

Copyright  
by  
Javier Jasso  
2022

**The Dissertation Committee for Javier Jasso certifies that this is the approved  
version of the following Dissertation:**

**How Does Parent Input Influence Bilingual Children's Knowledge and  
Use of Spanish Subjunctive? A Dyadic Study**

**Committee:**

Craig Champlin, Supervisor

Almeida Jacqueline Toribio

Mirza Lugo-Neris

Belem López

Lisa Bedore

Elizabeth Peña

**How Does Parent Input Influence Bilingual Children's Knowledge and  
Use of Spanish Subjunctive? A Dyadic Study**

**by**

**Javier Jasso**

**Dissertation**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**Doctor of Philosophy**

**The University of Texas at Austin**

**August 2022**

## **Acknowledgements**

This dissertation study could not have happened without the support of many. To my committee members, for supporting my work and helping me develop my ideas further. I would like to express my gratitude to my HABLA Lab mentors Lisa Bedore and Liz Peña, for encouraging my varied research interests; to my dissertation supervisor, Craig Champlin, for your words of wisdom and calm pragmatism; to Mirza Lugo-Neris and Belem Lopez, for your continuous support and much-needed empujones; and to Jacqueline Toribio, for your expertise and kindness in helping me navigate academia.

To my community of researchers/collaborators and dear friends in HABLA Lab and the #100Days writing community. For their invaluable help offering feedback at various stages of this project and helping share my recruitment flyer. A special thank you to Amy Pratt, Cecy Pérez, Katie Winters, Katy Durant, Monica Hudnall, and Sonny Nguyen. Thank you for lifting me up and supporting me throughout the doctoral process and beyond!

I would also like to thank my family: my mom Blanca (mami), sister Karla, and brother Jonathan for your support in every way imaginable. Mami, gracias por creer en mí y por tener la certeza (y la razón) de que todo va a salir bien.

This dissertation was funded by National Science Foundation through a Doctoral Dissertation Research Improvement Grant and additional funding through a CDIP Mini-Grant from California State University, which has allowed me funds to compensate families and work with two research assistants.

Lastly, a sincere thank you to all the community leaders and partners at St. Patrick's church for welcoming me into their spaces. And most of all, to the families who

took time out of their busy lives to participate and help me learn about their language practices. Thank you all!

## **Abstract**

### **How Does Parent Input Influence Bilingual Children's Knowledge and Use of Spanish Subjunctive? A Dyadic Study**

Javier Jasso, PhD

The University of Texas at Austin, 2022

Supervisor: Craig Champlin

Previous studies in bilingual research have used adult heritage bilinguals' performance to make strong developmental claims about the acquisition process. While bilingual children's grammatical systems are not structurally or functionally identical to other language users such as their parents, it is unclear where in the process these changes occur. Child data that can be situated in its specific language learning context and is appropriately modeled can contribute to this area of inquiry by examining both linguistic and experiential factors that contribute to observed linguistic performance. This dissertation study examined bilingual dyads' acceptability ratings and imitated productions of Spanish mood selection, where a verb like *lavar* (wash) takes on indicative (*lava*) and/or subjunctive forms (*lave*) depending on the clause. We tested difference in mood selection performance across dyad members, and whether this depended on mood variability. Next, we investigated the effects of age and same-language experience on structural knowledge. Finally, we tested a possible 'partner effect' of parent comprehension on their child's grammatical knowledge.

A geographically and ethno-linguistically diverse sample of bilingual parent–child dyads ( $n = 12$ ) participated in two web-based tasks—a spoken acceptability judgment task and sentence repetition task—as well as a language history questionnaire to document dyad members’ language experiences. Both experimental tasks contained obligatory and variable subjunctive contexts using the same set of mid-frequency verbs. Mood acceptability ratings and sentence repetition accuracy scores were the primary outcomes. Research questions were tested using multiple regression, linear mixed-effects models, and an actor-partner interdependence model (APIM).

At the dyad member level, parents more accurately rejected unacceptable sentences and outperformed their children’s subjunctive productions. Results showed some support for both dyad member and mood variability effects. Contrary to our prediction, the effects of age and current Spanish exposure were not observed. Implications are discussed in relation to the future of web-based studies, particularly for recruiting and retaining participants from racialized and/or language minoritized groups, as well as considerations for dyadic studies examining bilingual language acquisition.

