23) **Partial overlap:** Spanish syllabification of *plata* ‘silver’ (and all non-sC-clusters) and English syllabification of ‘play’ and ‘slow’ (and all non-sC-clusters and SSP obeying sC-clusters)



And finally, the structures shown in (24) are also *partially overlapping*, but in the opposite direction. Here, the bisyllabic structure is used for all Spanish sC-clusters, regardless of sonority, while this structure is only used for English sC-clusters that violate sonority restrictions. Another difference here, not in the structures but in the associated segments is that Spanish requires the insertion of /e/, whereas English allows the empty nucleus in the initial syllable.

(24) **Partial overlap 2:** Spanish syllabification of *escuela* ‘school’ and *eslora* ‘length’ (and all (historical) sC-clusters regardless of sonority) and English syllabification of ‘school’ (and all SSP violating sC-clusters)



2.7.4 Summary of sC-clusters in Spanish and English

In summary, Spanish has a one-to-one association between the presence of sC-clusters and the syllabic structure, whereas English does not. Spanish words with non-sC-onset onset clusters all have a branching onset structure (e.g., /plata/ → [pla.ta], /grano/ → [gra.no]), while those with sC-clusters have a bisyllabic structure that is instantiated with a prothetic /e/ (e.g., /slora/→ [es.lo.ra], /skwela/ → [es.kwe.la]). On the other hand, English has multiple structures for surface sC-clsters. English non-sC-clusters and sC-clusters that abide by the SSP and English's MSD have branching onsets (e.g., /pleɪ/ → [pleɪ], /slɑt/ → [slɑt]), while the sC-clusters that violate sonority preferences are syllabified across two syllables (e.g., /stɑp/ → [s.tɑp], /smɛl/ → [s.mɛl]).

In addition to this structural difference, the bisyllabic sC-cluster words in Spanish require the vocalization of the nucleus as /e/, whereas in English this syllabic position is null and thus not phonetically realized.

2.8 HG-GLA predictions for Spanish sC-clusters

As previously discussed, Spanish onset clusters can consist of an stop + /l, r/ or /f/ + /l, r/; sC-clusters in Spanish differ from those of English in that they are always repaired with a prothetic /e/ (/sC/ → [es.C]). Past analyses of Spanish sC-clusters have assumed that all sC-clusters are all syllabified in the same way regardless of the sonority of the second consonant (i.e., s+stop versus s+liquid). This should not be assumed though because languages that allow surface sC-clusters, like English, have been argued to have different representations for s+stop versus s+liquid clusters. The HG-GLA, described above in Section 2.3, was used here to make predictions about the surface forms and hidden syllable structure of sC-cluster inputs in Spanish to see if different sC-clusters are syllabified in a uniform way or a distinct way like in English.

2.8.1 Methodology

2.8.1.1 Corpus input

The HG-GLA was given sets of input-output data that included all possible onset cluster types in Spanish, shown in Table 2.4. The output forms in the final column are the correct output forms for that given input; the learner was not given any information on the syllabification of the output forms.

As a batch learner, the HG-GLA requires information about the relative frequencies of each onset cluster types. These frequencies were calculated from the SUBTLEX-Spanish corpus, a Spanish corpus with subtitle word frequencies (Cuetos, Glez-Nosti, Barbón, & Brysbaert, 2012), and given to the learner. Although Pater and Staubs (2013) in the modeling of Dutch syllable types included all syllable types in their frequency calculations, the frequencies here do not include all syllable types: syllables with singleton onsets or vowel-initial syllables (that are not followed by an sC-cluster) were entirely excluded. All onset clusters were considered, including clusters preceded by a prothetic /e/ (e.g., *especial* 'special'); hence, the frequency is the percentage of that cluster type compared to all clusters. In the table in 2.4, /sp/ represents all observed tokens of initial /s/ followed by a stop consonant (i.e., /p, t, k, m, n/); all of these sC-intial words in the corpus are loanwords, such as 'snowboard' or 'Starbucks', and their output forms were set as /esC/ following the canonical phonotactic system of Spanish. There were no instances of words beginning with /sl/, so the frequency of this onset cluster is zero. /esp/ stands for all words that begin with /es/ and then /p, t, k, m, n, l, r/. Finally, /pl/ refers to any 'typical' onset cluster with a branching onset in Spanish that consists of a stop or /f/ plus a liquid. Furthermore, the learner was not given information about the hidden structure (i.e., syllabification), so the probability of identical surface forms was 1.

Input	Frequency	Output forms	Correct output
/sp/	0.043	.sp s.p es.p e.sp p	esp
/esp/	0.41	es.p s.p .sp e.sp	esp
/sl/	0	.sl s.l es.l e.sl l	esl
/pl/	0.55	.pl p.l ep.l e.pl p	pl

Table 2.4: Spanish onset cluster input for HG-GLA

2.8.1.2 Constraints

The following constraints, defined below, were selected for analyzing the Spanish input-output pairs with the HG-GLA. Note that additional constraints were needed to accommodate the patterns of Spanish sC-cluster phonotactics versus English in Section 2.7.2.

Markedness constraints

***Complex**: Assign one violation for every output candidate that includes an onset cluster (McCarthy & Prince, 1999).

SSP: Assign one violation for every output candidate that has two adjacent consonants within a syllable whereby the first has a higher sonority rank than the second. Adjacent consonants within a syllable abide by the SSP and rise in sonority (Selkirk, 1984; Clements, 1990; Kager, 1999; Gnanadesikan, 1995/2004).

SylCon: Assign one violation for pair of adjacent consonants separated by a syllable boundary in which the first segment is more sonorous than the second.

The final segment of a syllable should be more sonorous than the initial segment of the following syllable (Gouskova, 2004).

***sC**: Assign one violation for every onset cluster that consists of an /s/ + another consonant. /s/ should never the the initial segment in an onset cluster (Tetzloff, 2020).

Fill-Nuc: Assign one violation for every syllable that lacks a nucleus. All syllabi should have a nucleus (A. Prince & Smolensky, 1993/2004).

R-CodaCond: Assign one violation for every coda consonant that has a place specification that is not coronal. Only coronal and placeless codas are allowed (A. Prince & Smolensky, 1993/2004).

Faithfulness constraints

Dep-V: Assign one violation for every vowel in the output that does not have a corresponding vowel in the input. Do not insert vowels (McCarthy & Prince, 1995).

Max-V: Assign one violation for every vowel in the input that does not have a corresponding vowel in the output. Do not delete vowels (McCarthy & Prince, 1995).

The markedness constraints were assigned an initial weight of 10, and the faithfulness constraints, an initial weight of 0. This imposed a markedness over faithfulness (m>>f) learning bias (Gnanadesikan, 1995/2004; Smolensky, 1996; Boersma, 1998). A m>>f bias is supported by the fact that children's initial productions are unmarked compared to the target, adult language. One such example in is that there is robust evidence that children simplify onset clusters, which are (implicationally) marked compared to singleton onsets. English targets like “clean" [klin] and “draw" [dɹɑ] are often produced as [kin] and [dɑ] with simplified, singleton onsets (Gnanadesikan, 1995/2004). Similar patterns have been reported in other languages as well, such as the Portuguese word “open" [abɾɨ] being produced as [abɨ] (Freitas, 2003) and the Hebrew word "ice cream" [glidɑ] being produced as [gidɑ] (Bloch, 2011).

The learning rate was set to 0.01, and the algorithm was run for 10,000 iterations.

2.8.2 Results

The final constraint weights that the HG-GLA converged on are shown below in Table 2.5. The markedness constraints FILL-NUC and R-CODACOND were assigned the highest weights (10.0), followed by SYLCON (8.6), then followed by the faithfulness constraint MAX-V (2.8). The remaining constraints were assigned a weight of zero.

Constraint	**Final weight**
FILL-NUC	10.0
R-CODACOND	10.0
SYLCON	8.6
MAX-V	2.8
*COMPLEX	0.0
SSP	0.0
*sC	0.0
DEP-V	0.0

Table 2.5: Final constraint weights

These constraint weights yielded the following output predictions, shown in Table 2.6. The first column has the input, the second column shows the possible output candidates (i.e., possible mappings), and the third column shows the MaxEnt probabilities assigned to each output candidate. The outputs favored by the phonological learner have a box around them in the table.

The HG-GLA assigned the most probability to faithful outputs for /esp/ and /pl/ inputs, in the second and fourth rows, as expected. The output candidate for /pl/ with the highest probability syllabified this cluster as a branching onset. For /esp/ cases, it was expected that /es.p/ would be assigned the greatest probability, following past analyses of Spanish syllabification; however, /es.p/ and /e.sp/ were assigned equal probability. Because syllable structure is hidden structure, the important thing is that the observed surface forms together accounted for virtually all of the probability for /esp/ inputs, as the faithful output forms are what are observed in speakers.

Input	Output	Probability
/sp/	.sp	0.33
	s.p	0
	es.p	0.33
	e.sp	0.33
	p	0
/esp/	s.p	0.03
	.sp	0
	es.p	0.48
	e.sp	0.48
/sl/	.sl	0.5
	s.l	0
	es.l	0
	e.sl	0.5
	l	0
/pl/	.pl	.99
	p.l	0
	ep.l	0
	e.pl	0
	p	0

Table 2.6: HG-GLA determined probabilities for each output candidate

The critical information in Table 2.6 are in the first and third rows, which have sC-cluster inputs. The HG-GLA did not predict the same outputs for s+stop and s+liquid clusters. For /sp/ inputs, 67% of the probability was assigned to the output candidates that have the prothetic /e/ (both /es.p/ and /e.sp/, while the remaining probability was assigned to the faithful output with a taughtosyllabic /.sp/ cluster). These results predict variation in the Spanish grammar which may be expected since speakers have some exposure to s+stop clusters through loanwords. Nevertheless, /e/ prothesis is preferred with an s+stop input. The predicted syllable structure for s+liquid inputs differed from that of s+stop inputs. Less probability (50%) was assigned to a surface form including the prothetic /e/ than in s+stop clusters. The remaining 50% of the probability was assigned to a faithful output candidate /.sl/. In both of these cases, the syllable structure is bisyllabic, unlike in s+stop clusters where syllabification varies among winning candidates.

2.8.3 Discussion

These results differ from the general assumptions previously made about Spanish syllabification in sC-clusters. As discussed in Sections 2.6-2.7, past analyses of Spanish sC-cluster phonotactics have assumed that all sC-clusters are repaired in the same way, by inserting an initial /e/ and syllabifying the /s/ as the coda of that syllable and the second consonant as the onset of the next syllable (sC → es.C). The HG-GLA, however, made a different prediction: that s+stop and s+liquid clusters may have both different surface forms and syllable structures. For s+stop inputs, the learner assigned the majority of the probability to surface forms with the prothetic /e/ (esC), while one third of the probability was assigned to the faithful onset cluster /.sp/. Additionally, the s+stop input yielded output candidates with both taughtosyllabic and bisyllabic structures. On the other hand, for s+liquid clusters equal probability was assigned to the two surface forms /sl/ (50%) and /esl/ (50%). Both winning output candidates had a taughtosyllabic, or branching onset, structure, which is in contrast to what was predicted for s+stop clusters. Given that the learner had zero experience with s+liquid clusters, it not only predicts variation in this context but also that the variation differs from that predicted for s+stop clusters.

Syllable structure is an example of hidden structure, so the different output candidates that have the same surface forms (es.C and e.sC) are phonetically realized in the same way, making this prediction hard to test. What can be tested is the acceptability of faithful (sC) and unfaithful (esC) outputs and if there are measurable differences in the acceptability of s+stop versus s+liquid clusters.

Furthermore, given that the MaxEnt probabilities in HG can be understood as reflecting the degree of well-formedness in phonotactics, the results of the HG-GLA lend themselves to the hypothesis that sC-clusters without the prothetic /e/ also have some degree of acceptability in Spanish. The insertion of the initial /e/ in Spanish sC-clusters has been argued to be a repair for initial sC-clusters, but if the cluster is syllabified as a true onset cluster, the need for the prothesis of /e/ is less clear. This is particularly important for the s+liquid clusters since 100% of the probability was assigned to output candidates with tautosyllabic a taughtosyllabic structure. The phonotactic motivation for /e/ prothesis is less clear if /sl/ sequences are syllabified as a branching onset, as the phonological learner predicted. If /e/

is not needed to repair s+liquid clusters, it is possible they will show increased variability compared to s+stop clusters, which are more likely to have the prothetic /e/ in perception.

Not only is this a testable prediction for Spanish monolinguals, but it is particularly important when considering the phonotactic systems of English-Spanish bilinguals, since the tautosyllabic sC-cluster structure is assumed to be the hidden structure for English for s+liquid clusters. The experiments run in this dissertation address this question of acceptability in sC-cluster phonotactics in Spanish monolinguals and English-Spanish bilinguals in the perception of both s+stop and s+liquid clusters.

2.8.4 Summary of HG-GLA predictions for Spanish sC-clusters

A corpus of Spanish subtitles was used to calculate the frequency of Spanish onset clusters, including sC-clusters. These frequencies were used to derive a grammar of weighted constraints using a batch Harmonic Grammar Gradual Learning Algorithm (HG-GLA). The result of this phonological learner made predictions that are inconsistent with previous analyses of Spanish sC-clusters. The HG-GLA predicted that s+stop and s+liquid clusters may not have parallel output forms, but that s+stop clusters are more likely to be 'repaired' with a prothetic /e/ compared to s+liquid clusters, which always have a taughtosyllabic structure despite the potential /e/ prothesis. This, along with the assumption that the assigned probabilities correspond to phonotactic well-formedness, leads to the hypothesis that Spanish sC-clusters may show variability in both monolingual and bilingual speakers. Furthermore, the fact that s+liquid clusters are more likely to be syllabified as branching onsets and less likely to show /e/ prothesis compared to s+stop clusters leads to the additional hypothesis that s+liquid clusters may exhibit more variability than s+stop clusters in perception, particularly for English-Spanish bilinguals whose English grammars have the same hidden structure for these clusters and allows them to surface.

CHAPTER 3

BILINGUAL VARIABILITY

Language systems have been said to be *"où tout se tient"* (Meillet, [1906] 1921: 16), meaning that all parts of a language are interconnected. The essence of this statement is still at the core of much linguistic research in the present day, as many linguists are interested in understanding the cross-domain interfaces. Thus, there is a general assumption that phonology, for example, has some sort of relationship with morphology, syntax, and even semantics, and vice versa, both within monolingual and bilingual speaker populations. One example of this is the presence of allomorphy, which refers to morphemes that have multiple phonological forms that vary based on the phonological context. Additionally, the field of morphosyntax focuses predominantly on this interdependence between the morphological and syntactic domains which cannot be entirely teased apart from each other. The syntax-semantics interface is yet another broad area of research that studies the intimate relationship between these two linguistic domains.

Furthermore, clinical research on patients with language disorders has provided more evidence to confirm this hypothesis that different linguistic domains are connected beyond the various linguistic interfaces. For example, speech delayed children and patients with specific language impairment, both of which primarily target morphosyntax, may also present with ancillary phonological deficits (Joanisse & Seidenberg, 2003); similarly, patients with the logopenic (or phonological) variant of primary progressive aphasia show the largest deficit in the phonological domain, but some also face morphosyntactic difficulties that are not typically associated with this disorder (Tetzloff et al., 2018).

3.1 Bilingualism

This interdependence of linguistic systems extends to the study of bilingualism (i.e., bilingual variability), but it is important to first define what being bilingual means, as there many differing assumptions about what constitutes bilingualism. Thiery (1978), for example, argues that in order to be bilingual, one must have equal and native-like proficiency in two or more languages, all learned from birth (i.e., two L1s). This is a deficit view of bilingualism, as many bilinguals are native speakers of two or more languages but do not use those languages as monolinguals would. In contrast, others have posited that bilingualism means that one is able to produce meaningful utterances in two or more languages, even in a very limited capacity (Beardsmore, 1986; Hakuta, 1986). This view of bilingualism is perhaps too broad, as someone with knowledge of certain lexical items in a second language may be able to produce some sort of meaningful utterance but may have no knowledge of the grammatical structure or complexities of that language. The language behavior of bilinguals depends of different factors such as language pairings, when and how each language was learned, and how each language is used in daily life. For the purpose of this paper, bilingualism will be defined as having lexical and a specified degree of grammatical competence in both languages, as determined by inter-rater judgment and self-ratings of speaking and listening comprehension; self-ratings of language competence have been widely used in past studies of bilingualism and have been shown to reliably reflect language abilities (Bachman & Palmer, 1989; MacIntyre, Noels, & Clément, 1997; Ross, 1998).

3.1.1 Variability in bilingualism

Assuming that bilinguals have some lexical and grammatical competence in each language, it has been long established that they are not two monolinguals in one and do not have one combined grammar for both languages (Genesee, 1989; Paradis & Genesee, 1996; Paradis, 2001). Decades of research have not only studied the interfaces between linguistic domains but also if a bilingual's two linguistic systems are also interconnected or if they are autonomous. Bilingual language interaction has been examined both in the morphosyntactic (Paradis & Genesee, 1996; Sánchez, 2003; Bullock & Toribio, 2004; Haznedar, 2010; Hatzidaki, Branigan, & Pickering, 2011) and phonetic/phonological (Paradis, 2001; Keshavarz & Ingram, 2002; Flege, 2003; Fabiano & Goldstein, 2005; Carlson et al., 2016; Carlson, 2019) domains, yet the mechanisms of bilingual language interaction are still unclear. Answering this question is paramount to creating a model of language systems that are able to account for all speakers of the world, both mono- and bilingual.

Past research has shown compelling evidence for the non-autonomy of linguistic systems in bilinguals across all linguistic domains. Behavioral studies have shown that bilinguals simultaneously activate both of their languages in perception and production, even in scenarios that strongly favor one language over the other, regardless of their age of exposure or linguistic background (Colomé, 2001; Marian & Spivey, 2003; Kroll & Tokowicz, 2005). Neuroimaging research has further supported this claim (Rodriguez-Fornells, Rotte, Heinze, Nösselt, & Münte, 2002; Marian, Spivey, & Hirsch, 2003; Martin, Dering, Thomas, & Thierry, 2009). For example, findings from Abutalebi and Green's (2007) fMRI study suggest that the two languages of a bilingual converge in the sense that they utilize the same neural pathways and connections for both languages, but that other cognitive strategies like inhibition resolve the competing forms and select one language over another. This non-monolingual-like language behavior is due to both language-internal (e.g., locus of variability) and language-external factors (e.g., age of exposure, language dominance, language use, etc.).

3.2 Language internal effects on bilingual variability

3.2.1 Types of bilingual variability

Much of the research interested in language-internal factors affecting cross-language interaction has attempted to define parts of grammar that are the most vulnerable to bilingual effects. Bilingual variability that has been observed is not random but is constrained linguistically and can be categorized as either qualitative or quantitative variability.

Qualitative variability refers to the use of features that are not instantiated in the target language. An example of this in Spanish would be using a pronominal subject as the experiencer in psychological predicates, shown below in (1). With psychological predicates in Spanish, the theme of the sentence is in subject position, meaning the verb agrees in person and number with this noun, and the indirect object pronoun is used to indicate the experiencer. Thus, the verb *gustar* must agree with *la comida*, and the person-pronoun must have dative (*me*) rather than nominative (*yo*) case. Here, the overgeneralization of a transitive structure to a non-transitive verb is an example of qualitative variability.

(1) a. *Yo gust-o la comida.
1-NOM.SG. please-1-SG. the food
I like the food.

b. Me gusta la comida.
1-DAT.SG. please-3-SG. the food
I like the food.

Quantitative variability has two types: target-like or non-target-like variability. Target-like variability, also referred to as 'inherent variation' (Lavandera, 1978), occurs in bilinguals when the structure of a language allows for optionality but the bilinguals use one option or the other in a way that differs from that of monolinguals. For example, Spanish object pronoun clitics in complex infinitival sentences can grammatically appear pre- or post-verbally, as in (2). English differs in that object pronouns invariably occur post-verbally. English-Spanish bilinguals rarely use these pronominal clitics ungrammatically, but they have been shown to use them with different rates of pre-verbal versus post-verbal placement compared to Spanish monolinguals (Davies, 1995; Montrul, 2010b; Pérez-Leroux, Cuza, & Thomas, 2011; Thomas, 2012).

(2) a. Ana la quiere comprar.
Ana $_{Clitic-3S}$ wants buy-$_{INF}$
Ana wants to buy it.

b. Ana quiere comprar-la.
Ana wants buy-$_{INF}$-$_{Clitic-3S}$
Ana wants to buy it.

Non-target-like variability is the variable distribution of grammatical features in ways that may be unattested in an adult monolingual grammar. For example, in Spanish, objects that are both animate and specific must be marked with the *'a-personal'*, a form of direct object marking (DOM), shown in (3). In (3-a), DOM is needed because the object, *Pablo*, is both animate and specific, but in (3-b), DOM renders the sentence ungrammatical because the object, *la tienda*, is not animate. In (3-c), the presence of the *a*-personal yields a semantically infelicitous reading: if Paco is searching for a new doctor, he likely doesn't have a specific doctor in mind, so DOM is unnecessary. (3-d) shows that the DOM is grammatical and felicitous because Paco is looking for a specific new doctor.

(3) a. Paco busca a Pablo. [+animate, +specific]
Paco looks for Pablo.
b. Paco busca *a la tienda. [-animate, +specific]
Paco looks for the store.
c. Paco busca #a un nuevo doctor. [+animate, -specific]
Paco looks for a new doctor.
d. Paco busca al nuevo doctor. [+animate, +specific]
Paco looks for the new doctor.

Monolingual children acquire DOM by age three (Rodríguez-Mondoñedo, 2008), however English-Spanish bilinguals (both early and late) do not reliably show control over this structure (Montrul & Bowles, 2009; Ticio, 2015). This can be attributed to the nonidentical overlap in structures between the two languages. English does not have DOM, and Spanish has sentences without DOM that are grammatical, so there is overlap between the two languages. However, this is an example of non-identical overlap because there is not a one-to-one mapping of this feature across English and Spanish.

The observation of bilinguals' variable distribution of features has led to the proposal that bilingual variability often occur when there is non-identical structural overlap between the two languages (Hulk & Müller, 2000; Müller, 2003). If language X uses A in environment 1 and B in environment 2, while language Y uses A in both environments 1 and 2, bilinguals may overextend their use of A to environment 2 in language X (i.e., incorrect contexts). In other words, one language has two structures mapping to two grammatical features, while the other language has just one structure that maps to both grammatical features.

3.2.2 Bilingual variability in phonology

Although the vast majority of the work on language internal effects in bilingualism has focused on morphosyntactic variability, it now generally accepted that the two phonological systems of a bilingual are separate but not entirely autonomous. Like in morphosyntax, the locus of bilingual variability in phonology tends to be where there is non-identical overlap between the grammatical structures of the two languages. This has been observed both in production and perception.

3.2.2.1 Bilingual variability in phonetic and phonological production

Paradis (2001) investigated bilingual phonology by looking at English-French bilingual children's truncation patterns in relation to word stress. Children often shorten, or truncate, polysyllabic words early in acquisition (e.g., 'ba.ˈna.na' → 'ˈna.na'), and generally stressed syllables and/or word final syllables are the preserved syllables. English stress is largely quantity sensitive, meaning that heavy syllables (i.e., syllables with a coda) tend to attract stress; this leads to many different word stress patterns, and, consequently, multiple different truncation patterns. French, on the other hand, has a fixed word stress that is not quantity sensitive, with stress always falling on the final syllable; this yields consistent truncation patterns across words. This study found that, in English, bilingual children did not follow monolingual truncation patterns. They did not show any sensitivity to syllable weight, but the truncation patterns did not match those of French either. This is an example of bilingual variability in phonology that is not a result of language to language transfer, and it occurs in a place of non-identical structural overlap in prosodic structure. This is evidence that bilinguals do something different than monolinguals in phonology.

Another example comes from English-Spanish bilinguals' production of voiceless stops (/p, t, k/) and their respective voice onset times (VOTs) (Flege, 1991; Thornburgh & Ryalls, 1998; Yavaş, 2002; Tetzloff & Thomas, in prep.; among others). VOT is the duration of time between the release of the stop closure and the onset of voicing of the following vowel (Lisker & Abramson, 1964; Abramson & Lisker, 1965; Cho & Ladefoged, 1999); Spanish voiceless stops have relatively short VOTs (∼0-25ms), whereas English voiceless stops have longer

VOTS (~30-90ms). These studies have found that when late English-Spanish (English L1 - Spanish L2) bilinguals produce Spanish voiceless stops, the VOT is non-native-like, with VOT durations falling between what is acceptable for Spanish versus English (i.e., transfer). Although this compromised VOT value may be an instance of direct transfer, other cross-language interactions regarding VOT have been observed. Tetzloff and Thomas (in prep.) found that early bilinguals of English and Spanish show native-like VOTs in Spanish but their English VOTs do not match those of English monolinguals. However, the effect is not one of compromise: the bilinguals' English VOTs were significantly longer than what is seen from monolinguals. Flege and Eefting (1987) also found this effect in advanced Dutch speakers of L2 English.

Of particular relevance to this dissertation, Escartín Ortiz (2005) aimed to determine which language internal factors influence the degree of /e/ prothesis in English-Spanish late bilinguals (Spanish L1 - English L2). As presented in the previous chapter, sC-clusters in English and Spanish show non-identical structural overlap, so this may be a locus of cross-language interactions due to variable distribution of features. The speakers in Escartín Ortiz's (2005) study were categorized into groups of beginner, intermediate, or advanced English proficiency. They were presented with two English sentences containing the same sC-initial word in the same position, but one of the sentences was grammatical and the other was ungrammatical. They were asked to read out loud the sentence that they believed to be grammatical. The sC-clusters included /sp, st, sk, sm, sn, sl/. After an acoustic analysis of the sC-cluster productions, the results revealed that the type of the sC-cluster had a significant effect on the production of a prothetic /e/: s+stop clusters (/sp, st, sk/) and /sl/ clusters were more likely to result in /e/ prothesis than s+nasal (/sm, sn/) clusters. This result was unexpected since /sl/ clusters have an increase in sonority, while s+stop clusters have a sonority fall. However, further analysis revealed that the preceding environment of the sC-inital word was actually a better predictor of /e/ prothesis. sC-initial words that followed a vowel were less likely to be pronounced with an initial /e/, and sC-initial words that followed a consonant or a pause were more likely to be pronounced with an initial /e/. Furthermore, beginner English speakers were most likely to produce the prothetic /e/, followed by intermediate speakers, with advanced speakers having the lowest amount of vowel insertion.

3.2.2.2 Phonotactic effects on speech perception

As discussed in the previous chapter, speakers of all languages are able to make phonological generalizations from the patterns present in their lexicon, which allows them to learn their language's phonotactic patterns and restrictions (i.e., the language specific constraint weights in an HG framework) (Coleman & Pierrehumbert, 1997; Vitevitch, Luce, Charles-Luce, & Kemmerer, 1997; Frisch & Zawaydeh, 2001; Daland et al., 2011; Albright, 2009; Bailey & Hahn, 2001; among others). Part of acquiring this phonotactic knowledge involves being able to make predictions about sequences that the speaker may have never before encountered. For instance, Daland et al. (2011) showed that English speakers judge /bd/ as better than /lb/ in onset cluster sequences, despite the fact that both of these sequences are unattested in the lexicon and are ungrammatical due to SSP and MSD violations in English. These results, which have been further supported by results of other similar studies, provide evidence that speakers of any given language have active knowledge of language universals, such as the SSP. This knowledge of language universals, in addition to the phonotactic generalizations learned from patterns in the lexicon, is then used to make judgments or distinctions between sound patterns that they have never before experienced (Davidson, 2006; Hayes, Siptár, Zuraw, & Londe, 2009; Hayes & White, 2013; Becker, Ketrez, & Nevins, 2011; Jarosz & Rysling, 2017; White & Chiu, 2017; among others).

It is well-established that phonotactic knowledge, regardless of whether it stems from language-internal patterns or from universal phonological tendencies, affects speech perception. Often times when listeners hear a sequence of sounds that is illicit in their native language, they report perceiving it as conforming to the phonotactics of their native language (Dupoux, Kakehi, Hirose, Pallier, & Mehler, 1999; Flege, 2003; Hanulikova, Mitterer, & McQueen, 2011; Lentz & Kager, 2015; among others). For instance, the phonotactics of Japanese do not permit most sequences of adjacent consonants. Dupoux et al. (1999) showed that when native speakers of Japanese hear an auditory nonce-word input like [ebzo], they perceive /ebuzo/, with an illusory, or physically non-existent, epenthetic /u/. They subconsciously insert an epenthetic /u/ into the illegal consonant clusters during perception so that what they perceive conforms to their phonotactic grammar. Likewise in English, past studies have demonstrated that English listeners exhibit a similar perceptual illusion with marked, unattested onset clusters, often perceiving an illusory, epehtetic schwa between the two onset consonants e.g., [lb] perceived as /ləb/) (Pitt, 1998; Berent, Steriade, Lennertz, & Vaknin, 2007; Berent, Lennertz, Smolensky, & Vaknin-Nusbaum, 2009). This same phenomenon has also been reported in other languages, including Korean (Kabak & Idsardi, 2007; Berent, Lennertz, Jun, Moreno, & Smolensky, 2008; Durvasula & Kahng, 2015), and French (Hallé, Segui, Frauenfelder, & Meunier, 1998) to name a few.

Spanish is no exception to this type of perceptual repair. The above predictions of the HG-GLA learner with Spanish sC-clusters demonstrates that the Spanish phonotactic system does not prefer output forms with sC-onset clusters, and when this sequence appears in the input, it is repaired with a prothetic /e/ (i,e., /sC/ → [es.C]). This analysis is not only supported by the lack of surface sC-clusters in the Spanish lexicon, but also by the fact that native speakers of Spanish reliably produce sC-cluster words with the prothetic vowel (Carlisle, 1991; Rauber et al., 2002).