## Table of Contents

List of Tables .....	11
List of Figures .....	12
Chapter 1: Introduction .....	13
1.1 Empirical Findings and Limitations in Studies on Heritage Grammars .....	14
1.2 Bringing Parent–Child Designs and Dyadic Analyses to the Study of Heritage Grammars .....	18
1.3 Increasing Diversity in Language Sciences .....	20
1.4 A Brief Overview of Spanish Subjunctive .....	21
1.5 Task Considerations for U.S. Heritage Speakers .....	23
1.6 Present Study and Specific Aims .....	25
Chapter 2: Method .....	27
2.1 Participants.....	27
Selection Criteria .....	27
Sample Size Calculation .....	29
2.2 Materials .....	29
Acceptability Judgment Task .....	30
Instructions and Familiarization Procedures .....	30
Sentence Stimuli .....	31
Target Verb Characteristics .....	32
Mood Clauses .....	33
AJT Pilot and Feasibility .....	34
Sentence Repetition .....	34



Instructions and Familiarization Procedures -----	35
Demographics and Language History Questionnaire-----	36
2.3 Procedures -----	38
AJT Acceptable/Unacceptable Ratings-----	40
Sentence Repetition Accuracy -----	41
2.4 Planned Analyses -----	41
Chapter 3: Results .....	44
3.1 Data Output and Data Cleaning -----	44
3.2 Descriptive Statistics -----	45
Acceptable/Unacceptable Ratings-----	45
Filler Trials -----	45
Grammatical/Ungrammatical Trials-----	45
Sentence Repetition Subjunctive Accuracy -----	45
3.3 Dyad Member and Mood Variability Effects -----	48
3.4 Child-Level Predictors of Mood Selection -----	49
3.5 Partner Effects -----	49
Chapter 4: Discussion .....	50
4.1 Group-Level Differences -----	50
4.2 Mood Variability Performance Across Groups -----	51
4.3 Age and Spanish Use on Child Performance-----	51
4.4 Considerations for Diversity Research -----	52
4.5 Limitations -----	53
4.6 Conclusion -----	54

References .....	55
------------------	----

## **List of Tables**

Table 1:	Dyad demographic information ( $n = 12$ ). -----	29
Table 2:	Sentence repetition items with subjunctive targets. -----	36

## List of Figures

Figure 1:	Participant screen during ATJ experiment, displaying rating scale and progress bar.-----	31
Figure 2:	Selection of Q-BEx modules used.-----	38
Figure 3:	Experiment workflow across platforms for child and parent participants.-----	39
Figure 4:	SEM model testing for actor (yellow arrows) and partner effects (blue arrows) -----	43
Figure 5:	Dyad members' mean grammatical ratings by mood variability. -----	46
Figure 6:	Dyad members' mean ungrammatical ratings by mood variability. -----	47
Figure 7:	Dyad members' mean sentence repetition accuracy by mood variability. -----	48

## **Chapter 1: Introduction**

Most speakers across the globe acquire and regularly use multiple languages throughout their lifetime. In the United States, many acquire a home language in early development alongside the societally dominant language. Of these home or heritage languages, Spanish is the most common, being acquired by two thirds of all early bilingual children in the U.S., or approximately 16.1 million children (Child Trends, 2019). These heritage speakers (HS; Kupisch & Rothman, 2018; Valdés, 2000, 2005; Zyzik, 2016) typically acquire Spanish grammatical structures to varying degrees, and there has been recent interest in developing more informative theoretical frameworks, as well as methodological and analytical procedures, to study how heritage grammars are acquired in childhood and beyond (López-Beltrán & Carlson, 2020), as this has important implications for this large group of speaker and for, more generally, for language theories, which have long centered monolingual English speakers.

An important yet unresolved question in the study of heritage languages is the nature of the child learner's input—the specific structures they hear in their daily interactions with their multiple communication partners—and how and to what degree this input differs from the child's own internalized grammatical system (Polinsky & Scontras, 2020; Rothman, 2008). A large body of work on heritage Spanish has focused on speakers' attainment of specific theoretically interesting linguistic structures (e.g., mood selection) and has convincingly demonstrated that by adulthood, HS have comparable and divergent patterning to monolingual speakers living in a Spanish-monolingual society and to second language (L2) learners of Spanish living in the U.S. (Polinsky & Scontras, 2020). However, these group effects are not entirely surprising given the considerable differences between these three groups of speakers, who not only

differ in the amount of input, but also in the type of input they receive and can abstract meaning from; the specific variety/registers they acquire; and whether acquisition has been via naturalistic or instructed modes (López-Beltrán & Carlson, 2020; Montrul et al., 2008). In addition, even within the HS group, studies have reported high variability in performance (Polinsky & Scontras, 2020), rendering group mean comparisons potentially misleading and less informative if relevant participant characteristics (e.g., age group, age of acquisition, and use of the heritage language [HL]) are not accounted for or (at minimum) systematically documented. A greater understanding of heritage grammars requires accounting for variability in performance.

One way to advance our understanding of heritage grammars is by examining direct HL transmission through parent–child pairs. This solution avoids the need to rely on imperfect reference groups to draw conclusions that will inform linguistic theory. Parent–child approaches connect previous empirical work while moving beyond comparative reference frameworks. This dissertation study aims to address a methodological limitation in the study of heritage languages by 1). directly examining parent–child dyad performance, 2). testing participant-level predictors on grammatical competence, and 3). applying an innovative statistical technique that can be used to answer other research questions in HL research. The following sections will review current empirical findings in heritage grammars; argue for the use of parent–child data and offer an appropriate statistical procedure; consider sampling and tasks appropriate for this group of speakers; and provide a brief sketch of Spanish mood acquisition.