Despite the salience of this phonotactic repair in speech production, this repair strategy is also active in the perception of sC-clusters by monolingual Spanish speakers: they often perceive the acoustic signal [sp] as /esp/ because their perception is filtered by their phonotactic grammars (e.g., the constraint weights along with the violations incurred by each output candidate). This leads them to map the unfamiliar signal to a familiar grammatical form (i.e., the most harmonic output candidate) (Hallé & Segui, 2003; Theodore & Schmidt, 2003; Cuetos, Hallé, Domínguez, & Segui, 2011; among others). The perceptual repair happens below the level of consciousness but makes the seemingly 'illicit' surface form licit.

Theodore and Schmidt (2003) provide experimental support for this claim: they presented monolingual Spanish listeners with stimuli from a continuum of [stib]-[estib], where the stimuli ranged from having no initial [e] to having a robust [e]. They reported that the participants detected /e/ in all of the stimuli, even when there was no acoustic evidence of this initial vowel. Cuetos et al. (2011) had similar findings and also showed that monolingual Spanish listeners detected /e/ roughly 60% of the time upon being presented with nonce words beginning with an sC-cluster. Furthermore, in a lexical decision task, Hallé and Segui (2003) showed that native Spanish listeners accept Spanish nonce words like *special*, derived from *especial*, as real words of Spanish, recovering the initial /e/ in perception. However, the same listeners rejected similar nonce words like *stuto*, derived from *astuto*, as real words of Spanish. This is interpreted as an indication that their phonotactic grammars led them to perceive /e/ as the initial illusory vowel in both types of stimuli, rather than /e/ in *e*-intial words and /a/ in *a*-initial words. Together, the results of these experiments show that phonotactics are used to filter input in perception, leading to an active illusory vowel effect in Spanish listeners upon hearing sC-clusters.

These studies and many others have looked that the effects of native phonotactics on speech perception of a foreign language within monolinguals, but the reality is that the vast majority of the world's speakers are not monolingual. Even so, the work characterizing phonotactic grammars of bilinguals is quite limited. This leads to an interesting question of how bilinguals perceive input that is licit in one language yet illicit in their other language: do English-Spanish bilinguals experience this perceptual repair of sC-clusters in Spanish even though this sequence is permitted in their English grammars?

3.2.2.3 Phonotactic effects on bilingual variability in perception

If the two languages of a bilingual have distinct phonotactic patterns, the bilinguals must maintain two separate sets of constraint rankings that make different output predictions given the same input. However, these conflicting constraint rankings are likely influenced by one another in some way (i.e., bilinguals are not two monolinguals in one).

Altenberg and Cairns (1983) performed a lexical decision task with English-German adult bilinguals in order to determine if their two phonotactic systems interacted with each other. English-German bilinguals were orthographically presented with monosyllabic nonce words of four types: nonce words legal in both English and German, nonce words not legal in English nor German, nonce words legal in English but not German, and nonce words legal in German but not English. They were instructed to determine whether or not the nonce token could be a word of English or not, and the results were compared with those of English monolinguals. They found that bilinguals differed from monolinguals in responses and in response times. When the stimuli were legal in both English and German and when the stimuli were illegal in both languages, responses were more accurate and faster than responses to other tokens and than responses from English monolinguals. Furthermore, the bilinguals were slower to reject words that were legal in just one of the two languages compared to the monolinguals. They concluded that these results support the hypothesis that bilinguals' two phonotactic systems are not autonomous.

Freeman et al. (2016) investigated whether or not English-Spanish bilinguals activate Spanish phonotactic constraints in comprehension via a primed lexical decision task. Both bilingual and English monolingual participants performed the same task, in which they first heard English cognate primes, non-cognate primes beginning with an sC-cluster, or control primes without an initial sC-cluster. They then were asked to respond to visual word and nonce word stimuli in English. The nonce words included tokens that would have been phonologically activated in the prime regardless of cognate status (e.g., cognate prime: 'stable', target: 'esteriors' vs. non-cognate prime: 'stain', target: 'esteriors'). Non-cognate and phonologically distant primes were also included (e.g., prime: 'stable', target: 'hainsale'. Real word targets and fillers were also included. Monolinguals and bilinguals showed similar accuracy in responding to nonce versus real word stimuli, although monolinguals reaction times were significantly faster. Furthermore, both monolinguals and bilinguals were more accepting of nonce words beginning with 'esC' and were primed by sC-cluster words than other prime-target pairs, and bilinguals had faster reaction times in this context. Bilinguals also showed faster responses to phonologically similar targets that were primed with cognate versus non-cognate words (i.e., faster responses for prime: 'stable', target: 'esteriors'; monolinguals did not show this effect. They argue that these results are evidence that English primes beginning with sC-clusters activate phonotactic constraints in Spanish, the non-active language, since Spanish sC-clusters are realized as esC. They also claim that these results support the notion that English cognates activate the phonological neighborhood of Spanish words with similar initial segments. This leads them to the conclusion that bilinguals can access both phonotactic constraints and phonological neighborhoods of the non-active language, supporting the hypothesis that the two phonological systems of bilinguals interact with each other.

Although Freeman et al.'s (2016) study provided evidence that both languages' phonotactic systems can be activated at one time, it relied on written target stimuli and it tested the language that is less restrictive in its phonotactic rules (i.e., English). Using visual stimuli in investigating phonological and phonotactic behavior can be problematic because it has been shown that different parts of the brain are activated when responding to visual versus auditory stimuli (Strand, Forssberg, Klingberg, & Norrelgen, 2008). Furthermore, this design relies on the literacy of participants, which may not be equal, and nonce word recognition has been shown to be mediated by the target's orthographic neighbors both in monolinguals (Andrews, 1992; Ziegler, Muneaux, & Grainger, 2003) and bilinguals (Altenberg & Cairns, 1983; Van der Hulst, 1984).

Carlson et al. (2016) and Carlson (2018, 2019) addressed these methodological concerns and auditorily tested if the cross-language interaction in bilingual phonology and phonotactics is present when the grammars are in conflict, as is the case with sC-clusters in English and Spanish. Spanish does not allow sC-clusters and repairs them by inserting a prothetic /e/, while English does allow this structure. When a bilingual hears a single speech signal consisting of [sp], for example, it could be interpreted in two ways depending on the active language.

The study reported in Carlson et al. (2016) focused on this conflict in the phonotactics of English and Spanish by testing if knowledge of English affected the tendency of Spanish bilingual speakers to perceptually repair sC-clusters with the prothetic /e/ that has been reported in studies of monolingual Spanish perception. In other words, they wanted to see if there was an illusory vowel effect in bilinguals as in monolinguals of Spanish, described above in Section 3.2.2.2. Participants in this study included 83 early English-Spanish bilinguals; all participants had robust exposure to both languages by six years of age, and 43 of them reported that Spanish was their dominant of the two languages. For the first experiment, participants were asked to listen to iambic, Spanish-like nonce words that began with sC-clusters (e.g., /espid, asmid/, etc.). The initial vowel of these stimuli was divided into five sections, with the first segmentation deleting the vowel entirely; the following sections progressively added 2.5 periods of the vowel to the stimuli, such that there were six instances of each nonce word, each with a different length initial vowel. After hearing the stimuli, half of the participants were asked to decide whether or not they perceived an initial /e/, and the other half of the participants were asked to decide whether or not they perceived an initial /a/. When the vowel was at its longest, both groups of participants performed at ceiling, being highly accurate in their detection of both /e/ and /a/, respectively. When the vowel was absent entirely, the participants detected the presence of /e/, and when the /a/ was short, they tended to mis-identify the presence of /e/. These results show that bilinguals, like monolinguals, are susceptible to the illusory vowel effects induced by Spanish phonotactics. Furthermore, the bilinguals who self-reported that Spanish was their dominant language were more likely to misperceive an initial /e/, suggesting that proficiency in English somewhat modulates the illusory vowel effect.

In a second experiment, Carlson et al. (2016), 32 early English-Spanish bilinguals and 15 Spanish monolinguals performed an AX (same-different) discrimination task. The AX task involved hearing two stimuli and responding if they were identical or not. Using the same nonce words, they created sets of four tokens each, where one had a short [e] of 2.5 periods, one had a long [e] of 10 periods, one had a short [a] of 2.5 periods, and one had a long [a] of 10 periods (e.g., [esbid] with short and with long vowel, [asbid] with short and long vowel). Each AX pair was composed of two stimuli from the same set. Participants were divided into three groups: Spanish monolinguals, English-dominant bilinguals (n=17), and Spanish-dominant bilinguals (n=15). All participants were highly accurate in their discrimination, with no differences in accuracy observed between groups. However, there was an interaction of language group and vowel length that revealed that Spanish monolinguals showed a stronger illusory vowel effect than bilinguals, and within Spanish bilinguals, the Spanish-dominant participants showed a stronger illusory vowel effect than their English-dominant counterparts. These results, too, show that knowledge of English mitigates the effect of Spanish phonotactics on speech perception to some degree.

Following this study, Carlson (2018, 2019) further investigated this effect of English on Spanish phonotactics in early and late English-Spanish bilinguals by investigating if more recent use of English increased the effect of English phonotactics in Spanish perception of sC-clusters. The first experiment of the 2018 study was a replication of the first experiment of the Carlson et al. (2016) study, the vowel detection task. The 102 English-Spanish bilingual participants were early bilinguals, having learned both English and Spanish from an early age. All participants were highly fluent in both languages and were asked to rate their language dominance as Spanish-dominant, English-dominant, or balanced. The results from this task mirrored those of the previous study, showing that bilinguals do perceptually repair sC-clusters in Spanish, perceiving the prothetic vowel particularly when there was no acoustic evidence of a vowel. He noted that these early bilingual participants reported perceiving the prothetic vowel 22% of the time when there was no vowel present, which is a much lower percentage than the 56% that Cuetos Vega, Glez Nosti, Barbón Gutiérrez, and Brysbaert (2011) reported in a monolingual Spanish population. Carlson (2018) also showed differences between groups of participants based on their language dominance. When partial acoustic information was present (i.e., a brief initial /a/ or /e/), the early bilinguals in Carlson (2018) also reported perceiving the prothetic vowel regardless of the vowel quality, but the Spanish dominant participants were more likely to do so than the balanced or English dominant listeners.

In the second experiment, 32 of the early bilinguals from the previous task, along with 15 monolingual Spanish listeners from Spain, participated in a similar AX task that was focused on discriminating between vowel durations and vowel qualities in syllables with sC-clusters. The stimuli included Spanish-like nonce words beginning with /esC/ and /asC/ (e.g., /espid, aspid/); there were two versions of each nonce-word, one with a long vowel and one with a short vowel (shortening the vowel made the vowel quality more ambiguous). All possible combinations of these nonce words were included as stimuli pairs, including pairs where both nonce words were identical. For these identical pairs, the monolinguals reported that the two words were the same 94% of the time, while the Spanish dominant and English dominant bilinguals reported that they were the same 92% and 86% of the time, respectively. As English proficiency increased, participants were less likely to identify the identical pairs as being the same. For the critical trials (short /espid/ - long /espid/, short /aspid/ - long /espid/, short /espid/ - long /aspid/, short /aspid/ - long /aspid/)

both bilinguals and monolinguals showed evidence of a perceptual repair in sC-clusters, but the English dominant group of bilinguals did so to a significantly lesser extent. Across all participants, there were fewer ‘wrong’ responses when the long vowel was /a/ rather than /e/, but this effect was weaker for all bilinguals than for monolinguals. The results of these tasks led Carlson to conclude that bilinguals are vulnerable to the perceptual illusion of perceiving a prothetic /e/ in sC-clusters so that the nonce words are phonotactically licit in Spanish, but that knowledge of English dampens this effect. Thus, knowledge of English, a less restrictive phonotactic system, reduces the perceptual repairs in Spanish, a more restrictive phonotactic system, ultimately providing support for the notion that bilinguals’ two phonologies interact with one another.

This study was followed by an additional study, Carlson (2019), which implemented the same AX task as the previous study but with 32 English-Spanish late bilinguals, all having robust exposure to English beginning around age 14. An additional variable was also added: half of the bilinguals interacted with the experimenter only in English prior to the AX task (English-switch group), while the other half only interacted with the experimenter in Spanish during this time (no-switch group). Fourteen Spanish monolinguals from Spain were also tested.

For the AX task, response accuracy was at ceiling for trials where the two stimuli were identical and where they both had long vowels that were different in quality. When the nonce words had differing short vowels, monolinguals responded that they were the same at a higher rate than chance, whereas bilinguals performed at chance. When the vowels differed but the longer vowel was /a/, all participants showed high accuracy, and when the longer vowel was /e/, all participants performed around chance. However, bilingual responses were significantly slower in this condition. The author suggested that this was because both languages favor discrimination of /asC/, but only English favors the discrimination of /esC/. The longer reaction time thus reflects the processing cost involved in the inhibition of the English phonotactic system. Furthermore, the English-switch bilinguals had significantly longer reaction times than the no-switch bilinguals, which suggests that prior English use increases the strength of the English phonotactic rules in Spanish.

Carlson (2019) also performed a lexical decision task in Spanish with the 14 late English-Spanish bilinguals (Spanish L1, English L2) and seven Spanish monolinguals; bilinguals were again split into English-switch and no-switch groups, as in the first experiment in this study. Stimuli included real words with beginning with an initial vowel and sC sequence (e.g., *escuela, astuto*), nonce words lacking the initial vowel (e.g., *scuela, stuto*), and nonce words with the wrong intial vowel (e.g., *ascuela, estuto*). Responses for real words and for wrong-vowel words were highly accurate across participant groups. When there was no vowel for words that traditionally begin with /e/ (e.g., *scuela*), all participant groups consistently reported that the stimulus was a real word, indicating that they had perceived the missing initial /e/. In other words, the illusory vowel effect was active for both monolinguals and bilinguals. When the initial-vowel-less words were not derived from /e/-initial words but from words with a different initial vowel, performance was at chance. The English-switch bilinguals had slower responses than the no-switch bilinguals, particularly for the no-vowel trials, but this difference was not significant. These results suggest that bilinguals are slower to process words in which the two languages' phonotactics are conflicting because the auditory stimulus activates both sets of phonotactic constraints which results in two competing representations that the speaker must choose between, whereas monolinguals do not encounter this phonotactic competition. This shows that the constraint sets of both languages are active at the same time and are integrated.

Together, these studies provide evidence of bilingual variability in the phonotactic systems of English-Spanish bilinguals and that language dominance affects performance. However, many of results reported were at levels of chance in accurate perception of sC-clusters, which may be able to be accounted for by factors other than language dominance. These studies did not look at the language internal factor of sC-cluster type, and although both early and late bilinguals were tested, the results of these two groups were not directly compared.

3.3 Language-external effects on bilingual variability

In addition to understanding how the grammatical structure of a language affects bilingual variability, a second goal in bilingualism research is to understand what non-linguistic factors affect bilingual variability. Some of these language-external factors include age of exposure to bilingualism and language proficiency.

3.3.1 Age and bilingual variability

Age of exposure to a language has been shown to be a factor that affects language representations and use in bilingualism. Studies in language acquisition have shown that children who acquire their languages earlier in childhood have different language outcomes than those who acquire an L2 later in childhood or in adulthood. Children who are exposed to bilingualism earlier and have robust exposure to both languages through adulthood tend to have more similar language processes and outcomes to monolinguals. As an individual ages, the process and outcome of language tend to be more variable compared to earlier acquired languages.

Speakers who begin learning an L2 before puberty tend to resemble native speakers more so than those who begin learning their L2 in adulthood in both production and perception across linguistic domains (Johnson & Newport, 1989; Flege, 1991; McDonald, 2000; among others). Age-related effects in L2 learning affect all linguistic domains, although the degree of such effects can vary between the phonetic/phonological domain and the morphosyntactic domain. Going forward, speakers who learn both languages before puberty will be considered early bilinguals (Montrul, 2013), which can be further broken down into early child and late child bilinguals: early child bilinguals are those that have had robust exposure to both languages by age six (school age), and late child bilinguals are those whose exposure to the L2 began between six and 14 years. Bilinguals whose age of exposure to the L2 is 15 or greater are considered to be late bilinguals or L2 speakers (Montrul, 2008).

When looking at bilingualism in the U.S., heritage speakers are an important group of early bilinguals. Valdés (2000: 1) defines a heritage speaker as an individual who was 'raised in a home where a non-English language is spoken, who speaks or merely understands the heritage language, and who is to some degree bilingual in English and the heritage language'. In the context of Spanish in the U.S., Spanish is a heritage language and this study will adopt a linguistic approach to heritage speakers (i.e., one must have some degree of receptive or productive linguistic competence in Spanish). For this paper, all early child bilinguals are in fact heritage speakers.

Most researchers agree that for morphosyntax learning the L2 before the age of 12-16, or around puberty, is important to gain native-like proficiency (Patkowski, 1980; Johnson & Newport, 1989, 1991; Birdsong, 1999; Birdsong & Molis, 2001). A pivotal study in the field of L2 morphosyntactic acquisition is that of Newport and Supalla (1987). With 46 L2 English speakers (L1 was either Chinese or Korean), they administered an English grammaticality judgment test that included morphosyntactic structures with a variety of complexities to individuals with ages of acquisition varying from three to 39; they reported a decline in the level of L2 morphosyntax based on age of exposure to English until about 16 years, at which point no clear age-related effects were evident. This study has been replicated several times with relatively consistent results (Johnson & Newport, 1991; Johnson, 1992; DeKeyser, 2000; Birdsong & Molis, 2001; Hartshorne, Tenenbaum, & Pinker, 2018; among others).

An obvious trait of a late acquired language is the presence of a foreign accent, or notable differences in the pronunciation (i.e., phonetics/phonology). Foreign accents have been attributed to age of exposure, namely that the earlier one is exposed to the language, the more likely they are to sound like a monolingual speaker (Flege et al., 1995, 1999; Flege, 1999). Within the phonetic and phonological domain, it is believed that age threshold is younger than for morphosyntax, around six years of age, suggesting that this linguistic domain is actually more vulnerable to age-related effects of bilingualism (Flege & Eefting, 1987; Flege, 1991; Flege et al., 1999; Piske, MacKay, & Flege, 2001).

Flege and colleagues were among the first and most influential researchers to study these age-related phonetic effects in bilingualism. For example, Flege et al. (1995) investigated the relationship between age of exposure to an L2 and foreign accent with a group of 240 Italian immigrants to Canada. The ages of exposure to English among the participants ranged from two to 23 years, and they had lived in Canada for an average of 32 years. These participants, along with 24 native English controls, were recorded reading English sentences. Word-initial, -medial, and -final consonants were then rated via a binary judgment task by native English listeners. The results revealed that the Italian immigrants' age of exposure to English was significantly predictive of their perceived foreign accent. A foreign accent was perceived from L2 speakers who were exposed to English at as early as nine to 11 years old. Similar results were reported with Korean-English bilinguals in the United States (Flege et al., 1999; Yeni-Komshian, Flege, & Liu, 2000).

Like phonetic production, phonetic and phonological perception have shown strong age-related effects as well. At birth, infants have the ability to discriminate all phonetic contrasts (Eimas, Siqueland, Jusczyk, & Vigorito, 1971). For example, Werker, Gilbert, Humphrey, and Tees (1981) made the observation that six month old infants in an English monolingual environment were able to discriminate non-native consonant contrasts of Hindi, while adult English monolinguals were not.

However, this ability to discriminate all sound contrasts is refined before one year at which point only native-language sound contrasts are perceived (Werker & Tees, 1984; Best et al., 1994; Best, McRoberts, & Goodell, 2001; Kuhl et al., 1997, 2006; among others). One piece of evidence supporting this idea of phonetic/phonological perception being constrained by one's native language comes from Bosch and Sebastián-Gallés (2003a,b) who tested infants that were Spanish monolinguals, Catalan monolinguals, and Spanish-Catalan bilinguals on sound contrasts that is contrastive in Catalan but not in Spanish. At four months of age, all three groups were able to discriminate between the vowels, but by 12 months only Catalan monolingual and the Spanish-Catalan bilingual infants were able to discriminate these contrasts. It should be noted that contrast discrimination trajectory differed between the monolingual and bilingual infants but that by 12 months the outcomes were equivalent. Other studies have used neurolinguistic methodologies to also test the sound discrimination abilities of infants and similarly found that the ability to discriminate non-native contrasts disappears around the age of 11 months (Rivera-Gaxiola, Silva-Pereyra, & Kuhl, 2005; Ferjan Ramírez, Ramírez, Clarke, Taulu, & Kuhl, 2017).

3.3.1.1 Age-related effects on the L1 in bilinguals

Although many studies have compared the speech of early bilinguals to monolingual speakers of their L2 (Mack, 1989; MacKay, Flege, Piske, & Schirru, 2001; Flege, 1991; among others), fewer studies have compared their speech to monolinguals of their L1 (Mack, 1990; Mack, Bott, & Boronat, 1995) and to monolinguals of both languages (Baker & Trofimovich, 2005; MacLeod & Stoel-Gammon, 2009). This is important because it provides insight on if there are any linguistic consequences of being an early bilingual in the L1 in addition to the L2.

Some studies have shown that there is influence in the outcome of early bilinguals in one or both languages, despite speakers being immersed with both languages before puberty. This bilingual variability results in the bilinguals' speech differing from that of monolinguals, even though both groups are considered to be native speakers. One such example comes from Baker and Trofimovich (2005); they showed a bidirectional L1-L2 influence in the production of vowels in English-Korean early bilinguals. The earlier the speaker was exposed to the second language, the more distinct their L1 and L2 vowel sounds were, despite these vowel categories overlapping for monolinguals of each language. Thus, the bilinguals over-differentiated the contrast in both languages, compared to monolinguals. A similar finding was reported for English-Italian early bilingual vowels (Flege, Schirru, & MacKay, 2003).

Liu and Cao's (2016) meta-analysis of neuroimaging studies in bilingualism demonstrates that this age-related bilingual variability extends to physiological effects in language comprehension. They found that late bilinguals implemented additional cortical regions in processing the L2 compared to earlier bilinguals, which suggests that using the L2 is more cognitively demanding for later learners. Furthermore, the age of exposure to the L2 also had a physiological effect on the processing of the L1. When using the L1, the early bilinguals showed increased activation in the left temporal region compared to the late bilinguals, which may be attributed to a higher degree of co-activation between the two languages (i.e., the earlier the age of exposure, the more likely both languages will be activated simultaneously).

These results show that age of exposure to bilingualism not only affects outcomes in the L2 but also may have an effect in the L1. The earlier the exposure to the L2, the more likely the individual will show bilingual variability in the L2 and also in the L1.

3.3.2 Language proficiency in bilingualism

In addition to age of exposure to bilingualism, language proficiency has been shown to affect phonetic production in bilingual and L2 speech (Flege, 1988; Thompson, 1991). Language proficiency can be defined as a speaker's ability to 'correctly use the rules' of that language in phonological, syntactic, and lexical/semantic systems (Burt & Dulay, 1978); however, in a generative framework, the passive competence or implicit knowledge of the speaker, rather than their productive abilities, is potentially a better indicator of proficiency when proficiency is understood to be represented by the ideal monolingual grammar (Chomsky, 1965). Although an earlier age of acquisition often correlates to higher language proficiency, this is not always the case in bilingualism, particularly for heritage speakers who may have had Spanish as their L1 but have more robust exposure to the dominant language of the society, English. Therefore, it is important to operationalize a measure of proficiency in studies of bilingualism, as it may be a factor that can be used to predict bilingual variability.

Relative language proficiency in bilinguals' two languages (i.e., dominance) has been shown to have an effect on bilingual variability (Yip & Matthews, 2000, 2005; Pérez-Leroux, Cuza, & Thomas, 2011; Carlson et al., 2016; Castilla-Earls, Pérez-Leroux, Martinez-Nieto, Restrepo, & Barr, 2020; among others). Yip and Matthews (2000, 2005) investigated the relationship between language dominance and bilingual variability. In analyzing a corpus of two Cantonese-English early bilingual children, they showed that null objects, which are ungrammatical in most cases in English but always grammatical in Cantonese, were more often produced in English by the child who was more dominant in Cantonese.

Following this work, Pirvulescu, Pérez-Leroux, Roberge, Strik, and Thomas (2014) looked at the role of language dominance in English-French early bilinguals' production (or omission) of direct objects. Each subject was rated as completely fluent, somewhat fluent, or not fluent in both languages, and these scores were combined with an indicator of the home language of the child and the community language; there were eight French dominant, ten English dominant, and ten balanced bilinguals. Their results showed that balanced bilinguals showed higher rates of object omission than English-dominant bilinguals in English and French-dominant bilinguals in French. This suggests that bilinguals show less variability in their dominant language.

Language dominance was also shown to affect Spanish clitic placement and omission in English-Spanish bilinguals (Pérez-Leroux et al., 2011). The children in this study were assigned a language dominance score based on parental ratings and narrative complexity, and they were asked to repeat Spanish sentences with both pre-verbal and post-verbal object clitics. Bilinguals in general were more likely than their monolingual counterparts to reposition the clitic (never ungrammatically), but English-dominant bilinguals were even more likely to do so than Spanish-dominant bilinguals and they also omitted preverbal clitics more often than the Spanish dominant bilinguals. This is another example of language dominance affecting bilingual variability in morphosyntax.

Within the domain of phonology, the Carlson et al. (2016) study discussed previously also found that language dominance affects bilingual variability. In a perception task aimed at examining the illusory vowel effect in sC-clusters in Spanish, early English-Spanish bilinguals were asked to self-assess if they were Spanish-dominant, English-dominant, or balanced. All English-Spanish bilinguals exhibited weaker illusory vowel effects than the Spanish monolinguals,but Spanish-dominant bilinguals were more susceptible to the illusory vowel effect than their English-dominant counterparts. These results, together with those of morphosyntax, suggest that being dominant in the non-target language yields more variability in the target language.

The effect of language dominance on bilingual variability does not always go in this direction though; sometimes being dominant in the target language can also result in bilingual variability in the target language as in Castilla-Earls et al.'s (2020) study. They examined the effects of English proficiency on the production of Spanish articles and object clitic pronouns in English-Spanish bilingual children. Language dominance was determined by the Spanish English Language Proficiency Scale (Smyk, Restrepo, Gorin, & Gray, 2013), which assesses syntactic complexity, verbal fluency, and lexical diversity via story retell tasks, and the Structure Photographic Expressive Language Test (Dawson, Stout, & Eyer, 2003), which also assess morphosyntactic abilities via a picture description task. They found that children with lower English proficiency performed worse on Spanish tasks compared to children who had higher English proficiency or were monolingual Spanish speakers. This ultimately suggests that relative balance of proficiency may not always interact with bilingual variability in the most obvious way, but that language proficiency should be considered when studying bilingual variability.

3.3.2.1 Operationalizing bilingual dominance

It is clear that language proficiency or bilingual dominance is an important factor in predicting cross-language interactions, but there are many different ways that this can be measured, each having its own advantages and disadvantages.

In studies of bilingualism and L2 acquisition, there are two ways language dominance is determined: independent linguistic measures and participant self-ratings (or parental ratings). Self-ratings have been widely used as a way to measure one's language abilities (MacIntyre et al., 1997; Shameem, 1998), particularly in studies of bilingual phonology (Freeman et al., 2016; Freeman, Blumenfeld, & Marian, 2017; Carlson et al., 2016; Carlson, 2018, 2019), and have been shown to correlate strongly with various tasks that probe at independent linguistic measures such as morphosyntactic complexity (Ortega, Iwashita, Rabie, & Norris, 1999; Flege, MacKay, & Piske, 2002; Chaudron, Prior, & Kozok, 2005; Wu & Ortega, 2013; Gaillard & Tremblay, 2016; Prentza, Kaltsa, & Tsimpli, 2019; D. Anderson, 1980; Bachman, 1982; MacIntyre et al., 1997; Ross, 1998; Montrul, 2006). For self-ratings, participants are asked to report how well they feel that they speak, comprehend, read, and write in each language on a numerical scale, typically from 1-5 or 1-7, with the lowest value representing little to no proficiency and the highest value representing native or native-like proficiency in that category.

One of the most common forms of self-assessment is the Language Experience and Proficiency Questionnaire (LEAP-Q) (Marian, Blumenfeld, & Kaushanskaya, 2007). The LEAP-Q was developed in the context of bilingualism language studies with the view that L2 acquisition depends not only on the age of exposure to the L2 but also to many experiential, or language-external variables. This questionnaire aims to gather information about social factors that have been shown to influence L2 acquisition and proficiency in late bilinguals, like language preference, age of acquisition, mode of acquisition, prior language exposure, and current language use in various settings. These factors help determine a bilingual's language proficiency, dominance, and preference.

The Language and Social Background Questionnaire (LSBQ) (J. Anderson, Mak, Chahi, & Bialystok, 2018) is a similar assessment to gather information on the linguistic and social profiles of bilinguals. The LSBQ also collects information about the age of exposure to bilingualism, linguistic and social upbringing of the participant and their family members, current proportion of use of each language, and self-rating measures of language proficiency across speaking, listening, reading, and writing domains. This questionnaire differs from the LEAP-Q in that the authors have created a formula that gives a score of proficiency in each language, as well as a composite score of bilingual dominance.

3.4 OT grammars in bilingualism

As evidenced by the numerous studies discussed in this chapter, there are differences in the phonological systems of bilinguals compared to monolinguals of either language. Nonetheless, current theories of generative grammar like OT fail to account for bilingualism in that they do not straightforwardly allow for any predictable bilingual variability.

It is widely accepted that bilinguals have two separate grammars that interact rather than one shared grammar. For the phonological domain in an OT framework, this idea presumes that bilinguals have a language-specific constraint ranking (or weighting in an HG framework) of universal constraints for each language. Because the two grammars utilize the same constraint set, the phonological grammars should be comparable to each other and thus plausible that they would interact.

For example, Lleó (2002) and Lleó, Kuchenbrandt, Kehoe, and Trujillo (2003) show that the phonological systems of German-Spanish bilingual children differ from those of monolingual children in each language both in development of prosodic structure and acquisition of codas. They offer OT as a possible theoretical framework that can account for this bilingual variability due to its reliance on a universal constraint set. They discuss how Spanish is more restrictive than German in what it allows for coda consonants, yet German monolinguals acquire codas earlier than Spanish monolinguals. This is justified in an OT system by the role of input: though a markedness constraint against codas is ranked higher in Spanish (i.e., the language is more restrictive), German children have more input of syllables with codas which allows them to rerank this constraint faster. The German-Spanish bilinguals, however, acquire Spanish coda consonants earlier than the Spanish monolinguals, which is explained by the increased input of codas from the German input. Although this shows how input from one language can affect the constraint ranking of another language and is explained through the lens of OT, they make no proposals as to how the architecture of OT can systematically accommodate this observation.

Along these same lines, Shooshtaryzadeh et al. (2015) discusses the potential benefits of OT as a theoretical framework for looking at phonological disorders in bilinguals. The universal constraint set lends itself to the emergence of the unmarked across languages (Hancin-Bhatt, 2008). This helps explain why phonological disorders across languages have similar patterns. She expresses that OT is particularly useful in examining phonological disorders in bilingual children because a phonological error that is present in both of the child's languages can be explained by an error or inconsistency in the ranking of a particular constraint in both of the phonologies. But again, no suggestion is offered as to how to operationalize this concept in practice.

In trying to account for this interaction between the two competing constraint rankings of bilinguals, Gonzales-Diaz (2006) proposes markedeness constraints that, in an OT framework, help select which language is preferred in contexts of code-switching and code-mixing. These constraints are based on the activation and inhibition of the languages based on the social context: for example, she proposes constraints that prefer the dominant language of the addressee and that prefer the dominant language of the environment (i.e., in Spain, this constraint would only be satisfied by outputs in Spanish). Thus, for any given input, both languages generate candidate sets that compete against each other to win, and the ranking of these contextual markedness constraint then ultimately determine the optimal output.

Muysken (2013) similarly proposes that the two languages of a bilingual also interact depending on the social circumstance, which affects when and how code-switching is implemented. He uses OT to try and account for this variability in bilinguals and proposes markedness constraints that vary based on context but are distinct from those of Gonzales-Diaz (2006). His markedness constraints include ones that disprefer using the L1, that disprefer using the L2, and that disprefer code-switching in general. Additional markedness constraints are proposed that differ based on the context, including power relations between interlocutors, solidarity, pragmatic face, and perspective. These interact with a faithfulness constraint that prefers candidates that communicate the speaker's interpretation of their utterance. He shows how this model is used to predict which language is used in code-switching contexts. The author notes that one problem that this model does not address is that it assumes that the L1 and L2 proficiencies are asymmetric, even though there are many bilinguals who are 'balanced', with equal proficiency in both of their languages. Another consideration that the model should make is that many of the linguistic and pragamtic constraints of the model not only interact with each other but also with things like the perceived similarity between the two languages, the prestige or status of each language, the relative proficiency the speaker has in each language, and the speaker's attitude toward each language. This model has been implemented in other studies of code-switching and code-mixing (Bhatt & Bolonyai, 2011; Kheder & Kaan, 2021).

Although these proposals can predict which language a speaker may opt to use in a bilingual context, they do not straightforwardly address how a constraint based system can predict bilingual variability at the level of phonology.

3.5 Summary of bilingual effects

Decades of past research have shown that bilinguals are not two monolinguals in one, but rather bilinguals show systematic differences in their language competence, referred to as bilingual variability. For phonology, bilingual variability has been explored with sC-cluster phonotactics in English-Spanish bilinguals. Results of past studies have shown that English-Spanish bilinguals differ from Spanish monolinguals in their perception of sC-clusters in Spanish based on their language dominance. These studies have not taken into account the language-internal variable of sC-cluster type (i.e., s+stop versus s+liquid) or the language-external factor of age of exposure to bilingualism. Furthermore, there have been proposals for how constraint-based models of grammar like OT can accommodate bilingual language in terms of pragmatics and code-switching, but this type of model is not able to straightforwardly predict systematic variability in phonotactics. The experimental studies described in the following chapters aim to address the variables of sC-cluster type, age of exposure to bilingualism, and relative language dominance, followed by a discussion of how OT-type grammars can possibly be adapted to account for bilingual variability in phonotactics.

CHAPTER 4

RESEARCH QUESTIONS AND HYPOTHESES

Past analyses of sC-clusters in Spanish have argued for a uniform structure for all types of sC-cluster, regardless of the sonority of the second consonant. Because surface sC-clusters are illicit in Spanish, they are repaired with a prothetic /e/, which resyllabifies the sC-cluster across the syllable boundary (sC → es.C). However, in Chapter 2 the HG-GLA, a phonological learning algorithm, predicted that s+stop and s+liquid clusters do not have parallel outputs. The HG-GLA predicted that s+stop clusters are more likely to be 'repaired' with /e/ prothesis and that the syllabification of the cluster may vary (es.C versus e.sC). On the other hand, it predicted that s+liquid clusters enforce /e/ prothesis only half of the time but that the syllable structure is always taughtosyllabic (.sC versus e.sC). The first goal of this experimental study is thus to test the language internal factor of sC-cluster type. If the segmental and suprasegmental structure of sC-clusters differs based on the sonority of the second segment, observable differences in perception are expected, as was predicted by the HG-GLA. If s+liquid clusters are always tautosyllabic, the prothetic /e/ is not phonotactically necessary to make the structure grammatical, and if this is the case, variability in s+liquid clusters should be greater than for s+stop clusters (i.e., Spanish listeners will rely less on the prothetic /e/ for accurate perception of the sC-cluster).

Perception of sC-clusters may also yield different behaviors in Spanish monolingual versus English-Spanish bilingual speakers, especially since the syllable structures of s+liquid clusters predicted by the HG-GLA match those of English. Past studies in this area have shown evidence of bilingual variability in the perception of sC-clusters, but the sonority profile of the sC-cluster has not been taken into account (Carlson et al., 2016; Carlson, 2018, 2019). The second goal of this study is to investigate if any variability related to the type of sC-cluster differs based on the language profile of the speaker (i.e., monolingual versus bilingual).

Additionally, in the Carlson et al. (2016) study the bilinguals had been exposed to bilingualism at an early age, around six years old, whereas in the Carlson (2019) study, the participants were late bilinguals, having their first exposure to bilingualism around age 14. It is well known that age of exposure to bilingualism has a strong effect on the language outcomes (Birdsong, 1992; Flege et al., 1999; Flege, 1999, 2003; Flege et al., 1995; MacKay & Flege, 2004; DeKeyser, 2000; Ioup, 2008; among many others), so these results do not paint the whole picture. Therefore, the third goal of this dissertation is to investigate the age-related effects on bilingual variability within the domain of bilingual phonotactic systems.

The results of past studies have also taken language dominance into account, but only based on a self-assessment question of 'Which language are you more comfortable in?'. The fourth goal of present study is to look at proficiency effects on bilingual variability by creating a proficiency score for each language, based on self-ratings of both verbal and listening proficiency in each language. These scores can then be used to create a dominance score that can be used to analyze data from tasks of sC-cluster perception.

The final goal of the dissertation is to then discuss how the experimental results of the English-Spanish bilinguals can be used to adapt current models of phonological theory to accommodate bilingual variability.

To summarize, this dissertation aims to add to address the following questions:

1. **Question 1: Language internal variable.** Do native speakers of Spanish, both monolingual and bilingual, show differences in the perception of sC-clusters based on the outputs predicted by the HG-GLA in Chapter 2? In other words, does the perception of s+stop clusters differ from that of s+liquid clusters?

   Question 2: Language profile. Does being monolingual versus bilingual result in measurable differences in the perception of s+stop and s+liquid clusters? Are bilinguals less likely to rely on the prothetic /e/ in perception of sC-clusters, and particularly for s+liquid clusters where the structure parallels that of English?

Question 3: Age of exposure to bilingualism. Does the age of exposure to bilingualism affect the perception of sC-clusters? Does a later age of exposure to English result in more monolingual-like behavior?

Question 4: Language dominance. Does language dominance as measured by self-ratings of language proficiency affect the variability observed in sC-cluster perception?

These questions are important not only for the fields of bilingualism and phonology, but they are also useful for applied domains such as educational and clinical settings. In these settings, the language development and outcomes of bilinguals is typically compared against that of monolinguals, despite the fact that the former group is much larger than the latter group. This can be problematic due to different acquisition trajectories and outcomes between the two groups of speakers. Understanding more about the bilingual variability and cross-language interaction that exists for bilingual speakers can help researchers to create specific language standards and targets for bilingual speakers. Additionally, the results of this study can be used to expand current models of generative phonological theory (e.g., Harmonic Grammar), which have been developed to account only for monolingual grammars.

In summary, the goal of the present dissertation is to build on the literature discussed by further investigating the nature of bilingual variability, particularly in bilingual phonology. The results of multiple behavioral experiments provide empirical data that address the role of both language-internal and language-external factors on variability in the phonological systems of monolingual and bilingual speakers of Spanish.

CHAPTER 5

METHODOLOGY

In order to answer the research questions, two experimental tasks were implemented: an AX same-different task and a nonce word grammaticality judgment task. Both tasks were hosted online using the PennController for IBEx platform (Zher & Schwarz, 2018), an extension of Internet Based Experiments Farm (IBEx Farm) (Drummond, 2013).

5.1 Participants

A total of 47 participants were included in this study, all of whom are native Spanish speakers (i.e., all have Spanish as an L1). The mean age of participants was 34 years, with the youngest being 18 and the oldest being 66. All but six participants lived in the U.S. at the time of testing, and four were born in the continental U.S. Other places of origin included Puerto Rico (15), Mexico (9), Ecuador (8), Colombia (3), Venezuela (3), El Salvador (2), Spain (2), and Cuba (1).

All participants had Spanish as their L1, and 11 were monolingual Spanish speakers. Of the 36 bilinguals, the age of exposure to English varied from birth to adulthood and they were grouped based on age of exposure to English. If English exposure began before six years of age, they were classified as early child bilinguals (n=13). Bilinguals whose first English exposure began between six and 13 were classified as late child bilinguals (n=8), and those whose exposure to English happened after this were considered adult bilinguals (n=15).

Participants were also grouped based on their proficiency in both Spanish and English. Based on the criteria for these groupings, described in detail in Section 5.2, 12 bilinguals reported equal native-like proficiency in both languages, 18 were Spanish-dominant, and six were English-dominant.

5.2 Procedure and task

All participants signed informed consent forms in Spanish prior to beginning the experiment, which was approved by the University of Massachusetts IRB. All communication with participants was in Spanish.

First, participants completed a modified version of the Language and Social Background Questionnaire (LSBQ) (J. Anderson et al., 2018) through an interview with the tester (Appendix A). The LSBQ is an assessment that collects information on the linguistic and social profiles of bilinguals. It helps gather information about the age of exposure to each language, the linguistic and social upbringing of the participant and their family members, and the current proportion of use of each language. It also includes self-rating measures of language proficiency across speaking and listening for each language. This type of language background questionnaire with self-ratings has been widely used in past studies of bilingualism and have been shown to be reliable summaries of participants' language abilities (Bachman & Palmer, 1989; MacIntyre et al., 1997; Ross, 1998; among others).

As part of this questionnaire, participants were asked to rate both their speaking and comprehension abilities from one to five for both Spanish and English. The average of these two scores for each language was then used as way to operationalize proficiency. If the English score was greater than the Spanish score, they were classified as English-dominant; if the Spanish score was greater than the English score, they were classified as Spanish-dominant; and if both scores were equal, they were classified as balanced.

The participants then completed a same-different AX discrimination task and a nonce word judgment rating task, and order of these tasks was counterbalanced between participants.

5.3 Experiment 1: AX discrimination

The first experiment was a replication of the AX discrimination task performed in Carlson's (2016, 2018, 2019) studies. This task was designed to determine if English-Spanish bilingual listeners perceive an illusory /e/ even when the acoustic information provides very little, or ambiguous, information about the actual vowel quality. Unlike a vowel-detection task (e.g., "Does this word begin with /e/?") an AX task does not require as much metalinguistic knowledge, and so it provides a better view of the lower-level perception tendencies of the bilingual listeners.

5.3.1 Stimuli

A male native speaker of Puerto Rican Spanish was recorded saying nonce words with a high quality Zoom H5 recorder and a Shure head-mounted SM10A microphone. The nonce words were of the form of the form VsCid, which included the following: /espid/, /esfid/, /esmid/, /eslid/, /aspid/, /asfid/, /asmid/, /aslid/. Two stimuli were created from each of the nonce words by removing portions of the initial vowel. The first stimulus from each nonce word was created by leaving 10 periods of the initial vowel, while the second one was created by leaving just 2.5 periods of the initial vowel. 10 periods of the vowel allows for the full vowel quality to be perceived, whereas 2.5 periods of phonation renders the vowel quality ambiguous. Examples of short /e/ and long /e/ can be seen in Figure 5.1.

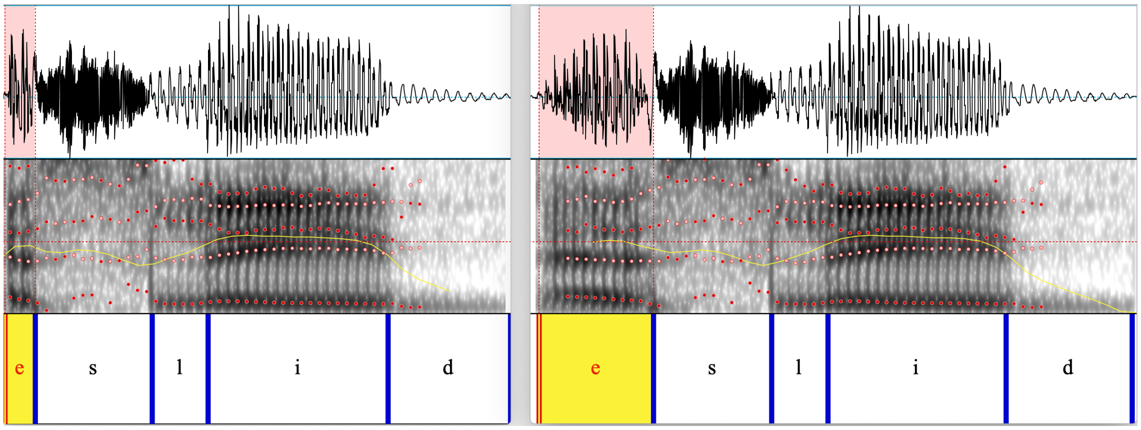


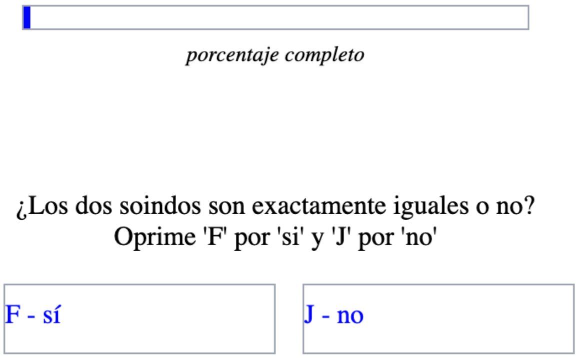
Figure 5.1: Comparison of /eslid/ with a short vowel of 2.5 periods (left) and a long vowel of 10 periods (right).

For each stimulus type (i.e., long /es/, short /es/, long /as/, short /as/) the exact same first syllable was spliced into the word so there were no phonetic differences between stimuli types; likewise, for each nonce word tail (i.e., /pid/, /fid/, /mid/, /lid/), the same token was combined with the initial syllable, again to avoid any fine-grained phonetic differences between stimuli types. The stimuli were also normalized to an intensity of 50dB. All modifications to the recorded stimuli were performed in PRAAT Phonetics Software (Boersma & Weenink, 2019). Stimuli with no initial vowel were not included because in Carlson et al.'s (2016) study, the bilinguals showed little evidence of a perceptual repair when there was no acoustic evidence for a vowel.

From these edited stimuli, 36 AX pairs were made with an interstimulis interval (ISI) of 250ms, following Davidson (2011), Carlson et al. (2016) and Carlson (2018, 2019), as the short ISI is intended to draw the participants' attention to the fine acoustic differences. In all AX pairs, the tail of the nonce word was the same. Thus, there were four different identical trials (identical trials (e.g., /espid - espid/), four trials in which vowel length remained constant across the stimuli but vowel quality varied (e.g., /espid - aspid/), four trials where only vowel quality was constant but vowel length differed (e.g., /espid - e:spid/), and four trials where both vowel quality and vowel length were distinct (e.g., /espid - a:spid/). Going forward, the differences in vowel quality and duration are denoted with the corresponding vowel (e or a) followed by 0 or 1, with 0 indicating that the vowel was short and 1 indicating that the vowel was long; for example, 'e1-a0' refers to a stimuli pair where the first part had a long /e/ and the second part had a short /a/ like in /e:spid - aspid/. This yielded a total of 16 different trial types per tail, for a total of 64 unique AX pairs: 16 identical and 48 non-identical (Appendix C). The non-identical pairs were presented twice (n=96), and the identical pairs were presented six times (n=96) so that there were an equal number of identical and non-identical trials.

5.3.2 Task

For each trial of the AX task, the participants heard one of the AX pairs, previously described. Each trial began with a fixation cross of 500ms, after which the AX stimuli pair was played. They were asked to respond with a key press if the two nonce words in the stimuli pair were identical or different. After responding the next trial began with the fixation cross. The key presses were counterbalanced across participants so that half responded with “F” for *same* and “J” for *different* responses, and the other half responded with “J” for *same* and “F” for *different* responses. If participants did not respond in five seconds or less, the task automatically advanced to the next AX pair; Figure 5.2 shows a screenshot of what participants would have seen for this portion of the experiment. The first ten trials were for training purposes, at which point participants were notified on the screen that the practice period was over and that the experiment would begin.



¿Los dos soindos son exactamente iguales o no?
Oprime 'F' por 'si' y 'J' por 'no'

Figure 5.2: Screenshot of what participants saw during the AX task

5.3.3 Statistics

The data from the AX task were analyzed using Signal Detection Theory (Green, Swets, et al., 1966; Pastore & Scheirer, 1974; Hautus, Macmillan, & Creelman, 2021). Signal Detection Theory is a way to measure participants’ ability to discriminate between stimuli while taking into account response bias. Response bias is an important consideration in tasks like an AX task because a participant could show 100% accuracy in saying all different pairs are different, but this could be the case because they responded ‘different’ to all trials, including identical stimuli pairs. Although the accuracy to the different pairs is at ceiling, these results do not give any information about how well the participant discriminated between the different pairs. On the other hand, a conservative participant could respond ‘same’ on all same and most of the different pairs, but this would show better discrimination than the previous participant who always answers ‘different’. Thus, percent correct of different pairs is not a meaningful measure of discrimination without accounting for response bias. Signal Detection Theory does exactly that.

Table 5.1 shows how an analysis using Signal Detection Theory organizes participants' responses. When the stimuli pair is different (e.g., e1-a0, e1-e0, etc.) and participants respond that they are different, this is a hit, and the raw number of hits is summed. When participants respond that the different trials are the same, these are false alarms, which are also summed. The hit rate (H) is then calculated by finding the probability of hits from all different trials (P(‘different’-response|different-stimuli)), and the false alarm rate (F) is the proportion of identical trials to which participants responded ‘different’ (P(‘different’-response|same-stimuli)). So, perfect discrimination would mean a hit rate of 1 and a false alarm rate of 0: the greater the difference between H and F, the better the participant's discrimination is.

	Stimuli: Different	Stimuli: Same
Response: ‘Different’	HIT	FALSE ALARM
Response: ‘Same’	MISS	CORRECT REJECTION

Table 5.1: Signal detection theory paradigm

The statistic d-prime (d‘) measures this difference by taking the difference of the z-transform of H and F (d‘=z(H)-z(F)). For the present study, A-prime (A’) scores were calculated rather than d‘ scores, which similarly measure the sensitivity in discriminating between same-different pairs, for multiple reasons. First, A’ is a score between 0 and 1, which can be more easily compared with the intuitive concept of ‘proportion correct’, whereas d‘ scores range from 0 to 4. Second, A’ calculations do not require corrections if responses are either all correct or all incorrect, in contrast from d‘ that does require this type of correction.

Third, A' allows false-alarm rates that are higher than hit rates, which d' does not (Hautus et al., 2021).

Linear mixed-effect regression models were then run on A' scores for each type of stimuli pair in order to look at language internal and external variables using the R packacke 'lme4' (Bates, Mächler, Bolker, & Walker, 2015). For all of these regressions, the type of sC-cluster was set as fixed effects and there was a by-subject random intercept included using the default polynomial trends contrast. When looking at the language internal variable, the interaction term was the stimuli pair, and when looking at the external variables of language profile, age of exposure to bilingualism, and bilingual dominance, the interaction terms were age of exposure group, language profile, and language dominance, respectively. The models were as follows:

> Q1 - Language internal effect: lmer(aprime ∼ stimulipair * sC-cluster + (1|subject), data=data)
>
> Q2 - Language profile: lmer(aprime ∼ sC-cluster * languageprofile + (1|subject), data=data)
>
> Q3 - Age of exposure: lmer(aprime ∼ sC-cluster * agegroup + (1|subject), data=data)
>
> Q4 - Language dominance: lmer(aprime ∼ sC-cluster * dominance + (1|subject), data=data)

5.4 Experiment 2: Nonce word judgment task

The second experiment is a nonce word judgment task. The goal of this experiment was to see if native Spanish speakers, both monolingual and bilingual, accept words with initial sC-clusters as possible words of Spanish, the results of which will provide insight on the language-internal and -external factors that affect the strength of an illusory vowel effect in this population.

Although Carlson (2019) had his participants perform a lexical decision task, a nonce word judgment task was performed instead because it is better at assessing phonotactic judgments rather than lexical judgments. The problem with a lexical decision task in a study of bilingualism is that the lexicons of the participants may vary based not only on their proficiency in Spanish but also on their dialect and context of acquisition. This could lead to an advantage for some speakers over others, which is completely independent of their phonotactic knowledge. In a nonce word judgment task, since none of the stimuli are real words, there should be little to no effect of lexical knowledge or robustness on the results.

Furthermore, a nonce word judgment task is ideal for investigating possibly non-categorical phonotactic patterns, as speakers' intuitions are rarely black and white but rather gradient (Anttila, 2008; Coetzee, 2008; Albright, 2009; White & Chiu, 2017). This gradience is seen in the form of intermediate judgments for nonce forms that are unattested yet grammatical; nonce words that are ungrammatical tend to receive significantly lower ratings, while nonce words that are grammatical and attested tend to receive significantly higher ratings. For example, a native speaker of English may give a nonce word 'blick' a high rating because the onset cluster is grammatical and attested, while giving a nonce word 'lbick' a low rating because the onset cluster is ungrammatical and unattested. A nonce word like 'shlick' may receive an intermediate rating because the onset cluster is unattested but abides by both the SSP and English's MSD.

The results of the present nonce word judgment task are then analyzed with respect to the language-internal factor of sC-cluster type and the language-external factors of language profile (i.e., monolingual versus bilingual), age of exposure to bilingualism, and language dominance.

5.4.1 Stimuli

The same male native speaker of Puerto Rican Spanish was recorded using a Zoom H5 recorder and Shure head-mounted SM10A microphone reading a list of nonce words. This speaker is also a native speaker of English, and so he had no difficulty accurately producing the sC-clusters word-initially. Each word was read aloud three times, but only the second repetition was used for the experiment in order to control for intonation across stimuli, and all stimuli were normalized for intensity in PRAAT after recording.

The target nonce words were created by combining complex onsets with tails of real Spanish words. The tails were selected from the Spanish SUBTLEX corpus (Cuetos et al., 2012) from bisyllabic or trisyllabic trochees (e.g., $(\sigma)\acute{\sigma}\sigma$), which is the regular Spanish stress pattern. The tails were created by removing the onset or onset cluster of the word in PRAAT, so that the first syllable of the tail began with the first vowel of the word (e.g., *paso* → *-aso*). All tails were selected from words with a log frequency between 0.75-2.5. The bisyllabic tails were randomly assigned to target heads, and other bisyllabic and trisyllabic tails were assigned to filler heads.

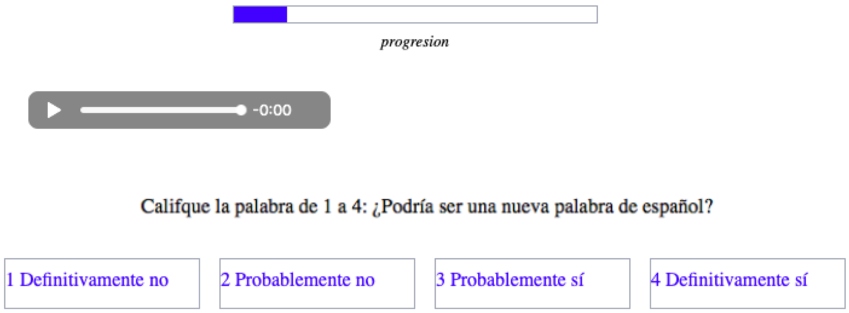
The target heads, or onsets, were as follows: /sp, st, sk, sm, sl/ and /esp, est, esk, esm, esl/. These were randomly assigned to the bisyllabic tails (Appendix D). Fillers were created with heads including licit Spanish onset clusters (/pr, pl, tr, kr, kl, fr, fl/) additional syllables (/asp, ins, tal, en, con, rem, pes, an, cal/), singleton onsets (/p, t, k, s, m, n, x, l, r/), and illicit onset clusters (/pn, tn, kn, bn, tl, fn/). The ungrammatical onset clusters were manually edited in PRAAT to remove intrusive vowels (i.e., /pnaso/ was produced as [pənaso], and the intrusive [ə] was removed for a stimulus of the form [pnaso]). Each tail that was associated with a target sC-cluster was also paired with an esC-initial word, a grammatical cluster, and an ungrammatical cluster (e.g., /spujo, espujo, prujo, fnujo/).

There were a total of 28 sC-clusters that violated the SSP and/or MSD; /sp, st, sk, sm/ were repeated seven times each with different tail. /sl/, the sC-cluster that is structurally distinct in English, was repeated 20 times with different tails. The /e/-initial counterparts to each of these onsets was used the same number of times with the same tails. Grammatical onset clusters, singleton onsets, ungrammatical onset clusters, and additional initial syllables were repeated four times each with different tails for a total of 120 fillers.

5.4.2 Task

In this task, the set of stimuli was presented aurally in a randomized order. Participants heard only one word at a time and were asked to rate the form on a scale from one to four on how likely this word could be a new word of Spanish. A rating of 1 meant that they believed it could never be a new word of Spanish; 2, that it probably could not be a new word of Spanish; 3, that it probably could be a new word of Spanish; and 4, that it definitely could be a new word of Spanish. Image 5.4 shows a screenshot of what participants would have seen for each of the 232 nonce words that was presented to them. The first five tokens presented were practice trials, so these responses were not included in the analyses. This portion of the testing lasted roughly 30 minutes.

Figure 5.3: Sample screen of what participants saw during the nonce word judgment task



Califque la palabra de 1 a 4: ¿Podría ser una nueva palabra de español?

Figure 5.4: Screenshot of what participants saw during the nonce word judgment task

5.4.3 Statistics

The dependent variable, the one-to-four rating from the nonce word judgment task, was analyzed using a cumulative link mixed-effects model (i.e., an ordinal regression) in R (R Core Team, 2013) ; the package utilized was ‘ordinal’ (Christensen, 2018). An ordinal regression was chosen due to the fact that participant responses were on an arbitrary scale from one to four, where only the relative ordering between the different values on this scale were relevant, as compared to a non-arbitrary continuous scale with linear regressions. For these regressions, the type of sC-cluster was set as a fixed effect, and language profile, age of exposure, and language dominance were used as interaction terms in subsequent models. For all regressions, there was a by-subject random intercept. The four models were as follows:

> Q1 - Language internal effect: clmm(response ∼ sC-cluster + (1|subject), data=data)
>
> Q2 - Language profile: clmm(response ∼ sC-cluster * languageprofile + (1|subject), data=data)
>
> Q3 - Age of exposure: clmm(response ∼ sC-cluster * agegroup + (1|subject), data=data)
>
> Q4 - Language dominance: clmm(response ∼ sC-cluster * dominance + (1|subect), data=data)

P-values from both tasks were corrected for multiple comparisons using Bonferroni correction.

5.5 Hypotheses

The goal of these experiments is to build off of past studies that have shown that English-Spanish bilinguals show evidence of an illusory vowel effect in Spanish. In the present study, there are four research questions that address language internal and external factors that could affect the illusory vowel effect in bilinguals in the context of sC-clusters. Here the research questions (Qs) are repeated, followed by the respective hypothesis (Hs).

> Q1: Does the type of sC-cluster (s+stop vs. s+liquid) affect the perception of these clusters in Spanish by native Spanish listeners?
>
>> H1: Based on the predictions of the HG-GLA that s+stop and s+liquid clusters do not have equivalent outputs, it is hypothesized that participants will treat s+liquid and s+stop clusters differently across both tasks. More specifically, it is hypothesized that s+liquid clusters will show more variability than s+stop clusters because the HG-GLA predicted that s+liquid surface forms are more probable than s+stop surface forms and because the HG-GLA also predicted that all s+liquid output forms have a taughtosyllabic structure, lessening the phonotactic ‘need’ for /e/ prothesis.