## **1.1 EMPIRICAL FINDINGS AND LIMITATIONS IN STUDIES ON HERITAGE GRAMMARS**

The current body of research on heritage grammars has provided insights about speakers' performance and explored several conditioning factors that likely influence

outcomes (e.g., Lynch, 1999; Polinsky & Scontras, 2020; Silva-Corvalán, 1994). These studies vary in their theoretical orientation and methodological approaches; these have different underlying assumptions that inform the types of questions asked and what counts as evidence of knowledge. The main findings from these traditions, as well as some limitations, are outlined below.

Formal generative frameworks have been the most common since the subfield of heritage grammars gained traction. Within these frameworks, competing theories are empirically tested, often using experimental paradigms such as judgment/preference tasks, elicited production tasks, or a combination of tasks (Ionin & Zyzik, 2014; López-Beltrán & Carlson, 2020). HS performance is compared to an experimental control or reference group such as monolingual Spanish speakers, L2 learners, and, more recently, Spanish-dominant speakers (e.g., Giancaspro et al., 2022). In terms of the participants sampled, these studies are often conducted with adult HS in university settings (e.g., Montrul 2007, 2009; Perez-Cortes, 2020). Many of these studies employ written tasks, presumably for historical reasons, as these were adapted from the field of SLA, and their ease of design and scoring responses. Results from these studies show HS–control differences across a range of structures, and these findings have been taken as evidence of specific developmental trajectories in the HS, including incomplete acquisition (Montrul, 2008), language attrition (Polinsky & Scontras, 2020), or protracted development (Cuza & Miller, 2015). These characterizations are not unproblematic and have been called into question by various researchers (Otheguy, 2013; Putnam & Sánchez, 2013; Rothman, 2008).

Comparatively fewer studies have been published from a usage-based perspective, although this framework has been used to examine specific constructions like Spanish mood selection or direct object clitics (Requena, 2022). The main empirical tools used

include spontaneous samples, cross-sectional or longitudinal studies, and monolingual corpus work. An important contribution of usage-based studies is the recognition that lexical forces can play an arguably more important role in language use (e.g., mood selection) than can be explained by either purely semantic or syntactic factors, as well as findings of gradience across monolingual and bilingual grammars (López-Beltrán & Carlson, 2020).

Sociolinguistic theories model the variation in HL speakers (Lynch, 1999; Otheguy & Zentella, 2012; Viner, 2018), usually through cross-sectional apparent time (vs. real time) studies on the production data of multiple generations of a speech community (Otheguy & Zentella, 2012), using adult speaker data that includes gender, age, and country of origin. These studies have examined language contact (e.g., with English) and dialectal contact phenomena (e.g., multiple varieties in a single geographic area, Aaron & Hernández, 2007; Klee & Lynch, 2009). An overall finding is that HL grammars show quantifiable changes when compared to more recently arrived immigrants (Klee & Lynch, 2009; Otheguy & Zentella, 2012). At the same time, much of this work shows that HS continue to show sensitivity to linguistic constraints (see Silva-Corvalán & Enrique-Arias, 2017, for a review of morphosyntactic/sociosyntactic studies in Spanish)

This review provides the reader an understanding of the current state of knowledge pertaining to HL grammars across these various empirical traditions. However, there are at least two major shortcomings of this work: one, the comparative framework and two, the restricted sampling procedures used. Comparative methods, where adult HSs' performance is compared to a control group, draw conclusions about developmental trajectories by examining the end-state adult grammar. Using this logic, divergent performance from a “native-speaker” baseline is often taken as evidence of



“incomplete acquisition” of specific forms or features; however, incomplete acquisition presupposes the availability of a target structure in the input, which is an assumption that is not typically empirically tested. Describing HL grammars as “incomplete” and “attrition” seems to be an ideological position rather than an empirical reality. At present, there is a paucity of studies that examine child HS performance and children’s real-world input sources (but see Cuza et al., 2013, for multiple child/adult comparisons). Without direct evidence from parent–child dyads, it remains an open question which aspects of children’s structural competence are best understood to be present in the input, yet not acquired and which, alternatively, can be traced back to the input itself (i.e., missing-input competence divergence; Flores et al., 2017; Pires & Rothman, 2009). For example, a study of Portuguese HS showed that the lack of exposure to a form only available in more formal registers is not acquired (Pires & Rothman, 2009).