Q2: Does being bilingual affect the degree of variability in the the perception of s+stop and s+liquid clusters in Spanish?

> H2: Because surface sC-clusters are licit in English, it is hypothesized that English-Spanish bilinguals will show increased variability in the perception of sC-clusters compared to monolinguals. If there is a language-internal effect of sC-cluster type, it is further hypothesized that bilinguals will show even more variability with s+liquid clusters since the syllable structure is the same across the two languages.

Q3: Does the age of exposure to bilingualism affect the degree of variability in the perception of sC-clusters in bilinguals?

> H3: It is predicted that an earlier age of exposure to bilingualism, or in this case L2 English, will result in more variability in perception of sC-clusters.

Q4: Does language dominance, as measured by self-ratings of proficiency, affect degree of variability in the perception sC-clusters in bilinguals?

> H4: It is predicted that bilinguals who are dominant in English will show more variability in the perception of sC-clusters than those who are balanced or Spanish-dominant.

CHAPTER 6

RESULTS

6.1 Question 1: Language-internal variable

The first research question targets if the language-internal variables affects the perception of sC-clusters by native Spanish speakers. For this phenomenon, the language-internal variable refers to type of sC-cluster. Although past literature has assumed that all types of sC-clusters in Spanish are syllabified across the syllable boundary, the results of the HG-GLA, presented in Chapter 2, predicted that s+stop and s+liquid cluster outputs have different distributions of each surface form and hidden syllable structures. s+stop clusters were predicted to show /e/ prothesis more than s+liquid clusters, and s+liquid clusters were predicted to always have a taughtosyllabic syllable structure unlike s+stop clusters. Therefore, a difference in the perception of one type of sC-cluster versus the other would support this dual model of sC-cluster structure for Spanish.

6.1.1 AX task

The results of all 49 native-Spanish speakers are reported in Table 6.1 and can be visualized in Figure 6.1. The table displays the mean A' score and standard deviation for each stimuli pair. A higher A' score indicates a greater sensitivity to the difference in the two adjacent stimuli. In other words, a high A' means that the participant showed accuracy in identifying these pairs of non-identical stimuli as different and showed accuracy in identifying identical pairs as identical; 0.5 represents responding at chance and 1 represents perfect discrimination.

In the graph, the different critical stimuli pairs lie along the x-axis, with each vowel representing the vowel quality (/a/ vs. /e/) and duration (0-short vs. 1-long) in the stimuli pair. For example, ‘e1-a0’ indicates a stimulis pair like *e:spid-aspid*, where the first vowel is a long /e/ and the second is a short /a/. The y-axis measures the mean A’ score for each stimuli pair, and the data are separated by sC-cluster. The error bars represent the 95% confidence interval.

Cluster type	Stimulus pair	Mean A’ (S.D.)
s+stop	a1-e1	0.81 (0.20)
	a0-e0	0.70 (0.20)
	a1-a0	0.61 (0.16)
	a1-e0	0.81 (0.20)
	e1-e0	0.58 (0.14)
	e1-a0	0.76 (0.17)
s+liquid	a1-e1	0.81 (0.18)
	a0-e0	0.63 (0.19)
	a1-a0	0.57 (0.17)
	a1-e0	0.79 (0.20)
	e1-e0	0.54 (0.16)
	e1-a0	0.70 (0.19)

Table 6.1: Mean A’ and standard deviation for each stimuli pair by sC-cluster type for all participants.

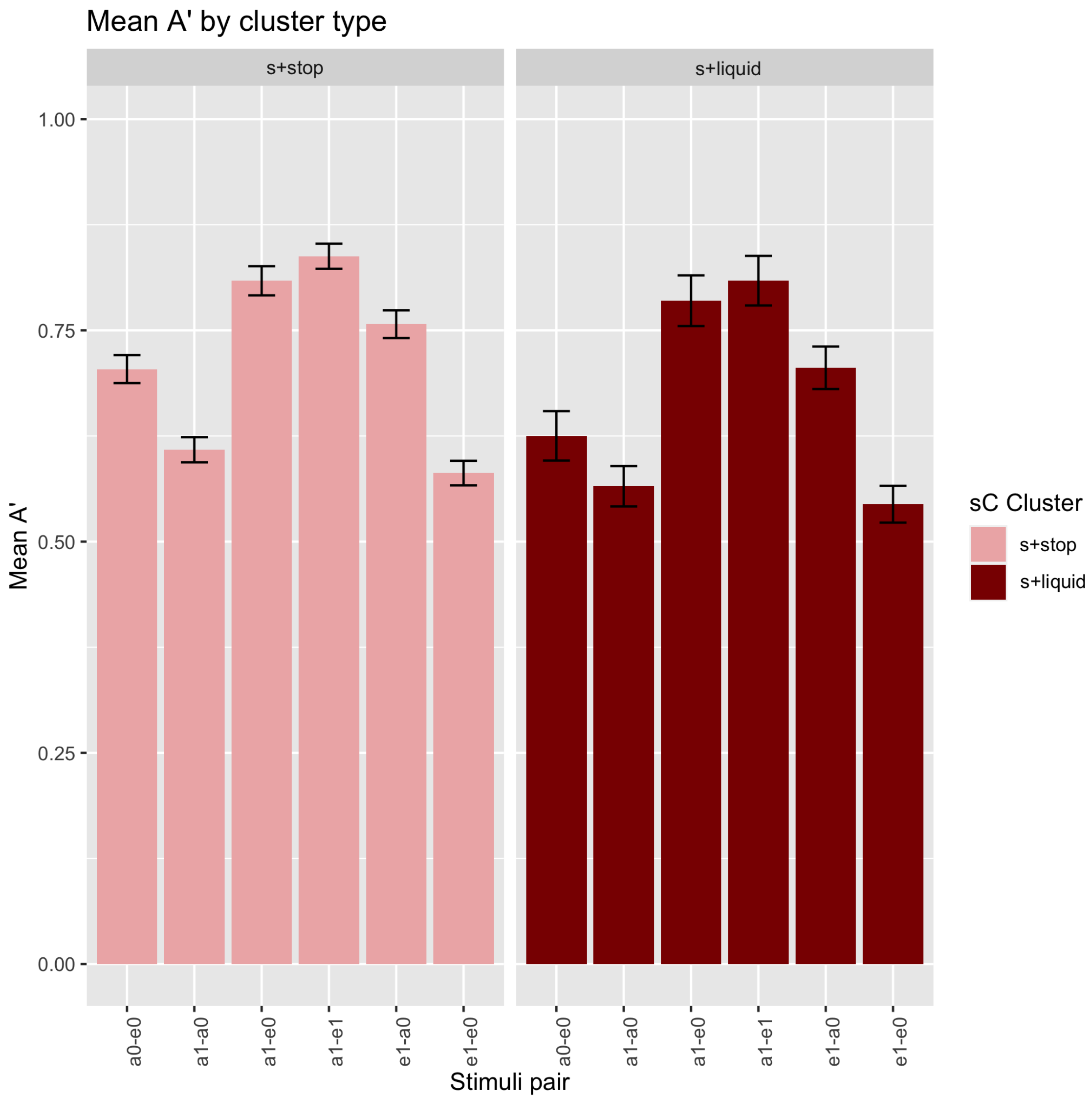


Figure 6.1: This figure shows the mean A' score of all participants for each of the critical stimuli pairs, separated by the initial sC-cluster. Error bars represent the 95% confidence interval.

Overall, s+liquid clusters yielded lower A' scores than s+stop clusters across stimuli pairs, and the discrimination trends were the same for each stimuli pair for both types of sC-cluster. When the vowels in the stimuli pair were both long and of different qualities (a1-e1), participants were able to correctly identify the stimuli as different for both types of sC-cluster (i.e., higher A' scores). It appears that the longer duration of the vowel provides enough acoustic evidence to be able to map the sound signal onto the /a/ or /e/, respectively.

When both vowels in the stimuli pair were short (a0-e0), the vowels did not have enough acoustic evidence to be mapped to a lexical vowel, yet participants performed above chance in discriminating these sounds; this was particularly the case for the s+stop pairs.

Following Carlson et al. (2016) and Carlson (2018, 2019), only the stimuli pairs that have vowels that differ in duration are considered the critical pairs (i.e., not a1-e1 or a0-e0 pairs); in the graph, the critical stimuli pairs begin with the third pair on the x-axis. Within this subset of stimuli pairs, participants were better at discriminating the stimuli pairs a1-e0 and e1-a0, where both the vowel quality and duration differed, than the pairs where only duration differed (a1-a0, e1-e0).

The critical stimuli pairs (a1-a0, e1-e0, a1-e0, e1-a0) were analyzed using a linear mixed effects regression model in order to determine if native Spanish speakers' ability to discriminate stimuli pairs differs depending on the type of sC-cluster (i.e., s+stop vs. s+liquid).[1] The results, found in Table 6.2, reveal a significant main effect of cluster type, with s+stop clusters more likely to have higher A' scores than s+liquid clusters (p=0.002). Furthermore, there is a main effect of stimuli pair, in which the pairs with different vowel qualities and durations (a1-e0, e1-a0) are more likely to have higher A' scores than the pairs that only differ in vowel duration (a1-a0, e1-e0) (p<0.0001).

Contrast: a1-a0 & s+stop	β estimate	S.E.	t-value	p-value
Intercept (a1-a0)	0.62	0.02	31.46	<0.0001*
a1-e0	0.19	0.02	12.51	<0.0001*
e1-a0	0.14	0.02	8.84	<0.0001*
e1-e0	-0.03	0.02	-1.62	0.11
s+liquid	-0.04	0.01	-3.06	0.0023

Table 6.2: Results of linear regression showing significant main effects of stimuli pair and sC-cluster. The asterisk denotes effects that remained significant after correction for multiple comparisons.

---

[1] The nonce words included in the 's+stop' category included /sp, sf, sm/ clusters all grouped together. However, the same linear regression run using sC-cluster (s+stop vs. s+liquid) as a predictor was run with consonant (sp, sf, sm) as a predictor and did not identify a significant effect of consonant within the s+stop group (R2=0.01, f(2,456)=2.55, p=0.10)

6.1.2 Nonce word judgment task

The data of 41 of the 47 participants were analyzed for the nonce word judgment task. Six participants were excluded due to a response bias in task performance. The remaining participants included eight Spanish monolinguals, 12 early child bilinguals, seven late child bilinguals, and 14 adult bilinguals.

The results of the nonce word judgment task are shown in Table 6.3, which includes the average rating and standard deviation for each word type, along with the percentage of each response for each word type. The mean response rating can also be seen in Figure 6.3.

Overall, participants rated sC-cluster nonce words as less acceptable than nonce words beginning with esC. The majority of esC-initial items were not rated as definitely or probably unacceptable, whereas the majority of sC-cluster items were rated as definitely or probably unacceptable. Within the sC-clusters, items beginning with s+liquid had a lower mean acceptability rating than s+stop clusters, due to the fact that more responses rated s+liquid clusters as definitely or probably unacceptable compared to s+stop clusters.

sC-cluster	Mean rating (S.D.)	1 Definitely no	2 Probably no	3 Probably yes	4 Definitely yes
s+stop	1.91 (0.99)	44.6%	29.0%	17.2%	9.3%
s+liquid	1.83 (0.95)	48.1%	28.5%	15.9%	7.4%
esC	2.63 (0.97)	14.5%	28.1%	37.0%	20.4%

Table 6.3: Mean rating, standard deviation, and proportion of each response type for all sC-cluster types across all participants.

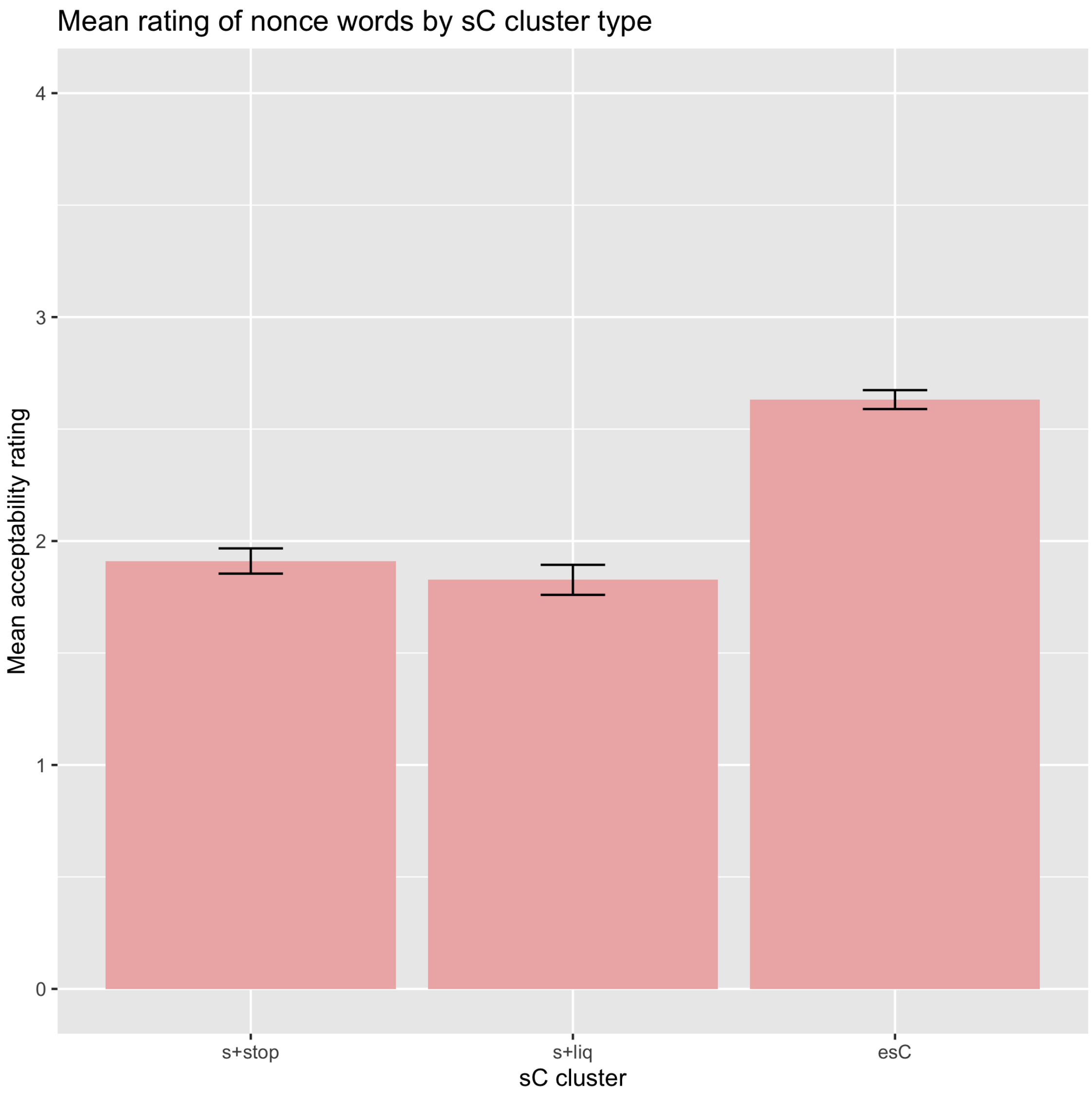


Figure 6.2: This figure shows the mean rating of each sC-cluster type by all participants. The rror bars represent the 95% confidence interval.

The results of the ordinal mixed-effects regression looking at the effect of cluster type on acceptability ratings are shown in 6.4; s+stop was set as the reference level. This model confirmed that esC words were rated as significantly more acceptable than sC-cluster words ($p<0.0001$), but the difference in acceptability between s+stop and s+liquid clusters was not statistically significant ($p=0.18$).

Contrast against s+stop	β-estimate	S.E.	z-value	p-value
s+liquid	-0.19	0.14	-1.33	0.18
esC	1.61	0.11	14.05	<0.0001*

Table 6.4: Results of the ordinal regression showing a significant main effect of nonce word type. The asterisk denotes effects that remained significant after correction for multiple comparisons.

6.1.3 Question 1 summary

These results show that, overall, native Spanish speakers have strong intuitions about what is licit versus illicit in Spanish. In the AX task, s+liquid stimuli pairs yielded more variability than s+stop pairs. However, participants showed high sensitivity to the fine-grained details in all stimuli pairs. In the nonce word judgment task, esC-initial nonce words were rated as significantly more acceptable than sC-initial nonce words, but there was no difference in acceptability based on the sC-cluster type.

6.2 Question 2: Language profile

The second research question aimed at comparing performance on the AX and nonce word judgment tasks of Spanish monolinguals (n=11) versus English-Spanish bilinguals (n=36), all of whom were native Spanish speakers, to investigate if simply being bilingual affects variability in sC-cluster perception.

6.2.1 AX task

Table 6.5 and Figure 6.3 show the mean A' scores of each critical stimuli pair by cluster type for Spanish monolingual versus English-Spanish bilingual participants. Bilinguals, on average, were better at discriminating the stimuli pairs, particularly for the s+stop pairs. The discrimination trends across the two language profiles were very similar and mirrored those of the previous section where all participants were combined in that stimuli pairs where both the vowel quality and duration differ resulted in more accurate responses.

Cluster type	Language profile	Stimulus pair	Mean A' (S.D.)
s+stop	Monolingual	a1-a0	0.59 (0.18)
		a1-e0	0.77 (0.22)
		e1-e0	0.58 (0.14)
		e1-a0	0.72 (0.19)
	Bilingual	a1-a0	0.61 (0.17)
		a1-e0	0.82 (0.20)
		e1-e0	0.58 (0.17)
		e1-a0	0.77 (0.19)
s+liquid	Monolingual	a1-a0	0.62 (0.21)
		a1-e0	0.73 (0.28)
		e1-e0	0.53 (0.13)
		e1-a0	0.66 (0.19)
	Bilingual	a1-a0	0.55 (0.14)
		a1-e0	0.80 (0.17)
		e1-e0	0.55 (0.15)
		e1-a0	0.72 (0.17)

Table 6.5: Mean A' score for each stimuli pair by language profile and sC-cluster type.

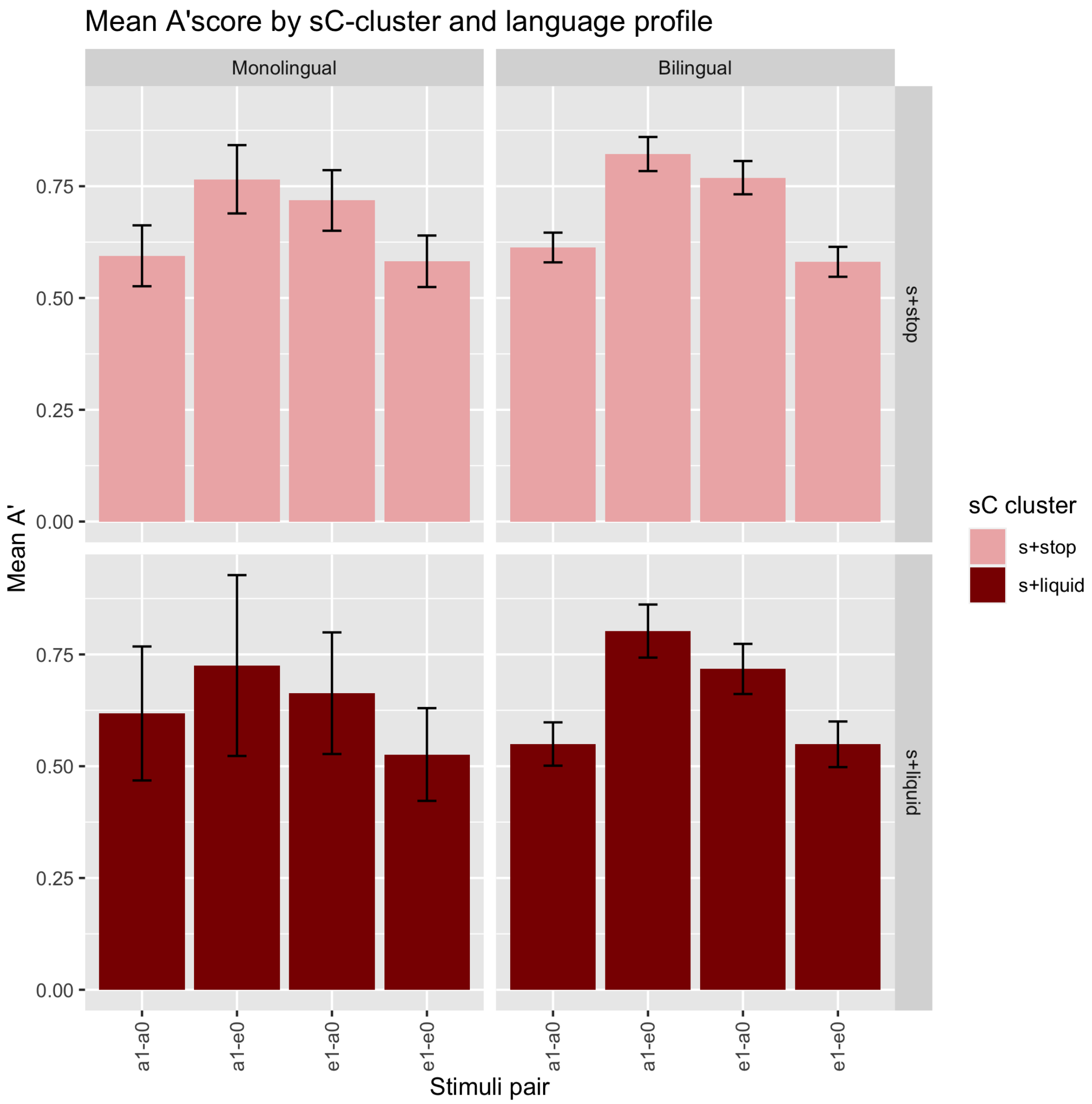


Figure 6.3: This figure shows the mean A' score of the critical stimuli pairs by language profile (monolingual vs. bilingual) and sC-cluster type. The error bars represent the 95% confidence interval.

The mean A' scores for these four stimuli pairs were analyzed using a linear mixed effects regression to determine if task performance differed between the monolingual and bilingual participants. These results, found in Table 6.6, show that there is no main effect of cluster type or language profile. In other words, s+stop and s+liquid clusters yielded similar A' scores, and bilinguals and monolinguals did not perform significantly differently.

Contrast: Monolingual * s+stop	β estimate	S.E.	t-value	p-value
s+liquid	-0.03	0.03	-0.98	0.33
bilingual	0.02	0.04	0.45	0.66
s+liquid * bilingual	-0.01	0.04	-0.34	0.74

Table 6.6: Linear regression with type of sC-cluster and language profile as interaction terms. P-values marked with an asterisk denote effects that remained significant after correction for multiple comparisons.

6.2.2 Nonce word judgment task

In the nonce word judgment task, bilinguals showed higher acceptability of all sC-cluster words, shown in Table 6.7 and Figure 6.4. All participants rated esC-initial words as the most acceptable, but monolinguals rated s+stop and s+liquid words as highly unacceptable. Bilinguals, on the other hand, rated s+stop and s+liquid words as more acceptable than the monolinguals did but still less acceptable than the grammatical esC-initial words.

sC-cluster	Language profile	Mean rating (S.D.)	1	2	3	4
s+stop	Monolingual	1.33 (0.79)	82.3%	7.8%	4.7%	5.2%
	Bilingual	1.99 (0.97)	37.9%	34.2%	18.6%	9.2%
s+liquid	Monolingual	1.26 (0.71)	86.2%	5.9%	3.9%	3.9%
	Bilingual	1.97 (0.95)	38.9%	34.0%	18.8%	8.2%
esC	Monolingual	2.56 (0.90)	13.3%	32.7%	39.0%	15.1%
	Bilingual	2.65 (0.98)	14.8%	27.0%	36.5%	21.6%

Table 6.7: Mean rating, standard deviation, and proportion of each response type for all sC-cluster types by language profile.

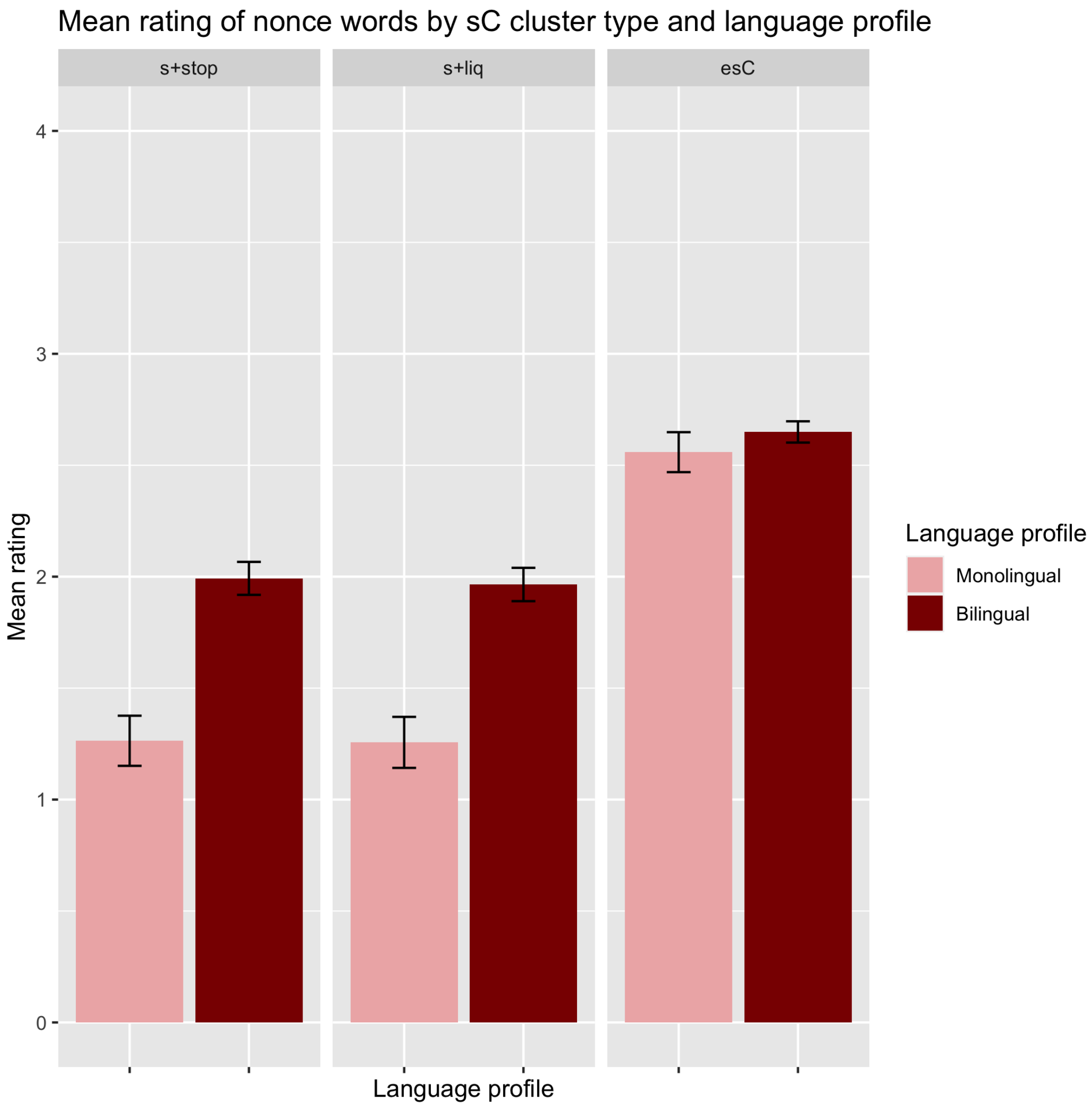


Figure 6.4: Mean acceptability rating by sC-cluster type by language profile. The error bars represent the 95% confidence interval

The results, shown in Table 6.8 of the ordinal mixed effects regression confirmed these observations. This model revealed a significant main effect of sC-cluster type, with grammatical esC-intial words yielding significantly higher acceptability rating scores than s+stop or s+liquid words across all participants ($p<0.0001$). There was also a significant main effect of language profile, with bilinguals' average responses being greater than monolinguals' ($p<0.0001$). A significant interaction between esC-words and bilingual participants indicates that the difference between grammatical esC-words and the ungrammatical sC-cluster nonce words was less than it was for monolinguals (i.e., monolinguals differentiated sC-cluster words from esC-intial words more than bilinguals did).

Contrast: Monolingual * s+stop	β estimate	S.E.	z-value	p-value
s+liquid	-0.006	0.33	-0.18	0.99
esC	3.63	0.25	14.44	<0.0001*
bilingual	2.38	0.41	5.75	<0.0001*
s+liquid * bilingual	-0.05	0.35	-0.14	0.89
esC * bilingual	-2.23	0.26	-8.44	<0.0001*

Table 6.8: Ordinal regression with type of sC-cluster and language profile as interaction terms. P-values marked with an asterisk denote an effect that remained significant after correction for multiple comparisons.

6.2.3 Question 2 summary

Comparing the Spanish monolingual versus the English-Spanish bilingual participants revealed inconsistent results across the two tasks. In the AX task, there is no effect of language profile: bilinguals and monolinguals perform the same in discriminating stimuli sC-cluster stimuli pairs. In the nonce word judgment task, on the other hand, the language profile of the participant does matter. Bilinguals were more accepting of the nonce words, particularly those that began with sC-clusters. However, there was no difference in acceptability ratings by either group between s+stop and s+liquid clusters.

6.3 Question 3: Age

The next research question probed at the language-external variable of age of exposure to bilingualism, or in this study exposure to English, and its effect on the perception of sC-clusters in Spanish. As mentioned in Chapter 5, participants were grouped based on their age of exposure to English in one of the following groups: monolingual Spanish, early child bilinguals, late child bilinguals, or adult bilinguals.