Participant characteristics (e.g., language exposure/use, number of interlocutors) are a likely source of observed cross-study differences that have received less attention. As noted earlier, HL studies have routinely used convenience samples of university students (Correa, 2011; Hurtado & Montrul, 2020; Lynch, 1999; Montrul 2007, 2009; Perez-Cortes, 2020). In addition, many samples are not carefully described in terms of bilingual characteristics that can be expected to contribute to outcomes, an issue that has been raised in bilingual research more broadly (Surrain & Luk, 2019). Both of these research practices pose threats to the generalizability of findings to other bilingual speakers. Thus, depending on the research questions, broader sampling of bilingual speakers and/or increased documentation of bilingual characteristics is warranted (see Surrain & Luk, 2019; López, 2020). One methodological/analytic aspect that can help to more directly connect input sources to HS performance in diverse samples is the study of parent–child pairs or dyads.

## **1.2 BRINGING PARENT–CHILD DESIGNS AND DYADIC ANALYSES TO THE STUDY OF HERITAGE GRAMMARS**

To better isolate the relationship between input and observed language use, we can leverage the fact that parents typically form the basis for children’s input. Therefore, collecting linguistic and demographic data from parent–child pairs (*dyads*) offers researchers the possibility to directly examine how input and other experiential factors shape HL grammars without the comparability issues outlined above (e.g., Kupisch & Rothman, 2018; Otheguy, 2013; Pascual y Cabo & Rothman, 2012; Rothman, 2008). That is, in contrast to previous cross-sectional studies of heritage grammars, child learners are compared to their most likely input source: a parent who speaks the HL and belongs to the same linguistic community (Pham & Tipton, 2018; Willard et al., 2015). Studying parent–child dyads may additionally help tease apart effects observed in intergenerational studies (done in apparent time) that examine different generations from the same geographical area (Lynch, 1999; Silva-Corvalan, 1994). Parent effects on child outcomes are routinely investigated in developmental psychology or language acquisition (e.g., Bedore et al., 2012; Gathercole et al., 1999). In HL research, leveraging parent–child data in this way opens the possibility to new research endeavors, allowing researchers to ask different types of questions about the nature of heritage grammars: for example, what structural knowledge in parent grammars is transmitted to the child during the language learning process, and what are some specific conditioning factors? These designs can employ regression models or similar procedures to compare parent–child grammars and examine the relative contributions of participant characteristics on performance (Baayen et al., 2008).

Among these predictors that have been shown to account for HL performance in school age are chronological age, input quantity, and proficiency. For example, children’s

current exposure has been found to predict morphosyntactic and lexical–semantic performance in school age (Bedore et al., 2012; Flores et al., 2017), although these relationships are less clear in middle childhood and adolescence (Soto-Corominas et al., 2020).

In addition to dyadic data using parent-level predictors on child outcomes, dyadic data has been used to investigate both members’ levels on the same predictors and outcomes (Kenny et al., 2006). While parent–child studies with one-sided measures (e.g., parent satisfaction, child attachment within a dyad) can employ regression models, reciprocal measures that investigate both members’ knowledge of specific linguistic forms are better modeled via a dyadic data analysis that recognizes the nested data structure being studied. This group of statistical procedures is common in other social sciences studying relationships and other mutual-influence phenomena (e.g., Wittenborn et al., 2013), but have been less adopted in studies on language acquisition or online language processing. In these designs, dyadic data is a type of clustered data, with the dyad as the primary level of analysis, and the procedure models the non-independence of dyads (i.e., due to linked scores) and is more appropriate for the data structure in question. One such type of dyadic analysis and the one that will be used in this dissertation study is the actor-partner interdependence model (APIM), which directly tests the effects of parents on child, as well as child on parent (Cook & Kenny, 2005; Wittenborn et al., 2013). A parent–child study includes distinguishable dyads (1 parent, 1 child in each dyad) and can investigate such mixed variables (i.e., variables that vary within the dyad and between dyads) as age and usage patterns. Any given mixed variable can be empirically tested for possible parent and child *actor effects* (i.e., levels on self), as well as parent-on-child and child-on-parent *partner effects* (i.e., levels on others). Of particular interest for this study is the parent-on-child partner effect, which tests for the

influence of the parents' comprehension scores on their children's production scores, and the direction and magnitude of this relationship (Fitzpatrick et al., 2016).

### **1.3 INCREASING DIVERSITY IN LANGUAGE SCIENCES**

In multilingual populations, bilingualism is a life experience best understood under its specific social and individual conditions (Tiv et al., 2020). However, language acquisition and use across the lifespan has historically been theorized with an implicit idealized monolingual speaker in mind, with other communication practices typically becoming exceptionalized, othered, and/or otherwise framed as deficient. Said another way, much of what is currently known about language acquisition and language processing is based on a restricted, rather homogenous sample of language users. While the existing literature has certainly advanced knowledge in this area, there is a critical need for empirical data from non-monolingual speakers to inform theories that reflect actual (vs. imagined) language use in communities.