6.3.1 AX task

The results of the same AX task were further analyzed by comparing participants based on their age of exposure to bilingualism and by sC-cluster type. Table 6.9 and Figure 6.5 show the mean A' scores for the critical stimuli pairs (a1-e0, a1-a0, e1-a0, e1-e0) by sC-cluster for each bilingual group.

The results of each participant group resemble the overall results presented in the previous section. All groups showed better discrimination for stimuli pairs that differed both in vowel quality and duration (a1-e0, e1-a0). However, the mean A' scores for these pairs was lower for monolingual Spanish speakers compared to bilinguals. A' scores between s+stop and s+liquid clusters was comparable within each bilingual group, and variation within each sC-cluster group, as measured by the standard deviations, was also comparable.

Cluster type	Age group	Stimulus pair	Mean A' (S.D.)
s+stop	Monolingual	a1-a0	0.59 (0.18)
		a1-e0	0.77 (0.22)
		e1-e0	0.58 (0.14)
		e1-a0	0.72 (0.19)
	Early child	a1-a0	0.57 (0.16)
		a1-e0	0.82 (0.22)
		e1-e0	0.65 (0.17)
		e1-a0	0.78 (0.22)
	Late child	a1-a0	0.60 (0.15)
		a1-e0	0.79 (0.22)
		e1-a0	0.75 (0.21)
		e1-e0	0.56 (0.16)
	Adult	a1-a0	0.66 (0.17)
		a1-e0	0.84 (0.17)
		e1-e0	0.53 (0.16)
		e1-a0	0.77 (0.16)
s+liquid	Monolingual	a1-a0	0.62 (0.21)
		a1-e0	0.73 (0.28)
		e1-e0	0.53 (0.13)
		e1-a0	0.66 (0.19)
	Early child	a1-a0	0.56 (0.17)
		a1-e0	0.81 (0.16)
		e1-e0	0.63 (0.16)
		e1-a0	0.75 (0.16)
	Late child	a1-a0	0.51 (0.08)
		a1-e0	0.76 (0.22)
		e1-e0	0.51 (0.08)
		e1-a0	0.65 (0.22)
	Adult	a1-a0	0.57 (0.13)
		a1-e0	0.82 (0.17)
		e1-e0	0.50 (0.14)
		e1-a0	0.73 (0.14)

Table 6.9: Mean A' score for each stimuli pair by bilingual group and sC-cluster type

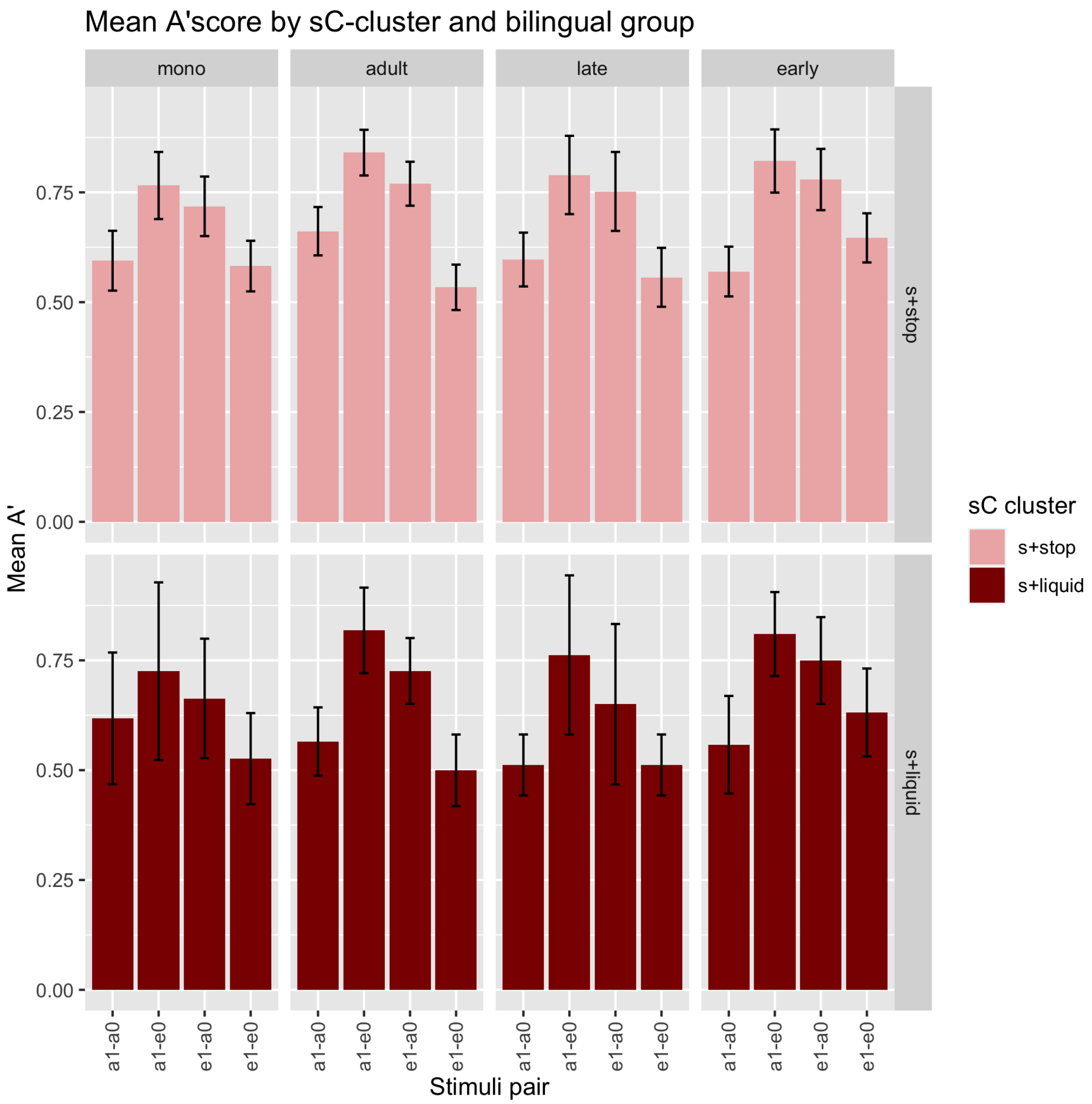


Figure 6.5: Mean A' score for critical stimuli pairs by sC-cluster type and bilingual group. Error bars represent 95% confidence intervals.

The results to the second linear regression are found in Table 6.10. Predictors included sC-cluster and bilingual group (age of exposure groups) as interaction terms; s+stop cluster and the monolingual group were set as the reference values.

This model revealed no significant interactions between sC-cluster type and age of exposure to bilingualism; overall, discrimination of stimuli pairs was consistent across all participant groups, regardless of their age of exposure to English and the type of sC-cluster.

Contrast: Monolingual * s+stop	β estimate	S.E.	t-value	p-value
s+liquid	-0.03	0.03	-0.98	0.33
early child	0.03	0.05	0.53	0.59
late child	-0.01	0.06	-0.23	0.82
adult	0.03	0.06	0.59	0.56
s+liquid * early child	0.01	0.04	0.33	0.75
s+liquid * late child	-0.02	0.04	-0.56	0.58
s+liquid * adult	-0.03	0.05	-0.71	0.48

Table 6.10: Linear regression with type of sC-cluster and bilingual group as interaction terms, monolingual and s+stop cluster as referecne values. P-values marked with an asterisk indicate a significant effect after correction for multiple comparisons.

6.3.2 Nonce word judgment task

Table 6.11 provides the mean acceptability rating and standard deviation of each cluster type by the bilingual group, along with the proportion of responses in each response category. Figure 6.6 also shows the mean ratings of the nonce word types by bilingual group.

For all four participant groups, the grammatical esC-initial nonce words were rated as the most acceptable, and s+stop and s+liquid words were rated as less acceptable. However, all bilingual groups rated the sC-cluster words as more acceptable than the monolinguals did. Furthermore, all participants except the late child bilinguals rated s+stop clusters as more acceptable than s+liquid clusters.

sC-cluster	Age group	Mean rating (S.D.)	1	2	3	4
s+stop	Monolingual	1.33 (0.79)	82.3%	7.8%	4.7%	5.2%
	Early child bilingual	2.07 (0.97)	34.5%	33.3%	22.7%	9.5%
	Late child bilingual	2.23 (0.90)	19.7%	48.8%	20.2%	11.3%
	Adult bilingual	1.95 (1.02)	44.1%	27.6%	18.0%	10.3%
s+liquid	Monolingual	1.26 (0.71)	86.2%	5.9%	3.9%	3.9%
	Early child bilingual	1.93 (0.97)	42.1%	30.7%	18.9%	8.3%
	Late child bilingual	2.26 (0.94)	21.8%	42.9%	22.6%	12.8%
	Adult bilingual	1.84 (0.91)	44.7%	32.3%	16.9%	6.0%
esC	Monolingual	2.56 (0.90)	13.3%	32.7%	39.0%	15.1%
	Early child bilingual	2.60 (1.04)	18.7%	26.2%	31.3%	23.8%
	Late child bilingual	2.75 (0.91)	9.9%	27.1%	41.4%	21.6%
	Adult bilingual	2.64 (0.95)	14.0%	27.7%	38.5%	19.8%

Table 6.11: Mean rating, standard deviation, and proportion of each resopnse type for all sC-cluster types by bilingual group.

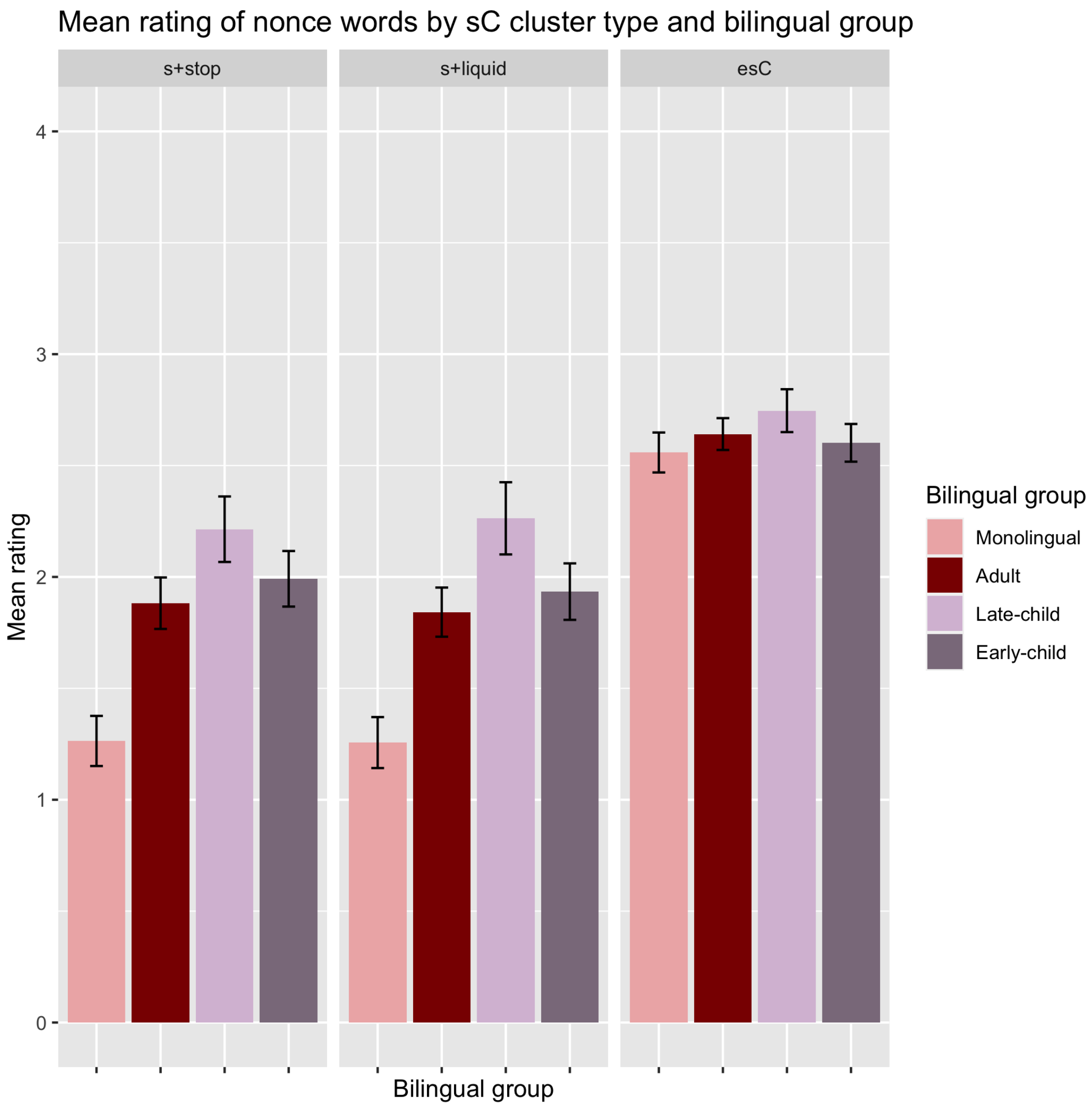


Figure 6.6: Mean acceptability rating for each type of nonce word by bilingual group. Error bars represent the 95% confidence interval.

Table 6.12 show the results of the ordinal regression where sC-cluster type and bilingual group are set as interaction terms; monolinguals and s+stop clusters were set as the reference values.

There were significant main effects of bilingual group, with all bilingual groups showing significantly higher mean acceptability ratings than monolinguals ($p<0.0001$ for all three bilingual groups), and there were no significant differences bilingual groups. There was no significant difference between acceptability of s+stop versus s+liquid clusters, but a main effect of sC-cluster group showed that esC-initial were significantly more acceptable than sC-cluster words across all participant groups ($p<0.0001$). Furthermore, significant interactions between esC-initial words and the three bilingual groups indicates that the difference in ratings between the esC- versus sC-initial words was less than for monolinguals (i.e., monolinguals differentiated esC-intial words from sC-cluster words more than the bilinguals did).

Contrast against s+stop and monolingual group	β-estimate	S.E.	z-value	p-value
s+liquid	-0.32	0.32	-1.0	0.32
esC	3.42	0.23	15.15	<0.0001*
adult	1.97	0.44	4.50	<0.0001*
early	2.32	0.45	5.16	<0.0001*
late	2.76	0.50	5.47	<0.0001*
s+liquid * adult	0.15	0.34	0.45	0.65
esC * adult	-1.85	0.24	-7.86	<0.0001*
s+liquid * early	-0.01	0.34	-0.04	0.97
esC * early	-2.25	0.24	-9.37	<0.0001*
s+liquid* late	0.34	0.36	0.93	0.35
esC * late	-2.41	0.26	-9.33	<0.0001*

Table 6.12: Ordinal regression with type of sC-cluster and bilingual group as interaction terms, monolingual group and s+stop clusters as reference values. P-values marked with an asterisk indicate a significant effect after correction for multiple comparisons.

6.3.3 Question 3 summary

In the AX task, age of exposure to bilingualism was not a predictor of performance on discriminating sC-cluster stimuli pairs, and there was no interaction with sC-cluster type. The nonce word judgment task also did not show an effect of age of exposure to bilingualism based on acceptability ratings of sC-cluster nonce words, nor were their interactions based on sC-cluster type. However, as in the previous section, bilinguals patterned together regardless of their age of exposure to English and showed less differentiation in acceptability between esC- and sC-initial words than monolinguals did.

6.4 Question 4: Language dominance

Language dominance was calculated for the bilingual participants based on a composite score of self-rated proficiency in Spanish comprehension and speaking and English comprehension and speaking, described in Section 5.2. Participants categorized as balanced, English-dominant, or Spanish-dominant. Spanish monolingual participants were not included in these analyses.

6.4.1 AX task

The data of the 36 bilingual participants were used to analyze the results of the AX task based on bilingual dominance. The mean A' scores and standard deviations are shown in Table 6.13, and these data can be visualized in Figure 6.7.

Balanced bilinguals were the worst at discriminating the stimuli pairs. English-dominant bilinguals were best at discriminating the vowel length distinction in the e1-e0 pair, but Spanish-dominant bilinguals showed better discrimination for all other pairs (a1-a0, a1-e0, e1-a0). Within each language dominance group, s+liquid clusters tended to have lower A' scores than the s+stop clusters.

sC-cluster	Language dominance	Stimulus pair	Mean A' (S.D.)
s+stop	balanced	a1-a0	0.57 (0.15)
		a1-e0	0.73 (0.23)
		e1-e0	0.55 (0.17)
		e1-a0	0.71 (0.21)
	English-dominant	a1-a0	0.57 (0.19)
		a1-e0	0.77 (0.22)
		e1-e0	0.67 (0.18)
		e1-a0	0.73 (0.23)
	Spanish-dominant	a1-a0	0.57 (0.15)
		a1-e0	0.73 (0.23)
		e1-e0	0.57 (0.16)
		e1-a0	0.78 (0.02)
s+liquid	balanced	a1-a0	0.52 (0.12)
		a1-e0	0.69 (0.22)
		e1-e0	0.48 (0.10)
		e1-a0	0.63 (0.17)
	English-dominant	a1-a0	0.60 (0.21)
		a1-e0	0.77 (0.12)
		e1-e0	0.66 (0.14)
		e1-a0	0.70 (0.19)
	Spanish-dominant	a1-a0	0.52 (0.16)
		a1-e0	0.69 (0.22)
		e1-e0	0.48 (0.10)
		e1-a0	0.78 (0.12)

Table 6.13: Mean A' and standard deviation for each stimuli pair by sC-cluster type and bilingual language dominance

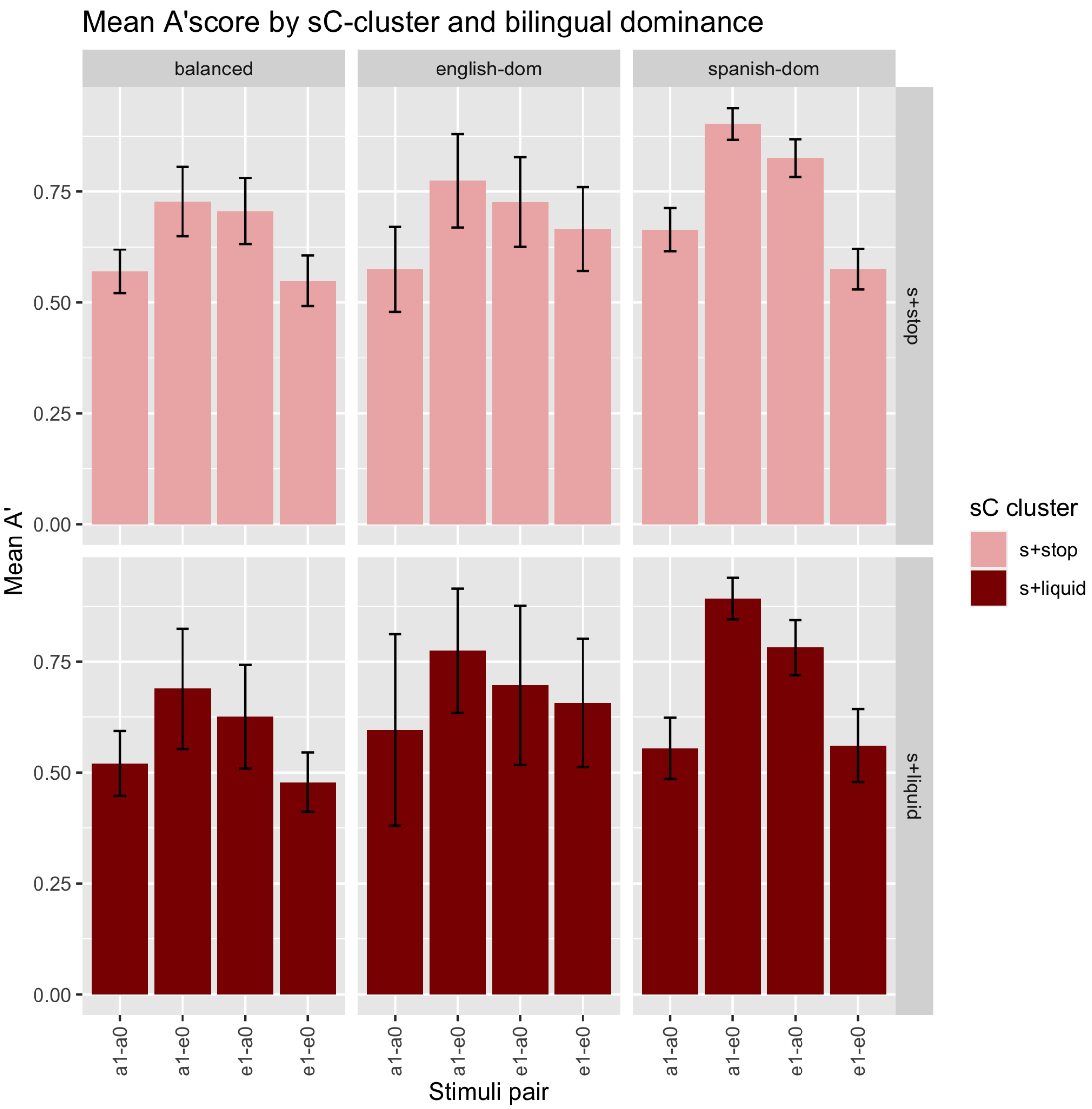


Figure 6.7: Mean A' for each stimuli pair by sC-cluster type and bilingual language dominance

Language dominance and type of sC-cluster were used as interaction terms in another linear mixed-effects regression, the results of which are shown in Table 6.14. Balanced bilinguals and s+stop clusters were set as the reference values.

There was a significant effect of bilingual dominance, in which Spanish-dominant participants were more likely to be able to discriminate the stimuli pairs than the balanced and English-dominant bilinguals (p=0.003). However, there were no significant interactions between sC-cluster type and dominance group, indicating that performance across stimuli pairs and cluster types followed the same trends across bilingual dominance groups.

Contrast: Balanced * a1-a0	β estimate	S.E.	t-value	p-value
s+liquid	-0.06	0.03	-1.92	0.06
English-dom	0.06	0.05	1.19	0.24
Spaish-dom	0.12	0.04	3.10	0.003
s+liquid * English-dom	0.04	0.05	0.86	0.39
s+liquid * Spanish-dom	0.01	0.04	0.32	0.75

Table 6.14: Linear regression with type of sC-cluster and language dominance as interaction terms, balanced and s+stop cluster as reference values. P-values marked with an asterisk indicate a significant effect after correction for multiple comparisons.

6.4.2 Nonce word judgment task

Acceptability of sC-clusters varied based on the language dominance of the bilingual participants, as seen in Table 6.15 and Figure 6.8. Balanced bilinguals rated the grammatical esC-initial words as slightly less acceptable than the English- and Spanish-dominant bilinguals. English-dominant bilinguals rated sC-clusters as more acceptable than the other two groups particularly for the s+liquid clusters. Spanish-dominant bilinguals rated both types of sC-clusters as less acceptable than did the English-dominant and balanced bilinguals.

sC-cluster	Dominance group	Mean rating (S.D.)	1	2	3	4
s+stop	Balanced	2.10 (1.01)	33.1%	36.2%	17.9%	12.8%)
	English-dom	2.12 (0.91)	28.7%	37.9%	25.9%	7.5%
	Spanish-dom	2.00 (0.99)	39.1%	31.6%	19.5%	9.7%
s+liquid	Balanced	2.00 (1.99)	38.4%	33.2%	18.4%	10.0%
	English-dom	2.12 (1.03)	35.1%	29.8%	22.8%	12.3%
	Spanish-dom	1.89 (0.90)	40.6%	35.9%	17.6%	5.9%
esC	Balanced	2.60 (0.98)	15.3%	30.0%	34.1%	20.6%
	English-dom	2.62 (0.99)	16.7%	25.5%	37.4%	20.4%
	Spanish-dom	2.69 (0.97)	13.9%	25.8%	37.6%	22.7%

Table 6.15: Mean, standard deviation, and proportion of each response for all sC-cluster types by language dominance.

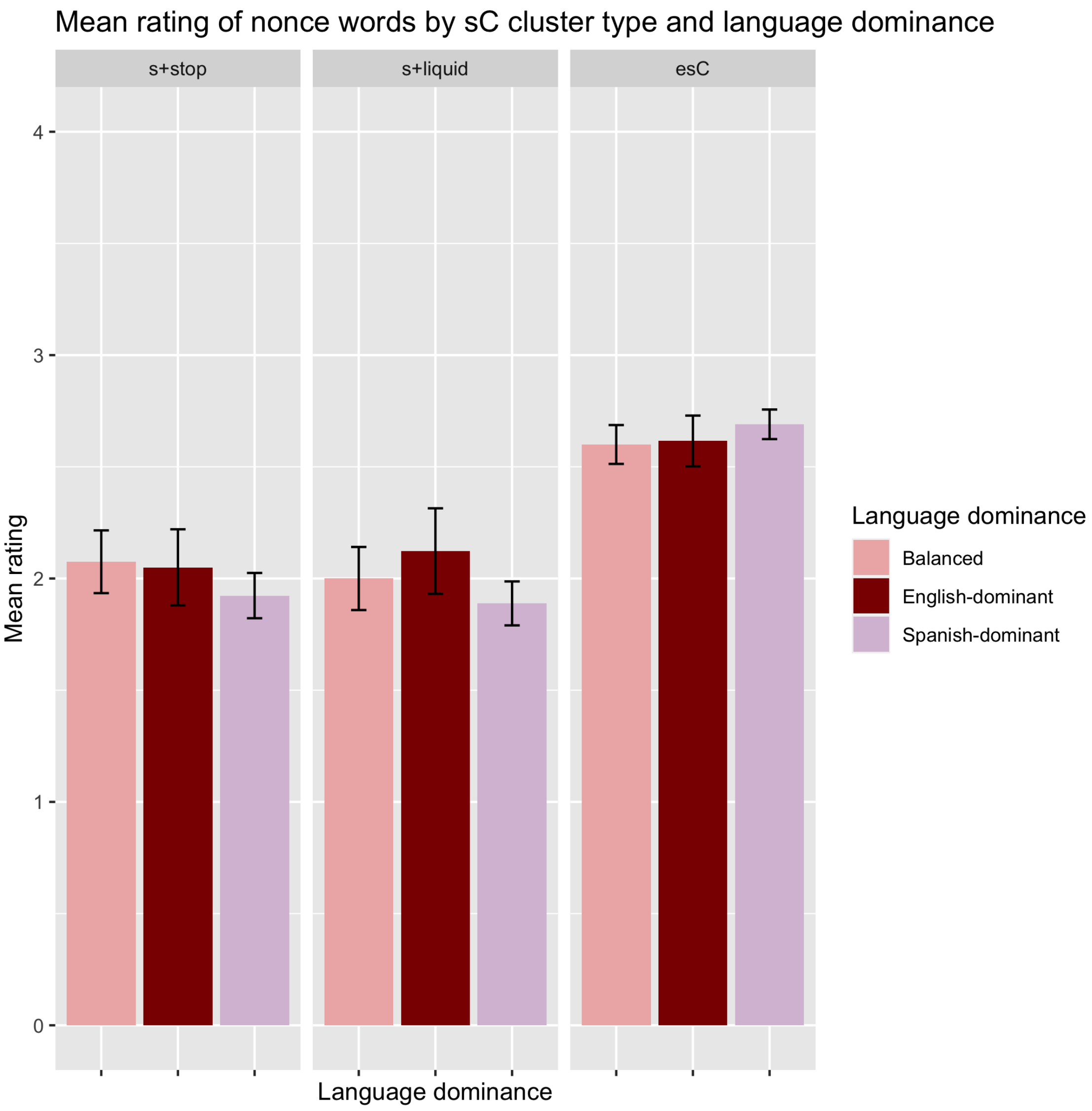


Figure 6.8: Mean acceptability rating for each type of nonce word by language dominance. Error bars represent 95% confidence interval.

Table 6.16 presents the results of the ordinal regression with the interaction terms of cluster type and language dominance group. Again, the model shows that esC-initial words were rated as significantly more acceptable by all bilingual participants. Additionally, there was a significant interaction between esC-initial words and the Spanish-dominant bilinguals (p=0.005), indicating that the difference in ratings between sC-clusters and esC-intial words was greater than for the other two bilingual groups.

Contrast against cluster-s+stop and dom-balanced	β-estimate	S.E.	z-value	p-value
s+liquid	-0.24	0.22	-1.21	0.26
esC	1.07	0.17	6.24	<0.0001*
English-dom	007	0.53	0.13	0.90
Spanish-dom	-0.28	0.41	-0.68	0.50
s+liquid * English-dom	0.21	0.29	0.71	0.48
esC * English-dom	-0.04	0.24	-0.17	0.87
s+liquid * Spanish-dom	0.04	0.22	0.17	0.86
esC * Spanish-dom	0.48	0.18	2.77	0.005

Table 6.16: Ordinal regression with type of sC-cluster and language dominance as interaction terms, balanced bilinguals and s+stop clusters as reference values. P-values marked with an asterisk indicate a significant effect after correction for multiple comparisons.

6.4.3 Question 4 summary

Bilingual dominance showed variable effects on the perception of sC-clusters in English-Spanish bilinguals in the two tasks. In the AX task, Spanish-dominant bilinguals showed better discrimination of stimuli pairs than English-dominant or balanced bilinguals. Their discrimination abilities in the AX task were better than the monolinguals whose results were presented in the results sections discussing Questions 2 and 3. The results of the nonce word judgment task showed that Spanish-dominant bilinguals showed more similar acceptability ratings to the monolinguals in Questions 2 and 3 than the English-dominant and balanced bilinguals did. This was evident from the lower acceptability ratings of sC-cluster words and higher ratings of esC-initial words, which was a significant interaction effect.