In consideration of the importance to move away from homogenous samples, for the present study, efforts were made to ensure appropriate recruitment and task design. A broad recruitment strategy was done intentionally to enhance generalizability, compared to convenience samples of university students in heritage instruction programs, common in previous studies (Correa, 2011; Hurtado & Montrul, 2020; Lynch, 1999; Montrul 2007, 2009; Perez-Cortes, 2020). Web-based data collection can offer advantages toward this end, yielding similar performance to in-person methods while allowing for more diverse sampling (Casler et al., 2013), although some technological literacy is assumed. Prior work has also suggested that shorter single-session studies with compensation may yield higher quality compared to those that are time-consuming or that occur over multiple sessions (Maniaci & Rogge, 2014). Increased participation in research by

marginalized communities, who may be distrustful of institutions that seek to study them, requires trust building and recognizing and reducing barriers affecting their participation (George et al., 2014).

#### **1.4 A BRIEF OVERVIEW OF SPANISH SUBJUNCTIVE**

Spanish subjunctive mood is used in this study to examine language comprehension and production in bilingual parent–child dyads. Subjunctive is a well-studied grammatical marker that occurs in obligatory and variable contexts, which makes it an ideal testing site for examining the effect of input variability on language acquisition (Otheguy & Zentella, 2012).

Mood selection, the use of subjunctive mood or the more common indicative mood, has been investigated in U.S.-residing adult HSs (Mikulski, 2010; Montrul 2007, 2009; Perez-Cortes, 2020; van Osch et al., 2017), cross-generational studies (Lynch, 1999; Viner, 2018), and developmental studies of monolingual and bilingual children (Blake, 1983; Castilla-Earls et al., 2018; Dracos & Requena, 2022; Flores et al., 2017; Pérez-Leroux, 1998; Silva-Corvalán, 1994, 2014). In terms of its diachronic use, subjunctive is present across most spoken varieties<sup>1</sup> of Spanish (Lipski, 2008; Lynch, 1999).

As can be seen the example below, indicative–subjunctive mood selection in non-matrix clauses is structurally transparent, with overt marking suffixed to the verb root for both indicative (1) and subjunctive (2) forms:

---

<sup>1</sup> Some exceptions include preferences for infinitival constructions (*antes de yo ir* [before I go]) rather than the equivalent subjunctive construction (*antes de que (yo) vaya*), observed in many Caribbean Spanish varieties (Lipski, 2008, pp. 112, 125, 137; Suñer, 1986), as well as the increasing use of indicative in historically subjunctive contexts in Argentine Spanish (Gallego & Alonso-Marks, 2014; Lastra & Butragueño, 2012).

(1). Elena sab-e                      que siempre **desayun-a**                      a las 7.

Elena know-3SG.INDIC that always eat.breakfast-3SG.INDIC at 7

‘Elena knows (that) they (sg) always have breakfast at 7.’

(2). Elena quier-e                      que **desayun-e**                      con los demás.

Elena want-3SG.INDIC that eat.breakfast-3SG.SVJV with the rest

‘Elena wants them (sg) to have breakfast with the group.’

However, the extent to which this transparent alternation is perceptually salient to the child learner is less clear, particularly in light of findings of centralization or raising of unstressed vowels (e.g., Brown et al., 2021; Colantoni et al., 2020; Ronquest, 2013). In addition, mood selection interacts with semantic and pragmatic information at the sentence and discourse level (Faulkner, 2021; Bosque, 2012, for an overview). Select constructions require integration of temporal information or specificity of antecedents.

Mood selection can be said to consist of obligatory or categorical constructions on the one hand (*para que \*llamaa/llame, porque llamaa/\*llame*) that disallow one mood form and, on the other, variable structures that allow for both forms (*cuando llamaa/llame, dice que llamaa/llame*), with accompanying semantic or pragmatic distinctions (van Osch et al., 2017). However, these distinctions are not uncontroversial; in fact, there is no agreed-upon consensus on which specific contexts are obligatory vs. variable (López-Beltrán, 2021).

When considering acquisitional patterns, developmental data suggest that subjunctive is acquired asymmetrically, on a by-construction basis, over a relatively wide developmental period (Blake, 1983; Flores et al., 2017). Complete acquisition in production may not emerge until approximately age 8, and are especially difficult to acquire in children with language disorders (Castilla-Earls et al., 2018). Corpus data

collected from Blake (1980) suggest target performance across all clause types by age 10. In generative traditions, mood selection exists both in the narrow syntax and at the interface with other linguistic and cognitive domains (van Osch et al., 2017), and this fact is reflected in the multiple stages in which it is naturally acquired: children first begin producing a few restricted subjunctive constructions (e.g., purposive *para que* [so] and adverbial future *cuando* [when]; Montrul, 2009); in contrast, pragmatically conditioned constructions (e.g., relative clauses of non-specific antecedents) are not fully acquired until after age 8 (Dracos et al., 2019; Pérez-Leroux, 1998). These later-acquired structures that interface with other linguistic/cognitive domains are largely the same contexts that appear most divergent in adult HS when tested experimentally, although this effect seems to be conditioned by lexical frequency (e.g., of the matrix verb) and morphological regularity (Giancaspro, 2017, 2019; Perez-Cortes, 2020).