CHAPTER 7

DISCUSSION

The main goal of phonotactics is to understand speakers' knowledge of what sound patterns are possible and impossible in their language. One phenomenon that has been widely studied in the field of phonotactics is sC-clusters because they tend to pattern differently than other onset clusters both within and across languages.

In English, for example, onset clusters must have a rise in sonority with the exception of sC-clusters. This has led to the proposal that sC-clusters are distinct from other onset clusters in their syllable structure. While onset clusters that have a rise in sonority are syllabified as branching onsets, sC-clusters that do not abide by the sonority preferences in English are syllabified as a coda /s/ followed by a singleton onset of the following syllable. It has been further proposed that sC-clusters that do not violate sonority restrictions in English are syllabified as branching onsets (i.e., s+liquid), which is supported by the increased variability of s+liquid versus s+stop production in acquisition.

The phonotactics of Spanish differ from English in that sC-clusters are never allowed to surface; it has been previously assumed that they are repaired by inserting a prothetic /e/, which is then followed by the coda /s/ and the singleton onset of the next syllable (/sC/ → [es.C]) regardless of the sonority profile of the cluster. However, the present study utilized a phonological learning algorithm, the HG-GLA, to predict both the surface and hidden structures of sC-clusters in Spanish. The predictions of the learner differed from past analyses: it predicted that sC-clusters show variability in their syllabification. s+stop clusters were predicted to surface with a prothetic /e/ 67% of the time and have an underlying bisyllabic structure 33% of the time. s+liquid clusters, however, were predicted to always be taughtosyllabic and only instantiate the prothetic /e/ half of the time. This led to the the hypothesis that perception of s+stop clusters would be different than that of s+liquid clusters by native speakers of Spanish.

Furthermore, past studies have shown that Spanish and English contradict each other regarding the phonotactics of sC-clusters (i.e., English allows surface sC-clusters and Spanish does not) and that English-Spanish bilinguals show non-monolingual-like behavior in perception. This non-monolingual-like behavior is referred to as bilingual variability, and has been shown to occur when there is non-identical structural overlap between the two languages, as is the case with sC-clusters.

In order to investigate the variability associated with sC-clusters in Spanish, two experiments were conducted: an AX task and a nonce word judgment task. The goal of these tasks was to answer the following research questions. The first question addressed the variability associated with the language internal factor and sought to determine if there was a difference in perception of sC-clusters in Spanish based on the different predictions of s+stop versus s+liquid clusters by the HG-GLA in native speakers of Spanish. The second question addressed bilingual variability in sC-cluster perception in Spanish by comparing the results of the two tasks between Spanish monolingual and English-Spanish bilingual speakers. Then, the language external factor of age of exposure to bilingualism was tested to determine if an earlier age of exposure to bilingualism resulted in increased variability in sC-cluster perception in Spanish. The final research question addressed how language dominance within English-Spanish bilinguals affects the perception of sC-clusters in Spanish.

The results of these studies are interpreted below one research question at a time, followed by a discussion of how generative models of phonology could be adapted to accommodate the phonotactic systems of bilingual speakers.

7.1 Question 1: Language-internal factor

The first research question addressed the question of if the type of sC-clusters affected their perception by native Spanish speakers. As discussed in Chapter 2, previous analyses of Spanish sC-clusters have assumed that all sC-clusters are syllabified in the same way: the sC-cluster is ‘broken up’ by the insertion of a prothetic /e/, which results in the resyllabification of the /s/ as a coda and the second consonant as a singleton onset (e.g., sC → es.C). However, the results of the phonological learner used in Chapter 2, the HG-GLA, made a different prediction. The HG-GLA predicted that s+stop clusters are more likely to instantiate /e/ prothesis compared to s+liquid clusters, but that s+liquid clusters always have a taughtosyllabic structure, unlike s+stop clusters. Thus, it was hypothesized that s+liquid clusters would show more acceptability in perception than s+stop clusters. This hypothesis was further motivated by the assumption that /e/ prothesis occurs in order to resyllabify the sC-cluster, but if the predictions of the HG-GLA are correct and s+liquid clusters are never bisyllabic, this mitigates the need for s+liquid clusters to invoke /e/ prothesis.

The AX and nonce word judgment tasks showed inconsistent results. The results of the AX task showed a significant effect of sC-cluster type. Participants were, overall, very good at discriminating the minute phonetic differences in all stimuli pairs, performing above chance for every critical pair. However, they were worse at discriminating stimuli pairs with s+liquid versus s+stop clusters. In other words, they were more likely to say that different stimuli pairs were the same when the initial cluster was s+liquid rather than s+top. The standard deviations of the A' scores for s+liquid clusters were also greater than those for s+stop clusters, indicating that there was more variability within responses for this cluster type as well. This result could be interpreted as support for the prediction of the HG-GLA that s+stop and s+liquid clusters may have different syllable structures and different likelihoods of perceiving the prothetic /e/, since native Spanish speakers did not treat them the same in this AX task.

The results of the nonce word judgment task, however, paint a different picture. In this task, native Spanish speakers rated s+stop and s+liquid clusters as equally unacceptable. No language-internal variability was observed. Both s+stop and s+liquid clusters were rated as unacceptable, whereas esC-initial nonce words were rated as significantly more acceptable. This demonstrates that native Spanish speakers have strong intuitions about what is versus what is not grammatical in the target language. The lack of language-internal variability here does not provide evidence of a structural difference based on sC-cluster type, which in turn does not support the predictions made by the HG-GLA. If the data had supported the predictions of the phonological learner, acceptability ratings for s+liquid clusters would have been higher or more variable than for s+stop clusters since the learner assigned higher probability to surface [sl] outputs compared to [sp] outputs, but this was not the case.

However, these results do not necessarily refute the phonological learner's predictions either. It is possible that the HG-GLA would have predicted more similar output candidate probabilities for the two types of sC-cluster types if had it been provided with any /sl/ inputs. No /sl/ words were provided as input because there were no Spanish or loanwords beginning with /sl/ in the corpus used. Furthermore, differences between s+stop and s+liquid clusters in English have been observed mostly in acquisition data but not in adulthood, so it is possible that, because the native Spanish speakers in the present study had fully developed grammars, they were not the ideal test population for revealing this possible difference in sC-cluster types. Because Spanish has very few words beginning with /esl/, acquisition data may not be as variable as it is in English. However, T. Prince (2014) showed that French speakers with stroke-induced aphasia showed variable production of sC-clusters in French. Atypical speakers of Spanish, such as those with stroke-induced or neurodegenerative forms of aphasia that primarily affect phonological processes, may show variability in sC-cluster perception more similar to what was predicted by the HG-GLA. Nevertheless, additional experimental tasks need to be developed to further examine if there is a difference in syllable structure between different sC-clusters.

7.1.1 Task-based differences

It is important to discuss why the two tasks yielded different results, and it is likely because the two methodologies were not testing the same thing. The AX task was performed as a replication of Carlson's past studies (Carlson et al., 2016; Carlson, 2018, 2019). He reported that his participants incorrectly responded that e1-e0 stimuli pairs were the same nearly 100% of the time and that e1-a0 stimuli pairs were the same roughly 75% of the time. This was interpreted as evidence for an illusory vowel effect. Since the short vowels did not supply enough acoustic evidence to be categorized as a specific vowel yet participants incorrectly responded that stimuli with these ambiguous vowels were identical to those with robust /e/s, he concluded that the ambiguous vowels were mapped to a prothetic /e/, the illusory vowel. The goal of the present study was then to determine if s+stop and s+liquid clusters were equally as susceptible this illusory vowel effect.

However, unlike the participants in Carlson's studies, the native Spanish speaking participants in the present study showed that they are highly sensitive to the minute, acoustic differences in these same stimuli pairs in the replicated task. Their good discrimination of these stimuli pairs suggests that they may not have been phonologically encoding the stimuli but instead were focusing on the phonetic differences. Various studies have shown that AX discrimination tasks are better for detecting fine-grained, phonetic differences rather than categorical or phonological differences (Gerrits & Schouten, 2004; Davidson & Shaw, 2012). Thus, these results are consistent with the observation that AX discrimination tasks are best suited to test phonetic discrimination and cannot rely on the assumption made by Carlson that discrimination in this particular task reflects phonotactic grammars. As such, the significant effect of cluster type found in the present study should not be interpreted as evidence for different phonological structures of the different sC-clusters.

Nonce word judgment tasks, on the other hand, seem to be better at targeting phonotactic knowledge (Coetzee, 2008; Albright, 2009). Therefore, the results of this task may provide more valid results to answer the present questions on how the language-internal factor of sC-cluster type affects variability in perception. There was no difference in acceptability between s+stop and s+liquid clusters, which does not support the predictions of the HG-GLA that the two types of sC-clusters may have different syllable structures. Because the results of the nonce word judgment task seem to better capture the phonotactic rather than phonetic knowledge of the participants in this study, the remainder of the discussion will focus on the results of this task and not the results of the AX task.

The use of this experimental paradigm can also help explain why past studies on sC-cluster perception in Spanish have reported robust illusory vowel effects but why the present study shows no evidence for this perceptual repair. Many scholars have used a vowel detection task to look at sC-cluster perception (Cuetos Vega et al., 2011; Cuetos et al., 2011; Carlson et al., 2016; Hallé et al., 2013). In these tasks, native Spanish speakers listened to words or nonce words with initial sC-clusters either with no vocalic material before the cluster or with a short, ambiguous initial vowel. They were then asked to determine if they perceived an initial /e/ or not. If participants know they are 'looking for' an initial /e/, this could influence their responses. In a nonce word judgment task, on the other hand, participants are not informed on what part of the word they should pay the most attention to. Rather, they are instructed to respond without overthinking if the word sounds like it could be a word of that language or not. In the present study, both grammatical and ungrammatical fillers were included to distract participants from realizing the task focused on sC-cluster perception.

In addition to vowel detection tasks, lexical decision tasks have also been utilized to investigate sC-cluster perception in Spanish (Hallé et al., 2013; Carlson, 2018, 2019). In these tasks participants listened to real words of Spanish that begin with esC that were presented with or without the initial /e/ (e.g., *especial* vs. *special*). Participants were then asked to determine if the word was a real word of Spanish or not. Although these studies have reported strong illusory vowel effects, this task conflates phonotactic and lexical knowledge. Nonce word judgment tasks do not rely on lexical knowledge and strictly target phonotactic judgments.

Although various tasks have been used to test sC-cluster perception in Spanish, the nonce word judgment task is the one that best tests phonotactic knowledge. The nonce word judgment task performed in this study showed that native Spanish speakers do not accept words with initial sC-clusters as possible words of Spanish. This result lends itself to the interpretation that sC-clusters are not susceptible to an illusory vowel effect (i.e., mis-perceived) in the absence of lexical knowledge. Given the vast amount of work that supports this perceptual illusion in Spanish, this result was unexpected. However, Tetzloff (2020) performed a similar nonce word judgment task with native Spanish listeners where sC-initial words were used as non-target fillers; they were rated as highly unacceptable by that set of participants as well. This further supports the notion that an illusory vowel effect may not be present in the perception of Spanish sC-clusters, as has been previously reported.

7.1.2 Language-internal effect summary

The AX task showed a significant effect of cluster type, with s+liquid clusters yielding worse discrimination than s+stop pairs, while the nonce word judgment task found no difference based on the type of the sC-cluster. Although these results are inconsistent, the results of the nonce word judgment task are interpreted as being more representative of Spanish speakers' phonotactic grammars because the task better targets phonotactic rather than phonetic knowledge. The lack of difference in acceptability between s+stop and s+liquid clusters does not support the predictions of the HG-GLA, but additional experiments are needed to further examine this language-internal variable.

7.2 Question 2: Language profile

The goal of the second research question was to determine if English-Spanish bilinguals differ from Spanish monolinguals in their perception of sC-clusters in Spanish. It was hypothesized that bilinguals and monolinguals would behave differently in their perception of sC-clusters. This hypothesis was supported, as the results of the nonce word judgment task showed that bilinguals rated sC-clusters as more acceptable than their monolingual counterparts. Both monolinguals and bilinguals reported that esC-initial words were the most 'Spanish-like', but bilinguals were more likely to accept sC-initial words. Spanish monolinguals consistently rated sC-initial words as highly unacceptable. The language profile of the speaker appears to be a factor that can predict variability in phonotactics.

It was further predicted that bilinguals would show higher acceptability of s+liquid clusters compared to s+stop clusters since the results of the HG-GLA posited that s+liquid clusters are more often syllabified in the same way in Spanish and English. This hypothesis was not supported by this group of participants for this task.

Past studies have reported that Spanish monolinguals are more accepting of sC-clusters than English-Spanish bilinguals because they have a stronger illusory vowel effect (Carlson, 2018, 2019). The present results are contradictory to those conclusions and instead suggest that Spanish monolinguals are more certain of what is phonotactically allowed or not allowed in Spanish compared to bilinguals. This discrepancy in results, however, may be the result of the type of task used, namely with the nonce word judgment task reflecting phonotactic knowledge un-confounded by lexical knowledge, as discussed in the previous section.

7.2.1 Post-hoc analysis

The higher acceptability of sC-clusters by English-Spanish bilinguals can be interpreted in a few different ways. One possibility is that bilinguals are more likely to perceive an illusory /e/ before sC-intial nonce words, which is consistent with the assumption that is made based on what monolinguals have showed in previous studies. However, in this study the Spanish monolinguals showed no evidence of a perceptual repair. Given that Spanish monolinguals should have a more restrictive grammar overall, compared to bilinguals whose second language allows sC-clusters, an illusory vowel effect would be more likely in the monolinguals.

Another possibility is that because English allows this structure, the bilinguals do not completely separate the two phonotactic systems and use all of their linguistic knowledge, which results in English phonotactics being active in their perception of Spanish.

Alternatively, a third possibility is that because bilinguals have a wider range of what is ‘allowed’ phonotactically compared to the more restricted set of possibilities within a Spanish monolingual, they are more likely to accept all ungrammatical structures. If bilinguals showed more acceptability overall, we could expect that they would also rate other ungrammatical structures better than monolinguals too.

This, however, was not the case. A post-hoc regression analysis was performed to test the difference in acceptability between sC-cluster nonce words and other nonce words with ungrammatical onset clusters. The other ungrammatical clusters included the nonce fillers, which began with /pn, tn, kn, fn/ (referred to collectively as *CC). The mean ratings for sC, *CC, and esC nonce words are shown in Table 7.1 and Figure 7.1. The *CC nonce words were rated as less acceptable than sC-clusters by bilinguals and equally as unacceptable by monolinguals.

Cluster group	Language profile	Mean rating (S.D.)	1	2	3	4
sC	Monolingual	1.26 (0.72)	78.8%	8.8%	5.5%	6.9%
	Bilingual	1.98 (0.96)	38.1%	33.8%	18.6%	0.5%
*CC	Monolingual	1.20 (0.67)	85.5%	9.9%	3.3%	1.3%
	Bilingual	1.45 (0.67)	64.1%	28.5%	6.2%	1.3%
esC	Monolingual	2.54 (0.90)	13.3%	32.7%	39.0%	15.1%
	Bilingual	2.64 (0.97)	14.8%	27.0%	36.5%	21.6%

Table 7.1: Mean rating, standard deviation, and proportion of each response type for sC, *CC, and esC nonce words by language profile.

Table 7.2 shows the results of the ordinal regression where nonce word type (sC, *CC, esC) and language profile (monolingual, bilingual) were set as interacting terms with a random intercept of subject included. These results show that *CC words were not rated as worse than sC words overall (p=0.91), but they were rated worse than esC words. Additionally, there was a main effect of language profile, with the bilinguals having higher acceptability ratings across these nonce word types (p<0.0001). There were significant interaction effects of nonce word type and language profile as well. The difference in rating between sC and *CC words was significantly greater for bilinguals than for monolinguals (p<0.0001), as monolinguals rated sC and *CC words the same.

Contrast against esC and Monolingual	β-estimate	S.E.	z-value	p-value
*CC	-0.03	0.29	-0.11	0.91
esC	3.62	0.19	19.02	<0.0001*
bilingual	2.37	0.35	6.69	<0.0001*
*CC * bilingual	-1.25	0.30	4.12	<0.0001*
esC * bilingual	-2.21	0.20	-11.04	<0.0001*

Table 7.2: Ordinal regression with nonce word type and language profile as interaction terms, monolingual and sC words as reference values. P-values marked with an asterisk indicate a significant effect after correction for multiple comparisons.

The fact that English-Spanish bilinguals rate sC-clusters as more acceptable than *CC clusters demonstrates that they are not just overall more accepting of ungrammatical forms in Spanish but that they are only more accepting of forms that are grammatical in English. This result falls in line with past results in bilingual morphosyntax that have proposed that bilingual variability is often observed in areas of non-identical structural overlap (Hulk &

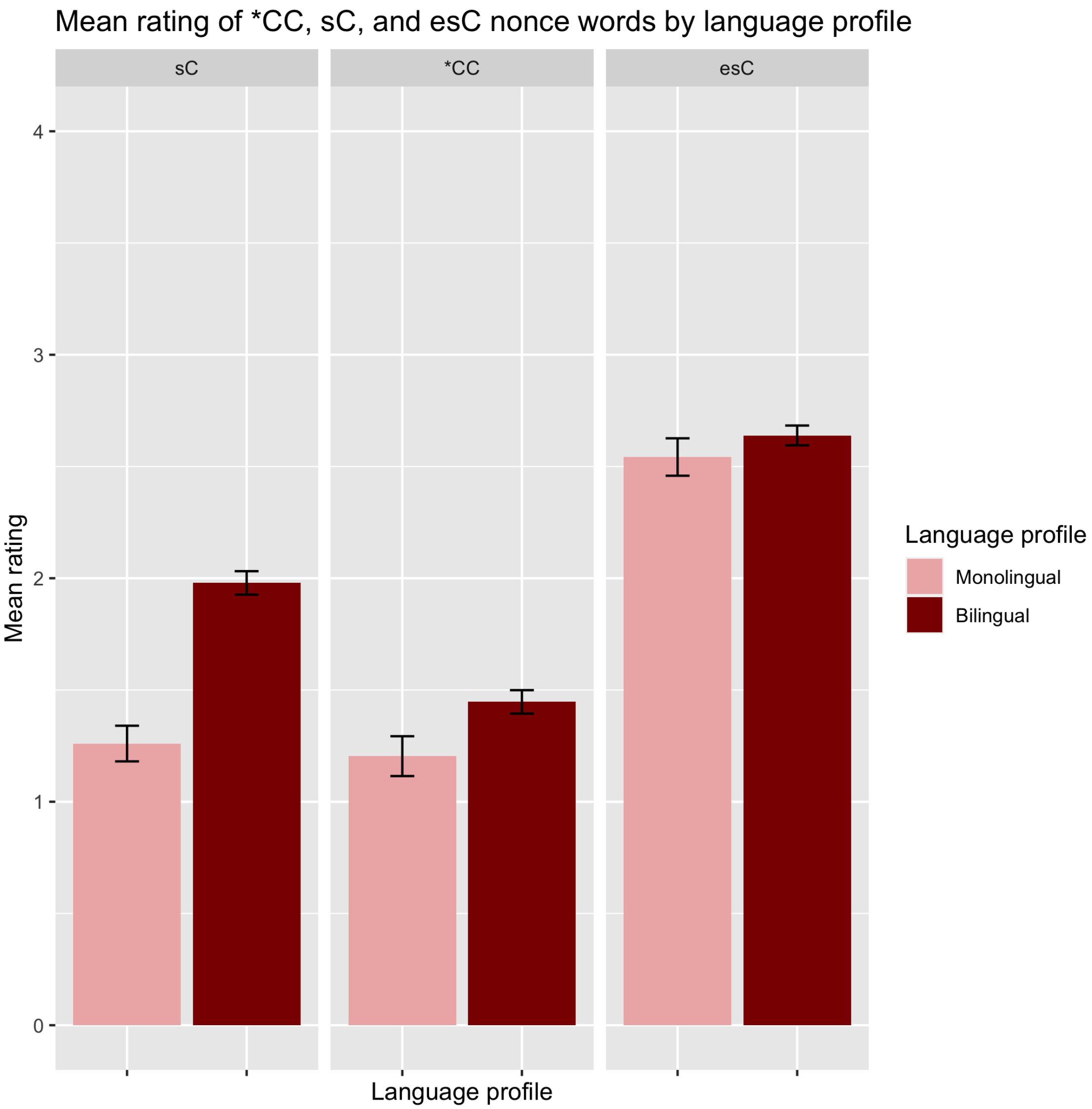


Figure 7.1: Mean rating of sC, *CC, and esC nonce words by language profile. Error bars represent 95% confidene interval.

Müller, 2000; Müller, 2003). With English-Spanish bilinguals, phonotactic forms that are grammatical in both languages are rated as highly acceptable (e.g., esC), and onset clusters that are ungrammatical in both languages are rated as highly unacceptable (*CC). It is with the sC-clusters, whose forms differ between the two languages, where the intermediate acceptability (i.e., phonotactic gradience) for bilinguals but not monolinguals occurs (i.e., bilingual variability).

7.2.2 Language profile summary

Whether a speaker is monolingual or bilingual has an effect on the perception of sC-clusters in Spanish. Spanish monolinguals rated sC-clusters as highly unacceptable, and English-Spanish bilinguals rated them as more acceptable than the monolinguals did but less acceptable than the grammatical esC nonce words. In order to confirm that the bilinguals' intermediate acceptability of the sC-clusters was specific to sC-clusters and not a result of higher overall acceptability, a post-hoc analysis was run comparing the ratings of sC-clusters with ungrammatical *CC clusters. The results showed that bilinguals rated *CC clusters as highly unacceptable, suggesting that English-Spanish bilinguals have gradient acceptability of illicit clusters in Spanish when those clusters are present in English.

7.3 Question 3: Age of exposure to bilingualism

The third research question asked if age of exposure to bilingualism affects the perception of sC-clusters in Spanish. An earlier age of exposure to bilingualism was predicted to correlate with increased bilingual variability in this domain, but results of the present study did not support this hypothesis, as there were no obvious age-related effects in the perception of sC-clusters by the bilingual participants. In the nonce word judgment task, bilinguals performed significantly differently than the monolinguals, but with there were no age-related effects within the different types of bilinguals (i.e., earlier exposure to English did not affect acceptability ratings). Furthermore, within bilinguals, s+stop and s+liquid clusters yielded the same results, indicating that there was no interaction between age of exposure to bilingualism and sC-cluster type.

The bilinguals in this study exhibited variability that distinguished them from the monolinguals, but the lack of age related effects within the bilingual participants could be due to the fact that they all acquired Spanish as an L1 and have had continued exposure throughout their lives. Carlson et al.'s (2016) and Carlson's (2018) studies showed evidence of bilingual variability in early English-Spanish bilinguals, and his following study (2019) showed a similar illusory vowel effect in late English-Spanish bilinguals. However, the two age categories in these studies were never directly compared against each other, so it is unclear if any age effects were present in his cohorts. The bilingual participants in his studies lived in El Paso, Texas, where both English and Spanish are widely spoken, and he noted that the participants had daily exposure to both languages like the participants in the current study. Given the similar results with both the early and late bilinguals across his two studies, it is possible that his participants did not show age related effects either but only an effect of being bilingual, as the results of the current study also show.

Other studies that have investigated the effects of bilingualism on the L1 have concluded that bilinguals tend to behave differently in phonology than monolinguals when there is reduced contact with the L1 (Au, Oh, Knightly, Jun, & Romo, 2008; Stölten, 2013; Ahn, Chang, DeKeyser, & Lee-Ellis, 2017). The participants in these studies were either adopted from another country or immigrated at an early age, resulting in diminished exposure to their L1. In these cases, earlier age of exposure to bilingualism had an effect on L1 language phonology since the exposure to the L1 was reduced or completely eliminated. However, English-Spanish bilinguals in the U.S. often differ from such populations because they can have continuous robust Spanish exposure even in an English-dominant community, particularly in Western Massachusetts where the current study took place. According to the 2020 Census, 54.6% of residents in Holyoke, MA are Hispanic, and this region is home to the largest population of Puerto Ricans outside of Puerto Rico. Thus, earlier age of exposure to L2 English does not necessarily correspond to reduced exposure to L1 Spanish in non-educational spaces for the bilingual participants in this study.

A study with Polish L1 - English L2 early bilinguals who had continued exposure to Polish showed that age of exposure to bilingualism tended to affect the L1 in production but not in perception tasks across linguistic domains (e.g., vocabulary, morphosyntax, and phonology) (Haman et al., 2017); all bilinguals in this study, regardless of age of exposure to L2 English, showed different language outcomes from the monolingual Polish speakers. These results fall in line with the results of the present study, since the present study did not reveal age-related effects in perception.

7.3.1 Age of exposure to bilingualism summary

The age of exposure to bilingualism does not appear to have an effect on the perception of sC-clusters in Spanish. Other studies that have shown age-related effects in the L1 included participants with less continued exposure to the L1 as the L2 was acquired. The bilingual participants in this study have consistent daily exposure and use in Spanish while living in an English language context. Their strong knowledge of Spanish in adulthood can explain why there are no obvious age-related effects in the perception of sC-clusters. Nevertheless, bilingual variability was evident, as the bilingual participants showed different behaviors than the Spanish monolinguals in the perception and acceptability of these sC-clusters.

7.4 Question 4: Language dominance

The final research question of the present study was to examine how language dominance affects bilingual variability as it pertains to sC-cluster perception in English-Spanish bilinguals. It was predicted that the English-dominant bilinguals would show greater bilingual variability, since the studies summarized in Section 3.4.2 showed a trend of increased bilingual variability when the non-dominant language was the target language. The results of the present study showed that Spanish-dominant bilinguals were less likely to accept sC-clusters in Spanish, which supports the hypothesis.

Although the Spanish-dominant participants rated sC-clusters as less acceptable than the balanced and English-dominant bilinguals did, there was no difference in acceptability ratings between these latter two groups. This was unexpected, as it was hypothesized that increased English-dominance would result in more variability. However, these results can best be interpreted by looking more closely at the dominance ratings of the participants in this study. Of the 41 participants who were included in the results of the nonce word judgment task, six were English-dominant, 15 were Spanish-dominant, and 10 were balanced. Language dominance was based on combining self-ratings where participants were asked to rate their speaking and comprehension abilities in each language from one to five. The speaking and listening scores were averaged for an overall Spanish proficiency score and an English proficiency score. The mean self-rating proficiency scores for Spanish and English by language dominance group are shown in Table 7.3. The Spanish-dominant bilinguals in the present study were more Spanish-dominant than the English-dominant bilinguals were in English. In other words, the Spanish-dominant participants had lower English abilities than the English-dominant participants had for Spanish. This imbalance suggests that the Spanish-dominant bilinguals may have been less likely to use any knowledge of English in the perception of Spanish sC-clusters than the balanced or English-dominant bilinguals.

Language dominance	Mean Spanish proficiency (/5)	Mean English proficiency (/5)
Balanced (n=10)	5	5
English-dominant (n=6)	3.75	5
Spanish-dominant (n=15)	5	2.08

Table 7.3: Mean Spanish and English language proficiency, based on self-ratings in speaking and listening comprehension, by language dominance group.

If more English-dominant participants were added to the study who are less proficient in Spanish than those already included, perhaps the difference between dominance groups would become more apparent, especially since variability was predicted to increase with English proficiency. Given these results and the results of past studies, it is predicted that a more robust set of English-dominant participants with less proficiency in Spanish would result in higher acceptability of sC-clusters by this group. This would then fall in line with past results: increased bilingual variability in the non-dominant, target language.

Another way that language dominance could be examined is on a linear scale rather than discrete groups. The mean English and Spanish ratings could be used or a composite score of both language proficiencies. This would better capture the nuances in bilingual variability associated with different degrees of language dominance. However, such an analysis would require a very large sample size.

7.4.1 Language dominance summary

The Spanish-dominant bilinguals in this study tended to rate sC-clusters as less acceptable than the balanced and English-dominant bilinguals did, but the balanced and English-dominant bilinguals showed similar acceptability ratings to the Spanish sC-clusters. This could be due to the fact that within the English-dominant participants, Spanish proficiency was relatively high. Expanding the sample to include more English-dominant participants may yield the predicted results of increased bilingual variability with increased English dominance.

7.5 Representing bilingual phonotactics in generative phonology

Any adequate generative theory of grammar must be able to account for all speakers, including bilinguals. Most models, including traditional OT and HG, are typically used to account for grammars of monolinguals. The results of the present study showed that bilingual variability is present in the phonotactic systems of English-Spanish bilinguals, evidenced by their higher acceptability of ungrammatical sC-clusters in Spanish compared to Spanish monolinguals. How can this be modeled using a generative model of phonology like HG?

As discussed in Chapter 3, past researchers have proposed sociolinguistic constraints that penalize output candidates based on pragmatic factors like the dominant language in the geographical region or the language dominance of the other interlocuter (Gonzales-Diaz, 2006; Muysken, 2013). This type of constraint would not be able to account for the bilingual variability observed in the perception of sC-clusters because the nonce word judgment task targeted phonotactic judgments and was not embedded in any sort of discourse. Furthermore, this type of pragmatic constraint was intended to predict the language of speech production in code-switching and code-mixing contexts and not language perception.

The first way to try and account for bilingual phonotactic grammars being different from those of monolinguals would be to group all language input together in one system, but some of the earliest work in this field concluded that bilinguals do not have one phonological systems for their two languages (Paradis, 2001). Feeding input from both Spanish and

English into one phonological grammar would not be able to predict the observed bilingual variability. For example, if the HG-GLA used in Chapter 2 were given both Spanish and English input together, it would assign more probability to sC-cluster outputs in Spanish, but it would also assign some probability to the mapping of sC-clusters to esC outputs in English (e.g., a word like *smile* could be mapped to *esmile* in English). English never inserts a prothetic vowel before sC-clusters like Spanish does, so the grammar should not over generate and predict this.