### **1.5 TASK CONSIDERATIONS FOR U.S. HERITAGE SPEAKERS**

Bilingual speakers receive input that is qualitatively different from that of both L2 learners and monolingually raised children (López-Beltrán & Carlson, 2020; Polinsky & Scontras, 2020; Rothman, 2008), acquiring a vernacular register of at least one Spanish variety in a naturalistic setting, while hearing fewer forms that are more frequent in formal registers (López-Beltrán & Carlson, 2020). Therefore, oral (vs. written) tasks designed to tap speaker-internal representations (vs. metalinguistic knowledge) are likely more adequate measures of speaker knowledge (Jegerski, 2018; van Osch et al., 2018), without penalizing speakers for a modality they may not regularly use, which would underestimate their linguistic knowledge.

Given the above considerations, two language tasks were designed—one comprehension, one production task—to capture both groups' performance while being

feasible for a web-based<sup>2</sup> study. For comprehension, an acceptability judgment task<sup>3</sup> (AJT) was chosen to measure dyad members' grammatical knowledge without requiring meta-linguistic awareness, which has been shown to underestimate HS knowledge (Jegerski, 2018; Sedarous & Namboodiripad, 2020; van Osch et al., 2018). This AJT used spoken sentences and a timed element, which minimize literacy effects and correlate with implicit knowledge (Ellis, 2005, discussed in Ionin & Zyzik, 2014). AJTs with a graded rating scale are appropriate for both children and adults (Ambridge et al., 2008; Blything et al., 2014), as they do not generally show extreme floor or ceiling effects (but see Montrul, 2009, for an exception with Spanish-dominant adults).

For production, a sentence repetition task was used to elicit subjunctive production, similar to what has been previously done in monolingual and bilingual children (Komeili & Marshall, 2013; Marinis & Armon-Lotem, 2015). Language imitation tasks like a sentence repetition task can appropriately measure production of specific forms and have been used extensively to elicit morphosyntactic forms across studies. Advantages of this experimental paradigm include greater experimental control of contextual/pragmatic factors and testing of specific structures that are otherwise infrequent in spontaneous samples or structured narratives (Shin, 2018). Finally, considering the appropriateness of the task for this web based-study, sentence repetition tasks are relatively quick and easy to administer, without much training needed for the participant. Sentence repetition data on U.S.-residing school-age bilingual children similar to the ages studied here have been used to provide information about general morphosyntactic performance (Jasso et al., in prep.; Peña, 2010–2014).

---

<sup>2</sup> I use the term *web-based* to avoid ambiguity with real-time language processing

<sup>3</sup> For an extended discussion of AJT tasks across studies, see Spinner & Gass (2019)



In terms of language experience, many paper and pencil language history questionnaires have been used to document bilingual participants' linguistic profile and relevant sociolinguistic variables (see Kaščelan et al., 2022). Some current examples for both bilingual adults and children include Language History Questionnaire (LHQ3; Li et al., 2020), Bilingual Language Profile (BLP; Birdsong et al., 2012), Bilingual Input Output Survey (BIOS; Peña et al., 2018), and Questionnaire for Parents of Bilingual children (PaBiQ; Tuller, 2015). A more recent survey that is an extension of the PaBiQ and designed specifically to be completed online by parent/child respondents, Quantifying Bilingual EXperience survey (Q-BEx; De Cat et al., 2022), was used for this study.

## **1.6 PRESENT STUDY AND SPECIFIC AIMS**

Prior work has established generational effects in knowledge and use of subjunctive forms in adult speakers across various linguistic communities in the U.S. At the same time, relatively few studies have examined subjunctive performance in children. Examining both parent and child performance can help contribute to the current knowledge and link to bodies of literature. This dissertation study will examine obligatory/variable mood contexts in parent-child dyads in both comprehension (judgment tasks) and production (sentence repetition) to explore the role of specific linguistic and social factors. Specifically, I asked the following questions:

1. Do parent and child groups differ in their mood selection (indicative/subjunctive) performance?
2. Does mood variability (obligatory/variable) predict performance in parent and child groups?
3. Do age and Spanish experience explain performance in child mood selection?

4. Is there a parent-to-child partner effect in mood selection performance?

Hypothesis 1: Based on previous work showing generational effects (e.g., Silva-Corvalan, 1994), parents were predicted to show greater knowledge and use of historical subjunctive than their children. This would be seen through higher grammatical ratings, lower ungrammatical ratings, and higher subjunctive accuracy in the parent group.

Hypothesis 2: An effect of mood variability favoring obligatory contexts was predicted for both groups (Erker, 2018; Otheguy & Zentella, 2012). This would be observed through higher grammatical ratings, lower ungrammatical ratings, and higher subjunctive accuracy in obligatory contexts.