The opposite of this would be to have entirely separate phonologies (i.e., language-specific constraint sets and weights) but would have no way to capture the bilingual variability observed either. If bilinguals had two autonomous systems, the English-Spanish bilinguals in the present study should have behaved identically to the Spanish monolinguals, which also was not the case. The vast majority of research in this field has agreed that bilinguals have two separate phonological systems that interact, yet very few proposals on how to model this interaction exist.

7.5.1 Current proposal

A generative model of grammar needs to be able to predict where bilingual variability will occur and where it will not. The data presented in this study showed that bilingual variability occurs in the perception of sC-clusters in Spanish by English-Spanish bilinguals. Bilingual variability was observed when the grammars were in conflict (e.g., sC-cluster perception) but not when the grammars were in agreement (e.g., esC-words, *CC-words). My only partially-developed proposal for how to capture this interaction between the two phonological grammars using MaxEnt probabilities in an HG framework is as follows.

All inputs come with a language-specific tag (Hsin, 2014); for example, *escuela*-$_{Spa}$ and 'school'-$_{Eng}$ would be tagged as Spanish and English words, respectively. It is not implausible to have language tags on words, since by one year of age bilingual babies are able to discriminate between their two languages (Bosch & Sebastián-Gallés, 2001, 2003b; Molnar, Gervain, & Carreiras, 2014). There are still unique, language-specific constraint weightings for each language. The tagged inputs are then evaluated in both sets of constraints. When the winning outputs of the two languages are the same, the MaxEnt probabilities assigned to the output candidates of the language that the input was tagged with remain. For example, an input of /esC/-$_{Spa}$ yields a faithful output ([esC]) in both English and Spanish. Because the winning outputs are the same, the MaxEnt probabilities assigned by the Spanish grammar do not change.

When the input is ungrammatical in both languages, again the winning candidates of each language are expected to be the same regardless of the language-tag on the input. An input like /pn/$_{Eng}$ would be repaired in both Spanish and English by the insertion of an epenthetic vowel to break apart the illicit onset cluster ([pVn]). Since both languages' constraint weightings agree on the winning candidate, the MaxEnt probabilities assigned by the English grammar would remain.

Up to this point, there is no difference between the proposed grammar and a monolingual grammar, since there are no conflicts between the two languages' phonotactic systems. Where this changes is when the predicted outputs differ between the two languages, like in the case of sC-clusters in English and Spanish. Under the Spanish constraint weighting, the winning candidate for an input of /sC/$_{Spa}$ is [esC], but under the English constraint weighting, the winning candidate would be [sC]. This would trigger the model and force it to penalize the winning candidate of the non-target language (English) less than other losing candidates in the target language (Spanish). This would result in more probability being assigned to the non-target (English) output form for bilinguals but not monolinguals, yielding some degree of bilingual variability.

This type of model could be developed in a similar way to the current HG-GLA, as an error-driven learner, meaning that when the learner's predicted output does not match the actual output, the weights of the constraint(s) responsible are adjusted so that that specific error is less likely in the future. However, in a bilingual model, if the output of the other language's constraint ranking is penalized less than other incorrect outputs, over time the majority of the probability would be split between the winners of the two language: this is the bilingual variability observed.

Alternatively, a model like the Dual Route model proposed in Becker and Tessier (2011) and Becker (2012) could also be adapted to yield the observed bilingual variability, as this type of model allows for variable outputs within one language. In the Dual Route model, there are two routes to selecting an output candidate: select the most harmonic candidate based on the current grammar or select an output from a Cache, where all previously produced forms are stored. Every time the learner produces a new output form with the current grammar, that form is stored in the Cache. Rather than triggering the learning update every time the learner's winner is different from the target output like the HG-GLA does, learning is triggered when a markedness constraint wrongly prefers a predetermined number of errors in the Cache. The learner then selects an error from the Cache to use to re-rank or re-weight the markedness constraint responsible for the error. That error is then removed from the Cache added to Support, the permanent repository of learning data. This model assumes that errors in the Cache that are not moved to Support decay over time. As these errors gradually decay, the probability of them being selected as an output from the Cache becomes less and less likely. This results in three stages of learning: 1) incorrect outputs, 2) variable outputs, 3) stable, target outputs.

Because the Dual Route model already allows for variability in output selection, it could potentially be adapted for bilingual grammars. In the case of Spanish sC-clusters, Spanish monolinguals would converge on a grammar that does not permit the surfacing of sC-clusters, and any error forms of sC-clusters in the Cache would decay over time. Bilinguals, however, would have continued evidence for sC-clusters in English, which would be stored in Support. It is plausible to imagine a system where bilinguals remain in the stage of variable outputs if the errors in the Cache for one language were also present in the Support of the other language.

This proposal is clearly not fully fledged out but remains as a goal for future research. Furthermore, such a model would need to be adapted for phonological alternations (i.e., not phonotactics) so that Spanish alternations do not start appearing in English and vice versa, if they are unattested examples of bilingual variability.

CHAPTER 8

CONCLUSIONS

8.1 General summary

All in all, the results of the current study culminated to one clear conclusion: bilingual variability is clearly present in phonotactics, as evidenced by the different behavior of English-Spanish bilinguals compared to Spanish monolinguals in the perception of sC-clusters in Spanish.

The type of sC-cluster was not shown to affect variability of sC-cluster perception in monolinguals or bilinguals. The HG-GLA predicted that s+stop clusters are more likely to be repaired with a prothetic /e/ compared to s+liquid clusters, but the experimental results showed that the two types of sC-clusters were treated the same in perception. This sheds light on the theoretical assumptions of sC-cluster phonotactics and supports a model where s+stop and s+liquid clusters are syllabified in the same way and have parallel output candidates.

The bilingual variability observed in the perception of sC-clusters was not conditioned by age of exposure to bilingualism, but rather it was an effect of simply being bilingual. English-Spanish bilinguals were more accepting of sC-initial words in Spanish since sC-clusters are licit in English. They are not, however, more accepting of all ungrammatical Spanish structures, which was evident from their low acceptability ratings of other illicit onset clusters. Language dominance showed a mild effect on bilingual variability, with decreased knowledge of English resulting in less bilingual variability in Spanish.

Finally, a sketch of a proposal for how to predict bilingual variability in a generative model of phonology was presented. Such a model does not assume one overarching phonology for both languages nor two autonomous language systems. Rather, it potentially resolves the issue that current models have failed to address: predicting non-random bilingual variability in observed areas, like when phonotactic constraints are in conflict across the two languages. Further development of this idea will be valuable to theoretical phonology, bilingualism, and their intersection.

8.2 Future directions

Although the present study provided empirical evidence on the nature of bilingual variability in phonotactics, there were various limitations with the experimental methods.

First, the use of the AX task did not target the phonological knowledge of the participants but rather their ability to distinguish between phonetically-different stimuli. As discussed, the AX task was chosen because the goal was to replicate the AX tasks presented in Carlson et al. (2016) and Carlson's (2018, 2019) studies. An ABX task would have been a better method for answering the questions of this study. ABX tasks are similar to AX tasks except that rather than comparing two adjacent stimlui, participants hear three stimuli and are asked to determine if the third (X) is the same as the first (A) or the second (B). This type of task is much more demanding on working memory and, as a result, has been shown to require some degree of phonological encoding (Davidson & Shaw, 2012). Replicating the AX task in an ABX paradigm may yield different results that provide better insight on the phonotactic, rather than phonetic, perception in bilinguals.

Additionally, it is worth exploring the predictions made by the HG-GLA with respect to the syllable structure of s+stop and s+liquid clusters in Spanish. The results of the HG-GLA assigned one third of the probability for s+stop clusters to an output with a bisyllabic structure ([es.p] versus [e.sp]), which is what was expected given past analyses of Spanish phonotactics. This was not the case for s+liquid clusters, as all probability was assigned to the taughtosyllabic structures ([sl] and [e.sl]). The experimental results did not show any observable difference in the perception of s+stop versus s+liquid clusters, but given that syllable structure is an example of hidden structure, further examinations could help determine if the syllabification predictions of the HG-GLA are observable. One way to test this would be through a production study of esC-initial words since only sC-cluster productions have been analyzed in previous studies. If the syllabification of s+stop and s+liquid are in fact different in Spanish, where s+stop is sometimes syllabified bisyllabically as the HG-GLA predicted, production may vary between these stimuli types. Many dialects of Spanish have significant /s/ aspiration or deletion in coda position. Thus, a word like *escuela* may be realized as [e.kwe.la], with the omission of the coda /s/. If speakers from an aspirating dialect aspirate or delete /s/ in s+stop clusters but not in s+liquid clusters (e.g., *escuela* as [e(h).kwe.la] but *eslora* as [e(h).lo.ra]), this would shed light on where the syllable boundary truly lies. If the two types of sC-clusters are syllabified differently in Spanish, /s/ aspiration should be present more often in es+stop but not in es+liquid contexts. Furthermore, conducting a similar study on the perception of sC-clusters in Spanish in a clinical population may also uncover whether or not there is a difference in the hidden structure of s+stop and s+liquid clusters in Spanish, as increased variability is often observed in these atypical populations.

Finally, I plan to implement my proposal for a phonological learning algorithm that is able to predict bilingual variability. This will be very useful for predicting other specific areas of bilingual variability in bilingual phonology of any two language pairings, which will be beneficial for future empirical studies examining bilingual variability in phonology.

APPENDIX A

LANGUAGE AND SOCIAL BACKGROUND QUESTIONNAIRE

Modified from Anderson et al. (2018)

1. **Participant's age** ______________

2. **Participant's place of birth/youth** ______________

   (a) If born/raised in a Spanish-speaking country,

   At what age did you arrive in the US? ______________

   What is the total time you have resided in the US? ______________

   (b) If born/raised in the US,

   Indicate any time spent in Spanish-speaking countries (year(s) and duration(s) of stay):

3. **Participant's occupation** ______________

4. **Participant's highest level of education** (select one)

   Primary Secondary College/Technical University

5. **Participant's language background**

   (a) **Languages and overall fluency rating**

   First Language ______________

Understand only   Poor fluency   Adequate fluency   Good fluency   Complete fluency/native

Second Language ________________

Age began learning 2nd language ________________

Understand only   Poor fluency   Adequate fluency   Good fluency   Complete fluency/native

Third Language ________________

Age began learning 3rd language ________________

Understand only   Poor fluency   Adequate fluency   Good fluency   Complete fluency/native

(b) **Home language(s) as child/adolescent**

English only   Mostly Eng/some Spa   Equal Eng/Spa   Mostly Sp/some Eng   Spanish only

Other: ________________

(c) **Language(s) spoken in elementary/middle school**

English only   Mostly Eng/some Spa   Equal Eng/Spa   Mostly Sp/some Eng   Spanish only

Other: ________________

(d) **Language(s) spoken in high school**

English only   Mostly Eng/some Spa   Equal Eng/Spa   Mostly Sp/some Eng   Spanish only

Other: ________________

(e) **Language(s) of post-secondary schooling**

English only   Mostly Eng/some Spa   Equal Eng/Spa   Mostly Sp/some Eng   Spanish only

Other: ________________

(f) **If your main language of school was not English, how did you learn it?** (select all that apply)

At home in youth   Classroom learning   In English community   Other ___________

(g) **If your main language of school was not Spanish, how did you learn it?** (select all that apply)

At home in youth   Classroom learning   In Spanish community   Other ___________

(h) **Have you ever taken grammar/language classes?**

In English? No _____   Yes (specify) _____________________

In Spanish? No _____   Yes (specify) _____________________

6. **Participant's current language use**

(a) **At home**

English only   Mostly Eng/some Spa   Equal Eng/Spa   Mostly Sp/some Eng   Spanish only

Other: ________________

(b) **At school**

English only   Mostly Eng/some Spa   Equal Eng/Spa   Mostly Sp/some Eng   Spanish only

Other: ________________

(c) **At work**

English only   Mostly Eng/some Spa   Equal Eng/Spa   Mostly Sp/some Eng   Spanish only

Other: ________________

(d) **In social situations**

English only Mostly Eng/some Spa Equal Eng/Spa Mostly Sp/some Eng Spanish only

Other: _______________

(e) **Language you watch TV/movies in**

English only Mostly Eng/some Spa Equal Eng/Spa Mostly Sp/some Eng Spanish only

Other: _______________

(f) **Language you currently feel most comfortable with**

English only Mostly Eng/some Spa Equal Eng/Spa Mostly Sp/some Eng Spanish only

Other: _______________

7. **Participant's language skills**

(a) **Speaking**

Spanish

Poor/limited Basic Adequate Good Excellent/native

English

Poor/limited Basic Adequate Good Excellent/native

(b) **Listening**

Spanish

Poor/limited Basic Adequate Good Excellent/native

English

Poor/limited Basic Adequate Good Excellent/native

(c) **Reading**

Spanish

Poor/limited   Basic   Adequate   Good   Excellent/native

English

Poor/limited   Basic   Adequate   Good   Excellent/native

(d) **Writing**

Spanish

Poor/limited   Basic   Adequate   Good   Excellent/native

English

Poor/limited   Basic   Adequate   Good   Excellent/native

Please make any comments about your past or current language use, acquisition, or learning that you think are important to know (especially concerning English and Spanish)

8. **Family language background and fluency rating**

(a) **Mother**

Place of birth ________________

First language ________________

Understand only   Poor fluency   Adequate fluency   Good fluency   Complete fluency/native

Second language ________________

Understand only   Poor fluency   Adequate fluency   Good fluency   Complete fluency/native

Other language ________________

Understand only   Poor fluency   Adequate fluency   Good fluency   Complete fluency/native

(b) **Father**

Place of birth ________________

First language ________________

Understand only   Poor fluency   Adequate fluency   Good fluency   Complete fluency/native

Second language ________________

Understand only   Poor fluency   Adequate fluency   Good fluency   Complete fluency/native

Other language ________________

Understand only   Poor fluency   Adequate fluency   Good fluency   Complete fluency/native

(c) **Other main caretaker/guardian**

Place of birth ________________

First language ________________

Understand only   Poor fluency   Adequate fluency   Good fluency   Complete fluency/native

Second language ________________

Understand only   Poor fluency   Adequate fluency   Good fluency   Complete fluency/native

Other language ________________

Understand only   Poor fluency   Adequate fluency   Good fluency   Complete fluency/native

APPENDIX B

NONCE WORD STIMULI PAIRS FOR AX TASK

Identical pairs	**Duration different**	**Vowel different**	**Duration & vowel different**
aspid - aspid	aspid - a:spid	aspid - espid	aspid - e:spid
espid - espid	espid - e:spid	espid - aspid	espid - a:spid
asfid -asfid	asfid - a:sfid	asfid - esfid	asfid - e:sfid
esfid -esfid	esfid - e:sfid	esfid - asfid	esfid - a:sfid
asmid - asmid	asmid - a:smid	asmid - esmid	asmid - e:smid
esmid - esmid	esmid - e:smid	esmid - asmid	esmid - a:smid
aslid - aslid	aslid - a:slid	aslid - eslid	aslid - e:slid
eslid - eslid	eslid - e:slid	eslid - aslid	eslid - a:slid
a:spid - a:spid	a:spid - aspid	a:spid - e:spid	a:spid - espid
e:spid - e:spid	e:spid - espid	e:spid - a:spid	e:spid - aspid
a:sfid -a:sfid	a:sfid - asfid	a:sfid - e:sfid	a:sfid - esfid
e:sfid -e:sfid	e:sfid - esfid	e:sfid - a:sfid	e:sfid - asfid
a:smid - a:smid	a:smid - asmid	a:smid - e:smid	a:smid - esmid
e:smid - e:smid	e:smid - esmid	e:smid - a:smid	e:smid - asmid
a:slid - a:slid	a:slid - aslid	a:slid - e:slid	a:slid - eslid
e:slid - e:slid	e:slid - eslid	e:slid - a:slid	e:slid - aslid

APPENDIX C

NONCE WORDS FOR JUDGMENT TASK

spandas
spasos
spato
spetas
spojas
spona
spujo
staros
steces
stenta
stias
stines
stolsas
stuyas
skargos
skene
skistes
skomas
skormas
sku no
skura
smanza
smeto
smilas
smobra
smocios
smoma
smuta
snancos
snara
snelta
snira
snitos
snobo
snotos
snuga
snumba
snupos
snurso
snurva
snusta
slajes
slancha
slantos
slara
slata
sleda
slegas
sleres
slicho
slitos
slobo
sloco
slodas
sloga
slojas
slotos
slumba
slurso
slusta
espandas
espasos
espato
espetas
espojas
espona
espujo
estaros
esteces
estenta
estias
estines
estolsas
estuyas
eskargos
eskistes
eskobra
eskoma
eskomas
eskormas

esku no
esmanza
esmere
esmeto
esmilas
esmocios
esmura
esmuta
esnancos
esnara
esnelta
esnira
esnitos
esnobo
esnotos
esnuga
esnumba
esnupos
esnurso
esnurva
esnusta

eslajes
eslancha
eslantos
eslara
eslata
esleda
eslegas
esleres
eslicho
eslitos
eslobo
esloco
eslodas
esloga
eslojas
eslotos
eslumba
eslurso
eslusta
bnantos
bnenta

bneto
bnitos
fnajes
fnobra
fnujo
fnupos
knadres
knargos
knasos
knuga
pnancos
pnara
pneda
pnormas
tnaros
tnobo
tnodas
tnuta
flancos
flandas
flormas

flurva
freda
frines
froga
fromas
klargos
klato
klistes
klocios
klusta
krancha
krelta
krilas
krodas
planza
plojas
plotos
pluyas
prajes
pru no
prujo

prupos
trara
tricho
troco
trojas
anadres
anargos
anona
ansistes
anuga
aspistes
aspojas
aspotos
aspura
aspuyas
calantos
caleto
calias
calitos
conines
conoga

conomas
conumba
enato
enocios
enurso
enurva
insara
insojas
insuyas
kantos
keto
kias
kitos
lato
legas
locios
lusta
manza
meces
mojas
motos
nara
noco
nojas
noma
pargos
pesanchas
pesaros
pesodas
pesuta
petas
pona
puga
pusta
randas
remancos
remandas
remicho
remormas
reres
rines
romas
sadres
sajes
su no
supos
talancha
talata
talilas
talolsas
tanchas
taros
tuta
xancha
xelta
xene
xolsas

REFERENCES

Abramson, A. S., & Lisker, L. (1965). Voice onset time in stop consonants: Acoustic analysis and synthesis. In *Proceedings of the 5th International Congress of Acoustics* (Vol. 51, p. A51).

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of neurolinguistics*, *20*(3), 242–275.

Ahn, S., Chang, C. B., DeKeyser, R., & Lee-Ellis, S. (2017). Age effects in first language attrition: Speech perception by Korean-English bilinguals. *Language Learning*, *67*(3), 694–733.

Albright, A. (2009). Feature-based generalisation as a source of gradient acceptability. *Phonology*, *26*(1), 9–41.

Altenberg, E. P., & Cairns, H. S. (1983). The effects of phonotactic constraints on lexical processing in bilingual and monolingual subjects. *Journal of Verbal Learning and Verbal Behavior*, *22*(2), 174–188.

Anderson, D. (1980). Cohesion and the cloze test. *Unpublished MA thesis, University of Illinois*.

Anderson, J., Mak, L., Chahi, A. K., & Bialystok, E. (2018). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods*, *50*(1), 250–263.

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(2), 234.

Anttila, A. (2008). Gradient phonotactics and the Complexity Hypothesis. *Natural Language & Linguistic Theory*, *26*(4), 695–729.

Archibald, J. (2014). *Phonological acquisition and phonological theory.* Psychology Press.

Au, T. K.-f., Oh, J. S., Knightly, L. M., Jun, S.-A., & Romo, L. F. (2008). Salvaging a childhood language. *Journal of Memory and Language*, *58*(4), 998–1011.

Bachman, L. F. (1982). The trait structure of cloze test scores. *Tesol Quarterly*, *16*(1), 61–70.

Bachman, L. F., & Palmer, A. S. (1989). The construct validation of self-ratings of communicative language ability. *Language Testing*, *6*(1), 14–29.

Bailey, T. M., & Hahn, U. (2001). Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language*, *44*(4), 568–591.

Baker, W., & Trofimovich, P. (2005). Interaction of native-and second-language vowel system (s) in early and late bilinguals. *Language and Speech*, *48*(1), 1–27.

Barlow, J. A. (1997). *A constraint-based account of syllable onsets: Evidence from developing systems.* (Unpublished doctoral dissertation). Ph.D. Dissertation at Indiana University.

Barlow, J. A. (2001). The structure of/s/-sequences: Evidence from a disordered system. *Journal of Child Language*, *28*(2), 291–324.

Barlow, J. A., & Gierut, J. A. (1999). Optimality Theory in phonological acquisition. *Journal of Speech, Language, and Hearing Research*, *42*(6), 1482–1498.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01

Beardsmore, H. B. (1986). *Bilingualism: Basic principles* (Vol. 1). Multilingual Matters.

Becker, M. (2012). Target selection in error selective learning. *Brill's Journal of Afroasiatic Languages and Linguistics*, *4*(1), 120–139.

Becker, M., Ketrez, N., & Nevins, A. (2011). The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language*, 84–125.

Becker, M., & Tessier, A.-M. (2011). Trajectories of faithfulness in child-specific phonology. *Phonology*, *28*(2), 163–196.

Berent, I., Lennertz, T., Jun, J., Moreno, M. A., & Smolensky, P. (2008). Language universals in human brains. *Proceedings of the National Academy of Sciences*, *105*(14), 5321–5325.

Berent, I., Lennertz, T., Smolensky, P., & Vaknin-Nusbaum, V. (2009). Listeners' knowledge of phonological universals: Evidence from nasal clusters. *Phonology*, *26*(1), 75–108.

Berent, I., Steriade, D., Lennertz, T., & Vaknin, V. (2007). What we know about what we have never heard: Evidence from perceptual illusions. *Cognition*, *104*(3), 591–630.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America*, *109*(2), 775–794.

Best, C. T., et al. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, *167*(224), 233–277.

Bhatt, R. M., & Bolonyai, A. (2011). Code-switching and the optimal grammar of bilingual language use. *Bilingualism: Language and Cognition*, *14*(4), 522–546.

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *13*(4), 525–531.

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language*, 706–755.

Birdsong, D. (1999). *Second language acquisition and the critical period hypothesis.* Routledge.

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and language*, *44*(2), 235–249.

Blevins, J. (1995). The syllable in phonological theory. In *Handbook of phonological theory* (pp. 206–244). Blackwell.

Bloch, T. (2011). *Simplification strategies in the acquisition of consonant clusters in Hebrew.*

Blutner, R. (2000). Some aspects of optimality in natural language interpretation. *Journal of Semantics*, *17*(3), 189–216.

Blutner, R., Bezuidenhout, A., Breheny, R., Glucksberg, S., & Happé, F. (2003). *Optimality Theory and pragmatics.* Springer.

Boersma, P. (1998). *Functional phonology: Formalizing the interactions between articulatory and perceptual drives* (Unpublished doctoral dissertation). University of Amsterdam.

Boersma, P. (2000). Learning a grammar in functional phonology. *Optimality Theory: Phonology, syntax, and acquisition*, 465–523.

Boersma, P., & Hayes, B. (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry*, *32*(1), 45–86.

Boersma, P., Levelt, C., et al. (2000). Gradual constraint-ranking learning algorithm predicts acquisition order. In *Proceedings of Child Language Research Forum* (Vol. 30, pp. 229–237).

Boersma, P., et al. (1997). How we learn variation, optionality, and probability. In *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* (Vol. 21, pp. 43–58).

Boersma, P., & Pater, J. (2008). *Convergence properties of a gradual learning algorithm for Harmonic Grammar.*

Boersma, P., & Weenink, D. (2019). Praat: Doing phonetics by computer (version 6.1.13)[computer software]. *Amsterdam: Institute of Phonetic Sciences*.

Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, *2*(1), 29–49.

Bosch, L., & Sebastián-Gallés, N. (2003a). Language experience and the perception of a voicing contrast in fricatives: Infant and adult data. In *Proceedings of the 15th International Conference of Phonetic Sciences* (pp. 1987–1990).

Bosch, L., & Sebastián-Gallés, N. (2003b). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, *46*(2-3), 217–243.

Bradley, T. G. (2006). Spanish complex onsets and the phonetics-phonology interface. *Optimality-theoretic studies in Spanish phonology*, 15–38.

Breen, M., Kingston, J., & Sanders, L. D. (2013). Perceptual representations of phonotactically illegal syllables. *Attention, Perception, & Psychophysics*, *75*(1), 101–120.

Bresnan, J. (2000). Optimal syntax. *Optimality theory: Phonology, syntax and acquisition*, 334–385.

Broselow, E. (1992). Transfer and universals in second language epenthesis. *Language Transfer in Language Learning*, 71–86.

Brown, C. (2000). The interrelation between speech perception and phonological acquisition from infant to adult. *Second Language Acquisition and Linguistic Theory*, *1*, 4–63.

Bullock, B. E., & Toribio, A. J. (2004). Introduction: Convergence as an emergent property in bilingual speech. *Bilingualism: Language and Cognition*, *7*(2), 91–93.

Bureau, U. C. (2020). *2020 census.* U.S. Department of Commerce.

Burt, M., & Dulay, H. (1978). Some guidelines for the assessment of oral language proficiency and dominance. *TESOL quarterly*, 177–192.

Cardoso, W., & Liakin, D. (2009). When input frequency patterns fail to drive learning: The acquisition of sC onset clusters. *Recent Research in Second Language Phonetics/Phonology: Perception and Production*, 174–202.

Carlisle, R. S. (1991). The influence of environment on vowel epenthesis in Spanish/English interphonology. *Applied Linguistics*, *12*(1), 76–95.

Carlson, M. T. (2018). Making room for second language phonotactics: Effects of L2 learning and environment on first language speech perception. *Language and Speech*, *61*(4), 598–614.

Carlson, M. T. (2019). Now you hear it, now you don't: Malleable illusory vowel effects in Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *22*(5), 1101–1122.

Carlson, M. T., Goldrick, M., Blasingame, M., & Fink, A. (2016). Navigating conflicting phonotactic constraints in bilingual speech perception. *Bilingualism: Language and Cognition*, *19*(5), 939–954.

Castilla-Earls, A., Pérez-Leroux, A. T., Martinez-Nieto, L., Restrepo, M. A., & Barr, C. (2020). Vulnerability of clitics and articles to bilingual effects in typically developing Spanish–English bilingual children. *Bilingualism: Language and Cognition*, *23*(4), 825–835.

Cenoz, J., & Genesee, F. (2001). *Trends in bilingual acquisition* (Vol. 1). John Benjamins Publishing.

Chaudron, C., Prior, M., & Kozok, U. (2005). Elicited imitation as an oral proficiency measure. In *14th World Congress of Applied Linguistics, Madison, WI.*

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, *27*(2), 207–229.

Chomsky, N. (1965). Aspects of the theory of syntax. *Cambridge, MA: MITPress.*

Chomsky, N. (1965/2014). *Aspects of the theory of syntax* (Vol. 11). MIT press.

Chomsky, N., & Halle, M. (1965). Some controversial questions in phonological theory. *Journal of Linguistics*, *1*(2), 97–138.

Chomsky, N., & Halle, M. (1968). The sound pattern of English.

Christensen, R. H. B. (2018). *ordinal—regression models for ordinal data.* (R package version 2018.8-25. http://www.cran.r-project.org/package=ordinal/)

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. *Papers in Laboratory Phonology*, *1*, 283–333.

Coetzee, A. W. (2008). Grammaticality and ungrammaticality in phonology. *Language*, 218–257.

Coffey, N. (2009, Mar). *The syllable in Spanish.* Retrieved from `http://www.lettermeister.javamex.com/spanish-pronunciation/spanish-syllable.shtml`

Coleman, J., & Pierrehumbert, J. (1997). Stochastic phonological grammars and acceptability. *arXiv preprint cmp-lg/9707017*.

Colomé, À. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language*, *45*(4), 721–736.

Crain, S., & Thornton, R. (2000). *Investigations in universal grammar: A guide to experiments on the acquisition of syntax and semantics.* MIT Press.

Cuetos, F., Glez-Nosti, M., Barbón, A., & Brysbaert, M. (2012). SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica*, *33*(2), 133–143.

Cuetos, F., Hallé, P. A., Domínguez, A., & Segui, J. (2011). Perception of prothetic /e/in # sC utterances: Gating data. In *Icphs* (pp. 540–543).

Cuetos Vega, F., Glez Nosti, M., Barbón Gutiérrez, A., & Brysbaert, M. (2011). SUBLTEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica: Revista de metodología experimental*, *32*, 133-143.

Cyran, E., & Gussmann, E. (1999). Consonant clusters and governing relations: Polish initial consonant sequences. *The Syllable: Views and Facts*, *45*, 219.

Daland, R., Hayes, B., White, J., Garellek, M., Davis, A., & Norrmann, I. (2011). Explaining sonority projection effects. *Phonology*, *28*(2), 197–234.

Davidson, L. (2006). Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics*, *34*(1), 104–137.

Davidson, L. (2011). Phonetic, phonemic, and phonological factors in cross-language discrimination of phonotactic contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(1), 270.

Davidson, L., & Shaw, J. A. (2012). Sources of illusion in consonant cluster perception. *Journal of Phonetics*, *40*(2), 234–248.

Davies, M. (1995). Analyzing syntactic variation with computer-based corpora: The case of modern Spanish clitic climbing. *Hispania*, 370–380.

Dawson, J. I., Stout, C. E., & Eyer, J. A. (2003). *SPELT-3: Structured photographic expressive language test.* Janelle Publications DeKalb, IL.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, *22*(4), 499–533.

De Lacy, P. (2006). *Markedness: Reduction and preservation in phonology* (Vol. 112). Cambridge University Press.

Dell, F., & Elmedlaoui, M. (1985). Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics*, *7*(2), 105–130.

Dell, F., & Elmedlaoui, M. (1988). Syllabic consonants in Berber: Some new evidence. *Journal of African Languages and Linguistics*, *10*(1), 1–17.

Dell, F., & Elmedlaoui, M. (2012). *Syllables in Tashlhiyt Berber and in Moroccan Arabic* (Vol. 2). Springer Science & Business Media.

Deuchar, M., & Quay, S. (1998). One vs. two systems in early bilingual syntax: Two versions of the question. *Bilingualism: Language and Cognition*, *1*(3), 231–243.

Drummond, A. (2013). *Ibex farm.* http://spellout. net/ibexfarm.

Duanmu, S. (2002). Two theories of onset clusters. *Chinese Phonology*, *11*, 97–120.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, *25*(6), 1568.

Durvasula, K., & Kahng, J. (2015). Illusory vowels in perceptual epenthesis: The role of phonological alternations. *Phonology*, *32*(3), 385–416.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*(3968), 303–306.

Escartín Ortiz, C. I. (2005). *The development of sc onset clusters in Spanish and English* (Unpublished doctoral dissertation). Concordia University.

Fabiano, L., & Goldstein, B. (2005). Phonological cross-linguistic effects in bilingual Spanish–English speaking children. *Journal of Multilingual Communication Disorders*, *3*(1), 56–63.

Ferjan Ramírez, N., Ramírez, R. R., Clarke, M., Taulu, S., & Kuhl, P. K. (2017). Speech discrimination in 11-month-old bilingual and monolingual infants: A magnetoencephalography study. *Developmental Science*, *20*(1), e12427.

Fikkert, P. (1994). *On the acquisition of prosodic structure* (Unpublished doctoral dissertation).

Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *The Journal of the Acoustical Society of America*, *84*(1), 70–79.

Flege, J. E. (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.

Flege, J. E. (1999). Age of learning and second language speech. In *Second language acquisition and the critical period hypothesis* (pp. 111–142). Routledge.

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*, *6*, 319–355.

Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. *Laboratory Phonology*, *9*, 353–382.

Flege, J. E., & Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, *6*(3), 185–202.

Flege, J. E., MacKay, I. R., & Piske, T. (2002). Assessing bilingual dominance. *Applied Psycholinguistics*, *23*(4), 567–598.

Flege, J. E., Munro, M. J., & MacKay, I. R. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America*, *97*(5), 3125–3134.

Flege, J. E., Schirru, C., & MacKay, I. R. (2003). Interaction between the native and second language phonetic subsystems. *Speech communication*, *40*(4), 467–491.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, *41*(1), 78–104.

Fleischhacker, H. (2001a). Cluster-dependent epenthesis asymmetries. *UCLA Working Papers in Linguistics*, *7*, 71–116.

Fleischhacker, H. (2001b). Experimental results on relative similarity. *Paper presented at Linguistic Society of America in Washington, D.C.*.

Flemming, E. (2007). Stop place contrasts before liquids. In *Proceedings of the 16th International Congress on Phonetic Sciences, Saarbrücken* (pp. 233–236).

Flemming, E. (2013). *Auditory representations in phonology.* Routledge.

Freeman, M. R., Blumenfeld, H. K., & Marian, V. (2016). Phonotactic constraints are activated across languages in bilinguals. *Frontiers in Psychology*, *7*, 702.

Freeman, M. R., Blumenfeld, H. K., & Marian, V. (2017). Cross-linguistic phonotactic competition and cognitive control in bilinguals. *Journal of Cognitive Psychology*, *29*(7), 783–794.

Freitas, M. J. (2003). The acquisition of onset clusters in European Portuguese. *Probus*, *15*(1), 27–46.

Frisch, S. A., & Brea-Spahn, M. R. (2010). Metalinguistic judgments of phonotactics by monolinguals and bilinguals. *Laboratory Phonology*, *1*(2), 345–360.

Frisch, S. A., & Zawaydeh, B. A. (2001). The psychological reality of OCP-Place in Arabic. *Language*, 91–106.

Gaillard, S., & Tremblay, A. (2016). Linguistic proficiency assessment in second language acquisition research: The elicited imitation task. *Language Learning*, *66*(2), 419–447.

Genesee, F. (1989). Early bilingual development: One language or two? *Journal of Child Language*, *16*(1), 161–179.

Gerrits, E., & Schouten, M. E. (2004). Categorical perception depends on the discrimination task. *Perception & Psychophysics*, *66*(3), 363–376.

Gnanadesikan, A. (1995/2004). Markedness and faithfulness constraints in child phonology. *Constraints in Phonological Acquisition*, 73–108.

Goad, H. (2011). The representation of sC clusters. *The Blackwell companion to phonology*, 1–26.

Goad, H. (2012). sC clusters are (almost always) coda-initial. *The Linguistic Review*, *29*(3), 335–373.

Goad, H., & Rose, Y. (2004). Input elaboration, head faithfulness and evidence for representation in the acquisition of left-edge clusters in West Germanic. *Constraints in Phonological Acquisition*, 109–157.

Goldsmith, J. A. (1990). *Autosegmental and metrical phonology* (Vol. 1). Basil Blackwell.

Goldwater, S., & Johnson, M. (2003). Learning OT constraint rankings using a maximum entropy model. In *Proceedings of the Stockholm Workshop on Variation within Optimality Theory* (Vol. 111120).

Gonzales-Diaz, M. (2006). *Activation and inhibition. Control mechanisms in a bilingual brain. An Optimality Theory modelling* (Unpublished doctoral dissertation). MA Thesis, University Van Amsterdam.

Gouskova, M. (2001). Falling sonority onsets, loanwords, and syllable contact. *CLS*, *37*(1), 175–185.

Gouskova, M. (2004). Relational hierarchies in Optimality Theory: The case of syllable contact. *Phonology*, *21*(2), 201–250.

Green, D. M., Swets, J. A., et al. (1966). *Signal detection theory and psychophysics* (Vol. 1). Wiley New York.

Hakuta, K. (1986). *Mirror of language. The debate on bilingualism.* ERIC.

Halle, M., & Jones, L. G. (1971). The sound patterns of Russian: A linguistic and acoustical investigation (with an excursus on the contextual variants of the Russian vowels by Lawrence G. Jones). In *Description and analysis of contemporary standard Russian no. 1.* The Hague and Paris: Mouton.

Hallé, P. A. (2008). Language-specific aspects of the perception of illegal clusters. *Workshop on 'Consonant Clusters and Structural Complexity', 31*.

Hallé, P. A., & Best, C. T. (2007). Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/clusters. *The Journal of the Acoustical Society of America*, *121*(5), 2899–2914.

Hallé, P. A., & Segui, J. (2003). *Special is especial: Perception and production of the /sp, st, sk/ clusters by native speakers of Spanish.* Communication au 13ème ESCOP, Granada, Espagne.

Hallé, P. A., Segui, J., Dominguez, A., Cuetos, F., Jaichenco, V., & Sevilla, Y. (2013). Special es especial pero stuto no es astuto: La percepción de una/e/protética en el habla y la escritura de hablantes de español. *Psicolingüística en español: Homenaje a Juan Segui*, 31–47.

Hallé, P. A., Segui, J., Frauenfelder, U., & Meunier, C. (1998). Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance*, *24*(2), 592.

Haman, E., Wodniecka, Z., Marecka, M., Szewczyk, J., Białecka-Pikul, M., Otwinowska, A., ... others (2017). How does L1 and L2 exposure impact L1 performance in bilingual children? evidence from Polish-English migrants to the United Kingdom. *Frontiers in Psychology*, *8*, 1444.

Hancin-Bhatt, B. (2008). in Optimality Theory. *Phonology and Second Language Acquisition*, *36*, 117.

Hanulikova, A., Mitterer, H., & McQueen, J. M. (2011). Effects of first and second language on segmentation of non-native speech. *Bilingualism: Language and Cognition*, *14*(4), 506–521.

Harris, J. W. (1983). Syllable structure and stress in Spanish: A nonlinear analysis. *Linguistic Inquiry Monographs Cambridge, Mass.*(8), 1–158.

Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*, *177*, 263–277.

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science*, *15*(6), 409–414.

Hatzidaki, A., Branigan, H. P., & Pickering, M. J. (2011). Co-activation of syntax in bilingual language production. *Cognitive Psychology*, *62*(2), 123–150.

Hautus, M. J., Macmillan, N. A., & Creelman, C. D. (2021). *Detection theory: A user's guide*. Routledge.

Hayes, B., Siptár, P., Zuraw, K., & Londe, Z. (2009). Natural and unnatural constraints in Hungarian vowel harmony. *Language*, 822–863.

Hayes, B., & White, J. (2013). Phonological naturalness and phonotactic learning. *Linguistic Inquiry*, *44*(1), 45–75.

Hayes, B., & Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, *39*(3), 379–440.

Haznedar, B. (2010). Transfer at the syntax-pragmatics interface: Pronominal subjects in bilingual turkish. *Second Language Research*, *26*(3), 355–378.

Hsin, L. (2014). *Integrated bilingual grammatical architecture: Insights from syntactic development* (Unpublished doctoral dissertation). Johns Hopkins University.

Hulk, A., & Müller, N. (2000). Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism: Language and Cognition*, *3*(3), 227–244.

Ioup, G. (2008). Exploring the role of age in the acquisition of a second language phonology. *Phonology and Second Language Acquisition*, *36*, 41–62.

Itô, J. (1982). The syllable structure of Russian. *University of Massachusetts, Amherst MS*.

Jäger, G. (2007). Maximum entropy models and stochastic Optimality Theory. *Architectures, rules, and preferences: variations on themes by Joan W. Bresnan. Stanford: CSLI*, 467–479.

Jany, C., Gordon, M., Nash, C. M., & Takara, N. (2007). How universal is the sonority hierarchy?: A cross-linguistic acoustic study. In *Proceedings of the ICPhS* (pp. 1401–1404).

Jarosz, G. (2010). Implicational markedness and frequency in constraint-based computational models of phonological learning. *Journal of Child Language*, *37*(3), 565.

Jarosz, G., & Rysling, A. (2017). Sonority sequencing in Polish: The combined roles of prior bias & experience. In *Proceedings of the Annual Meeting on Phonology* (Vol. 4).

Joanisse, M. F., & Seidenberg, M. S. (2003). Phonology and syntax in specific language impairment: Evidence from a connectionist model. *Brain and Language*, *86*(1), 40–56.

Johnson, J. S. (1992). Critical period effects in second language acquisition: The effect of written versus auditory materials on the assessment of grammatical competence. *Language Learning*, *42*(2), 217–248.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60–99.

Johnson, J. S., & Newport, E. L. (1991). Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language. *Cognition*, *39*(3), 215–258.

Kabak, B., & Idsardi, W. J. (2007). Perceptual distortions in the adaptation of English consonant clusters: Syllable structure or consonantal contact constraints? *Language and Speech*, *50*(1), 23–52.

Kager, R. (1999). *Optimality Theory.* Cambridge University Press.

Kawasaki, H. (1982). *An acoustical basis for constraints on sound sequences* (Unpublished doctoral dissertation). University of California, Berkeley.

Kaye, J. (1992). *Do you believe in magic? the story of s+ C sequences* (Tech. Rep.). SOAS working papers in Linguistics and Phonetics.

Kaye, J., Lowenstamm, J., & Vergnaud, J.-R. (1990). Constituent structure and government in phonology. *Phonology*, *7*(1), 193–231.

Keating, P. A. (1983). Comments on the jaw and syllable structure. *Journal of Phonetics*, *11*(4), 401–406.

Keshavarz, M. H., & Ingram, D. (2002). The early phonological development of a Farsi-English bilingual child. *International Journal of Bilingualism*, *6*(3), 255–269.

Kheder, S., & Kaan, E. (2021). Cognitive control in bilinguals: Proficiency and code-switching both matter. *Cognition*, *209*, 104575.

Kim, J. Y. (2011). L1-L2 phonetic interference in the production of Spanish heritage speakers in the US. *The Korean Journal of Hispanic Studies*, *4*, 1–28.

Kristoffersen, K. E., & Simonsen, H. G. (2006). The acquisition of # sC clusters in Norwegian. *Journal of Multilingual Communication Disorders*, *4*(3), 231–241.

Kroll, J. F., & Tokowicz, N. (2005). *Models of bilingual representation and processing: Looking back and to the future.* Oxford University Press.

Kuhl, P. K., Kiritani, S., Deguchi, T., Hayashi, A., Stevens, E. B., Dugger, C. D., & Iverson, P. (1997). *Effects of language experience on speech perception: American and Japanese infants' perception of /ra/and /la* (Unpublished doctoral dissertation). Acoustical Society of America.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*(2), F13–F21.

Lavandera, B. (1978). Where does the lingusitc variable stop? *Language and Society*, *7*(2), 171-182.

Legendre, G. (2001). An introduction to Optimality Theory in syntax. *Optimality-theoretic syntax*, 1–27.

Legendre, G., Miyata, Y., & Smolensky, P. (1990/2006). Harmonic Grammar - A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations. In *Proceedings of the 12th annual conference of the Cognitive Science Society* (pp. 884–891).

Legendre, G., Sorace, A., Smolensky, P., et al. (2006). The Optimality Theory - Harmonic Grammar connection. *The Harmonic Mind*, *2*, 339–402.

Lentz, T. O., & Kager, R. W. (2015). Categorical phonotactic knowledge filters second language input, but probabilistic phonotactic knowledge can still be acquired. *Language and Speech*, *58*(3), 387–413.

Liceras, J. M. (1985). The value of clitics in non-native Spanish. *Interlanguage Studies Bulletin (Utrecht)*, *1*(2), 151–168.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*(3), 384–422.

Liu, H., & Cao, F. (2016). L1 and l2 processing in the bilingual brain: A meta-analysis of neuroimaging studies. *Brain and Language*, *159*, 60–73.

Lleó, C. (2002). The role of markedness in the acquisition of complex prosodic structures by German-Spanish bilinguals. *International Journal of Bilingualism*, *6*(3), 291–313.

Lleó, C., Kuchenbrandt, I., Kehoe, M., & Trujillo, C. (2003). Syllable final consonants in Spanish and German monolingual and bilingual acquisition. *N. Müller (ed.)*, 191–220.

MacIntyre, P. D., Noels, K. A., & Clément, R. (1997). Biases in self-ratings of second language proficiency: The role of language anxiety. *Language Learning*, *47*(2), 265–287.

Mack, M. (1989). Consonant and vowel perception and production: Early English-French bilinguals and English monolinguals. *Attention, Perception, & Psychophysics*, *46*(2), 187–200.

Mack, M. (1990). Phonetic transfer in a French-English bilingual child. *Language Attitudes and Language Conflict*, 107–124.

Mack, M., Bott, S., & Boronat, C. B. (1995). Mother, I'd rather do it myself, maybe: An analysis of voice-onset time produced by early French-English bilinguals. *Issues and Developments in English and Applied*, 28.

MacKay, I. R., & Flege, J. E. (2004). Effects of the age of second language learning on the duration of first and second language sentences: The role of suppression. *Applied Psycholinguistics*, *25*(3), 373–396.

MacKay, I. R., Flege, J. E., Piske, T., & Schirru, C. (2001). Category restructuring during second-language speech acquisition. *The Journal of the Acoustical Society of America*, *110*(1), 516–528.

MacLeod, A. A., & Stoel-Gammon, C. (2009). The use of voice onset time by early bilinguals to distinguish homorganic stops in Canadian English and Canadian French. *Applied Psycholinguistics*, *30*(1), 53–77.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*.

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism*, *6*(2), 97.

Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brain imaging. *Brain and Language*, *86*(1), 70–82.

Martin, C. D., Dering, B., Thomas, E. M., & Thierry, G. (2009). Brain potentials reveal semantic priming in both the 'active' and the 'non-attended' language of early bilinguals. *NeuroImage*, *47*(1), 326–333.

McCarthy, J. J. (1986). OCP effects: Gemination and antigemination. *Linguistic Inquiry*, 207–263.

McCarthy, J. J., & Prince, A. S. (1995). Faithfulness and reduplicative identity.

McCarthy, J. J., & Prince, A. S. (1999). Faithfulness and identity in prosodic morphology. *The prosody-morphology interface*.

McDonald, J. L. (2000). Grammaticality judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics*, *21*(3), 395–423.

McLeod, S., Van Doorn, J., & Reed, V. A. (2001). Normal acquisition of consonant clusters. *American Journal of Speech-Language Pathology*.

Meillet, A. ([1906] 1921). *Introduction à l'étude comparative des langues indo-européennes.* Hachette.

Meisel, J. M. (2010). Age of onset in successive acquisition of bilingualism: Effects on grammatical development. *Language Acquisition Across Linguistic and Cognitive Systems*, *52*(225-248).

Molnar, M., Gervain, J., & Carreiras, M. (2014). Within-rhythm class native language discrimination abilities of Basque-Spanish monolingual and bilingual infants at 3.5 months of age. *Infancy*, *19*(3), 326–337.

Montrul, S. A. (2006). On the bilingual competence of Spanish heritage speakers: Syntax, lexical-semantics and processing. *International Journal of Bilingualism*, *10*(1), 37–69.

Montrul, S. A. (2008). *Incomplete acquisition in bilingualism: Re-examining the age factor* (Vol. 39). John Benjamins Publishing.

Montrul, S. A. (2010a). Dominant language transfer in adult second language learners and heritage speakers. *Second Language Research*, *26*(3), 293–327.

Montrul, S. A. (2010b). How similar are adult second language learners and Spanish heritage speakers? Spanish clitics and word order. *Applied Psycholinguistics*, *31*(1), 167.

Montrul, S. A. (2013). Bilingualism and the heritage language speaker. *The handbook of bilingualism and multilingualism*, 168–189.

Montrul, S. A., & Bowles, M. (2009). Back to basics: Differential object marking under incomplete acquisition in Spanish heritage speakers. *Bilingualism*, *12*(3), 363–383.

Moreton, E. (2002). Structural constraints in the perception of English stop-sonorant clusters. *Cognition*, *84*(1), 55–71.

Moreton, E., Pater, J., & Pertsova, K. (2017). Phonological concept learning. *Cognitive Science*, *41*(1), 4–69.

Müller, N. (1998). Transfer in bilingual first language acquisition. *Bilingualism: Language and Cognition*, *1*(3), 151–171.

Müller, N. (2003). *(In) Vulnerable domains in multilingualism* (Vol. 1). John Benjamins Publishing.

Muysken, P. (2013). Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition*, *16*(4), 709–730.

Navarro Tomás, T. (1980). *Manual de pronunciación española [Manual of Spanish pronunciation].* Madrid: Consejo Superior de Investigaciones Científicas. (Original work).

Newport, E., & Supalla, T. (1987). A critical period effect in the acquisition of a primary language. *Unpublished manuscript*, *262*.

Ortega, L., Iwashita, N., Rabie, S., & Norris, J. (1999). A multilanguage comparison of L2 syntactic complexity measures. In *12th AILA (World Congress of International Association of Applied Linguistics) Conference, Tokyo.*

Paradis, J. (2001). Do bilingual two-year-olds have separate phonological systems? *International Journal of Bilingualism*, *5*(1), 19–38.

Paradis, J., & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition*, *18*(1), 1–25.

Parker, S. (2002). *Quantifying the sonority hierarchy* (Unpublished doctoral dissertation). University of Massachusetts at Amherst.

Parker, S. (2012). *The sonority controversy* (Vol. 18). Walter de Gruyter.

Pastore, R., & Scheirer, C. (1974). Signal detection theory: Considerations for general application. *Psychological Bulletin*, *81*(12), 945.

Pater, J., & Barlow, J. A. (2003). Constraint conflict in cluster reduction. *Journal of Child Language*, *30*(3), 487–526.

Pater, J., & Staubs, R. (2013). *Modeling learning trajectories with batch gradient descent.* Paper presented at the meeting of NECPhon, MIT.

Patkowski, M. S. (1980). The sensitive period for the acquisition of syntax in a second language. *Language Learning*, *30*(2), 449–468.

Pérez-Leroux, A. T., Cuza, A., & Thomas, D. (2011). Clitic placement in Spanish–English bilingual children. *Bilingualism: Language and Cognition*, *14*(2), 221–232.

Pérez-Leroux, A. T., Cuza, A., & Thomas, D. (2011). Clitic placement in Spanish–English bilingual children. *Bilingualism: Language and Cognition*, *14*(2), 221–232.

Pérez-Leroux, A. T., & Glass, W. R. (1999). Null anaphora in Spanish second language acquisition: Probabilistic versus generative approaches. *Second Language Research*, *15*(2), 220–249.

Pinker, S. (2009). *Language learnability and language development, with new commentary by the author: With new commentary by the author* (Vol. 7). Harvard University Press.

Pirvulescu, M., Pérez-Leroux, A.-T., Roberge, Y., Strik, N., & Thomas, D. (2014). Bilingual effects: Exploring object omission in pronominal languages. *Bilingualism: Language and Cognition*, *17*(3), 495–510.

Piske, T., MacKay, I. R., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, *29*(2), 191–215.

Pitt, M. A. (1998). Phonological processes and the perception of phonotactically illegal consonant clusters. *Perception & Psychophysics*, *60*(6), 941–951.

Prentza, A., Kaltsa, M., & Tsimpli, I. M. (2019). The role of bilingualism, age of onset of L2 acquisition & literacy in sentence repetition: The case of Albanian-Greek speaking children. *Selected Papers on Theoretical and Applied Linguistics*, *23*, 350–363.

Prince, A., & Smolensky, P. (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, MA: Blackwell Publishing.

Prince, T. (2014). On the behaviour of sC clusters: French acquisition and aphasia data. In *Talk presented at the 22nd Manchester Phonology Meeting. Manchester, England.*

R Core Team. (2020). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Rauber, A. S., et al. (2002). *The production of English initial/s/clusters by Portuguese and Spanish EFL speakers* (Unpublished doctoral dissertation). Universidade Federal de Santa Catarina.

Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7-and 11-month-old American infants. *Developmental Science*, *8*(2), 162–172.

Rodriguez-Fornells, A., Rotte, M., Heinze, H.-J., Nösselt, T., & Münte, T. F. (2002). Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature*, *415*(6875), 1026–1029.

Rodríguez-Mondoñedo, M. (2008). The acquisition of differential object marking in Spanish. *Probus*, *20*(1), 111–145.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386.

Rosenblatt, F. (1962). Principles of neurodynamics: Perceptrons and the theory of brain mechanisms. Washington, DC: Spartan.

Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing*, *15*(1), 1–20.

Sánchez, L. (2003). *Quechua-Spanish bilingualism: Interference and convergence in functional categories* (Vol. 35). John Benjamins Publishing.

Sebastián-Gallés, N., & Bosch, L. (2002). Building phonotactic knowledge in bilinguals: Role of early exposure. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(4), 974.

Sebastián-Gallés, N., & Bosch, L. (2005). *Phonology and bilingualism.* Oxford University Press.

Sebastian-Galles, N., & Kroll, J. (2003). Phonology in bilingual language processing: Acquisition, perception, and production. *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*, 279–318.

Selkirk, E. O. (1984). On the major class features and syllable theory.

Serratrice, L., Sorace, A., & Paoli, S. (2004). Crosslinguistic influence at the syntax–pragmatics interface: Subjects and objects in English–Italian bilingual and monolingual acquisition. *Bilingualism: Language and Cognition*, *7*(3), 183–205.

Shameem, N. (1998). Validating self-reported language proficiency by testing performance in an immigrant community: The wellington indo-fijians. *Language Testing*, *15*(1), 86–108.

Shooshtaryzadeh, F., et al. (2015). Optimality theory and assessment of developing and disordered phonologies. *Journal of Indian Speech Language & Hearing Association*, *29*(2), 13.

Smolensky, P. (1986). *Information processing in dynamical systems: Foundations of harmony theory* (Tech. Rep.). Colorado Univ at Boulder Dept of Computer Science.

Smolensky, P. (1995). *On the internal structure of the constraint component con of UG.*

Smolensky, P. (1996). On the comprehension/production dilemma in child language. *Linguistic Inquiry*, *27*(4), 720–731.

Smyk, E., Restrepo, M. A., Gorin, J. S., & Gray, S. (2013). Development and validation of the Spanish–English language proficiency scale (selps).

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism*, *1*(1), 1–33.

Sorace, A., & Serratrice, L. (2009). Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism*, *13*(2), 195–210.

Steriade, D. (1982). *Greek prosodies and the nature of syllabification* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Stölten, K. (2013). *The effects of age of onset on VOT in L2 aquisition and L1 attrition: A study of the speech production and perception of advanced Spanish-Swedish bilinguals* (Unpublished doctoral dissertation). Centre for Research on Bilingualism, Stockholm University.

Strand, F., Forssberg, H., Klingberg, T., & Norrelgen, F. (2008). Phonological working memory with auditory presentation of pseudo-words—an event related fMRI study. *Brain Research*, *1212*, 48–54.

Tesar, B. (1995). *Computational optimality theory* (Unpublished doctoral dissertation). University of Colorado Boulder.

Tesar, B., & Smolensky, P. (1998). Learnability in optimality theory. *Linguistic Inquiry*, *29*(2), 229–268.

Tesar, B., & Smolensky, P. (2000). *Learnability in optimality theory.* Mit Press.

Tetzloff, K. A. (2020). Exceptionality in Spanish onset clusters. *Borealis–An International Journal of Hispanic Linguistics*, *9*(1), 245–278.

Tetzloff, K. A., & Thomas, D. (in prep.). Internal and external-induced phonetic (micro)variation in Spanish heritage speakers.

Tetzloff, K. A., Whitwell, J. L., Utianski, R. L., Duffy, J. R., Clark, H. M., Machulda, M. M., … Josephs, K. A. (2018). Quantitative assessment of grammar in amyloid-negative logopenic aphasia. *Brain and Language*, *186*, 26–31.

Theodore, R. M., & Schmidt, A. M. (2003). Perceptual prothesis in native Spanish speakers. *The Journal of the Acoustical Society of America*, *113*(4), 2256–2256.

Thiery, C. (1978). True bilingualism and second-language learning. In *Language interpretation and communication* (pp. 145–153). Springer.

Thomas, D. (2012). *Grammatical optionality and variability in bilingualism: How Spanish-English bilinguals limit clitic-climbing* (Unpublished doctoral dissertation). University of Toronto.

Thompson, I. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants. *Language Learning*, *41*(2), 177–204.

Thornburgh, D. F., & Ryalls, J. H. (1998). Voice onset time in Spanish-English bilinguals: Early versus late learners of English. *Journal of Communication Disorders*, *31*(3), 215–229.

Ticio, E. (2015). Differential object marking in Spanish-English early bilinguals. *Linguistic Approaches to Bilingualism*, *5*(1), 62–90.

Valdés, G. (2000). The teaching of heritage languages: An introduction for Slavic-teaching professionals. *The learning and teaching of Slavic languages and cultures*, 375–403.

Van der Hulst, H. (1984). *Syllable structure and stress in Dutch*. De Gruyter.

Vitevitch, M. S., Luce, P. A., Charles-Luce, J., & Kemmerer, D. (1997). Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech*, *40*(1), 47–62.

Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, 349–355.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*(1), 49–63.

Wexler, K., & Manzini, M. R. (1987). Parameters and learnability in binding theory. In *Parameter setting* (pp. 41–76). Springer.

White, J., & Chiu, F. (2017). Disentangling phonological well-formedness and attestedness: An ERP study of onset clusters in English. *Acta Linguistica Academica*, *64*(4), 513–537.

Wu, S.-L., & Ortega, L. (2013). Measuring global oral proficiency in SLA research: A new elicited imitation test of L2 Chinese. *Foreign Language Annals*, *46*(4), 680–704.

Yavaş, M. (2002). Voice onset time patterns in bilingual phonological development. *Investigations in Clinical Phonetics and Linguistics*, 341–350.

Yavaş, M. (2011). Patterns of cluster reduction in the acquisition of # sC onsets: Are bilinguals different from monolinguals? *Clinical Linguistics & Phonetics*, *25*(11-12), 981–988.

Yavaş, M., Ben-David, A., Gerrits, E., Kristoffersen, K. E., & Simonsen, H. G. (2008). Sonority and cross-linguistic acquisition of initial s-clusters. *Clinical Linguistics & Phonetics*, *22*(6), 421–441.

Yavaş, M., & Core, C. W. (2006). Acquisition of # sC clusters in English speaking children. *Journal of Multilingual Communication Disorders*, *4*(3), 169–181.

Yavaş, M., & Someillan, M. (2005). Patterns of acquisition of /s/-clusters in Spanish-English bilinguals. *Journal of Multilingual Communication Disorders*, *3*(1), 50–55.

y Cabo, D. P. (2015). Issues in Spanish heritage morphosyntax. *Studies in Hispanic and Lusophone Linguistics*, *8*(2), 389–401.

Yeni-Komshian, G. H., Flege, J. E., & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean–English bilinguals. *Bilingualism: Language and Cognition*, *3*(2), 131–149.

Yip, V., & Matthews, S. (2000). Syntactic transfer in a Cantonese–English bilingual child. *Bilingualism: Language and Cognition*, *3*(3), 193–208.

Yip, V., & Matthews, S. (2005). Dual input and learnability: Null objects in Cantonese-English bilingual children. In *Isb4: Proceedings of the 4th International Symposium on Bilingualism.*

Zher, J., & Schwarz, F. (2018). *PennController for Internet Based Experiments (IBEX).* https://doi.org/10.17605/OSF.IO/MD832.

Ziegler, J. C., Muneaux, M., & Grainger, J. (2003). Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory and Language*, *48*(4), 779–793.