Hypothesis 3: We hypothesized that children's current Spanish use and age would positively predict their acceptability ratings and subjunctive targets (Bedore et al., 2012; Castilla-Earls et al., 2018).

Hypothesis 4: Regarding the effects of structure-specific language input on children's subjunctive use, we predicted that parents' acceptability rating would positively and strongly contribute to children's production rates (Gathercole et al., 1999).

## **Chapter 2: Method**

### **2.1 PARTICIPANTS**

Twelve Spanish-speaking parent–child dyads participated in this study from an initial pool of 136 IRB-consenting dyads. Dyads consisted of a Spanish-speaking bilingual child living in the U.S. and one co-residing Spanish-speaking parent. This study was advertised online and in-person, in metropolitan and suburban areas in Philadelphia, Chicago, and Houston. Recruitment flyers in English and Spanish containing the study link were disseminated using professional listservs and mailing lists (COGSCI, American Speech-Language-Hearing Association [ASHA] Special Interest Groups), social media posts (Twitter, Facebook, Instagram), via specialist websites, and emails to professional contacts with IRB-approved language. Physical flyers were posted in community centers (e.g., local grocery stores, panaderías, churches, public libraries, laundromats, community clinics, and schools) to increase a more diverse participant sample (Sugden & Moulson, 2015). Finally, in-person testing opportunities were offered through a partnership with a volunteer health clinic at a local church serving several Spanish-speaking immigrant communities. This study was approved by the University of Texas at Austin and Widener University IRB Offices. Families received compensation (i.e., electronic gift card) for study participation. The majority of families who consented to participating via Qualtrics (91.17%) did not complete both studies. Families with incomplete data were contacted and given direct links that allowed them to resume the study; one family contacted the researcher on their own. Of those contacted, three resulted in complete data.

#### **Selection Criteria**

Children were second- or third-generation HS and met several inclusion criteria: 8–13 years of age and raised in a Spanish-speaking home in the continental United States,

excluding Puerto Rico, where Spanish is a societally dominant language. This study examined linguistic performance in older elementary and middle-school children, who are generally expected to have naturally acquired most grammatical structures in their languages—including mood selection across various clause types—but who have likely undergone a significant change in language exposure (Castilla-Earls et al., 2019). An additional consideration was parental concern of their child’s overall language abilities. To minimize including children with suspected language disorder, any child with expressed parental concerns based on survey responses was excluded from the analysis (Castilla-Earls et al., 2018; Tomblin et al., 1997; Norbury et al., 2016).

Participating parents were Spanish speakers raised in a Spanish-speaking Latin American country (92.31%) and the continental U.S. (7.69%). Information about their self-rated Spanish proficiency, country of origin, Spanish variety (e.g., Mexican Spanish, Colombian Spanish), and maternal education was collected. No restrictions were placed on the aforementioned variables to ensure a broad representation of speakers of U.S. Spanish was sampled. Table 1 shows the demographics of the 12 participating dyads with sufficient data for the analyses.

Table 1: Dyad demographic information ( $n = 12$ ).

Child age (years)	10.46 (1.71)
Sex (% female)	54.45%
Spanish use	.25 (.25)
Highest caregiver education	3.53 (0.78)
Spanish dialects	Mexican (57.69%), Central American (7.69%), South American (42.31%)
Parent age (years)	43.64 (8.05)
Spanish proficiency	4.92 (0.27)
English proficiency	3.50 (0.76)

### Sample Size Calculation

A power analysis was conducted to determine a minimum number of participants needed to detect an effect using simulations within the *simr* package in R (Green & MacLeod, 2016), which is especially important given the documented number of underpowered studies in psychological research (Brysbaert & Stevens, 2018). Using estimated effect sizes from the literature and a desired power level of .80, a minimum sample size of 90 dyads was recommended.

## 2.2 MATERIALS

After reading the study information and agreeing to participate in the study, parent and child groups completed two web-based experimental measures: an acceptability judgment task and a sentence repetition task, which were counterbalanced.

The total completion time was approximately 10 minutes. Following these experimental measures, parents also completed a language history questionnaire lasting approximately 30 minutes. These procedures are described in more detail in the following section.

### **Acceptability Judgment Task**

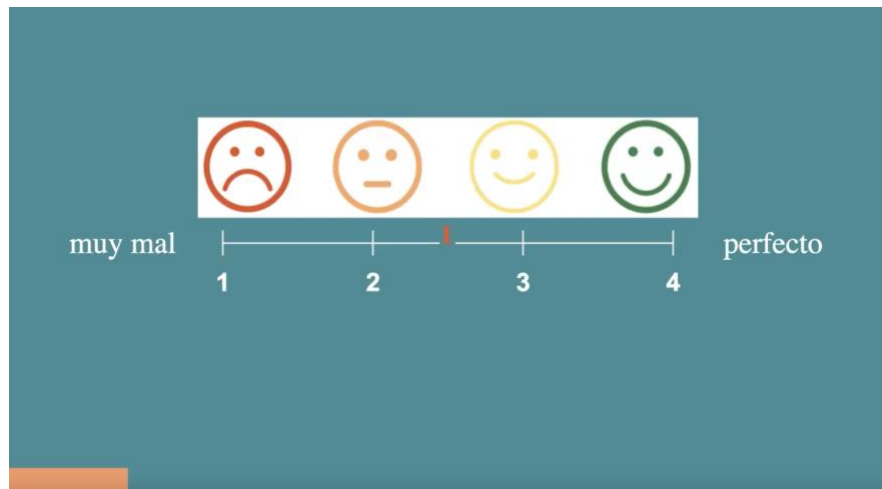
A researcher-designed timed acceptability judgment task (AJT) captured participants' off-line ratings and associated reaction times of mood selection across various clause types. Task instructions, practice trials, and experimental stimuli were audio-recorded by the researcher, a heritage speaker who uses a Mexican Spanish variety. Items were delivered pseudo-randomly across 6 possible versions to ensure that participants would not hear two same-condition items consecutively; this is routinely done in AJT studies to minimize the likelihood that participants will learn the nature of the task and develop an undesired response bias (Spinner & Gass, 2019).

### ***Instructions and Familiarization Procedures***

Pre-recorded spoken instructions introduced the task (Ambridge et al., 2008), and participants were asked to help a friendly dog named Bruno learn Spanish by telling him how acceptable each sentence was. After an explanation of the keys and setup, participants heard 10 non-randomized practice trials and were asked to use their keyboard or mouse to rate them on a 4-point scale from 1 (*muy mal* / *very bad*) to 4 (*perfecto* / *perfect*), as shown in Figure 1. Practice trials mirrored the experiment, with each of the clause types represented, but did not contain any of the target verbs. Participant familiarization to the task—but not to the structure of interest—was particularly important given the timed component. Participants were instructed to answer as quickly and as accurately as possible. Non-contingent feedback (*lo estás haciendo muy bien*,

*ahora vamos a hacer más* [you're doing great, let's do some more]) and a visual progress bar were provided. There was no explicit feedback regarding item accuracy.

Figure 1: Participant screen during ATJ experiment, displaying rating scale and progress bar.



### ***Sentence Stimuli***

A total of 50 sentences were recorded in a quiet environment using a Blue Yeti USB microphone. Pre-processing of recordings (e.g., compression, noise reduction) in Audacity ensured high quality and uniform audio quality. Each sentence item was approximately 350ms in duration. Sentences were constructed to be as similar as possible to each other in order to isolate the area of interest (Keating & Jegerski, 2015). One way this was done was to use the same set of target verbs across conditions, offering greater experimental control. Sixteen variable, 16 categorical, and 18 filler sentences were evenly divided between acceptable and unacceptable forms, similar to previous studies (Spinner & Glass, 2019; van Osch et al., 2017). Filler sentences contained a single clause with two-way contrasts between 1) grammatical third-person present indicative and ungrammatical second-person subjunctive forms (*levanta* ~ *\*levantes*) or 2) grammatical

periphrastic future third-person singular and ungrammatical plural forms (*va a levantar ~ \*van a levantar*). These two structures were chosen due to the similar site (i.e., following the verb root), as well as the early-acquired and stable nature of their respective semantic distinctions. Filler sentences served as a baseline to understand participants' general understanding of the task.

### ***Target Verb Characteristics***

Target verbs selected were mid-frequency (and likely to be part of HS' input) belonging to the –AR conjugation class, which is the most frequently occurring verb class. Both –ER and –IR class verbs were excluded, as they take exactly opposite forms to the –AR paradigm (noted in Silva-Corvalán, 2014), and including all classes would create a confound (Mikulski, 2010). Also excluded were verbs with homophonous roots in other classes (e.g., *sent-ar ~ sent-ir*, *as-ar ~ hac-er*, *pod-ar ~ pod-er*) and semantically bleached (all-purpose) verbs like *agarrar* and *coger* (both meaning “get”). Verbs with diphthongizing stems (e.g., *volar/vuel-*) were not excluded. An additional consideration was stigmatized status: verbs like *cachar* (catch), *puchar* (push), and *lonchar* (have for lunch) that otherwise fit these criteria but would likely be tagged as “incorrect” were excluded given the nature of the task (Parodi, 2008). Adult bilinguals reviewed the stimuli set and provided feedback to ensure the appropriateness across Spanish varieties.

While not part of the decision-making process, several other variables were collected. The relative frequency of the matrix verb (i.e., mood-triggering verb) and morphological regularity of the embedded target verb (coded as ‘regular’ or ‘stem change’) were collected (Giancaspro 2017, 2019; Perez-Cortes, 2020). Lemma frequency and subjective ratings including concreteness and familiarity were obtained from EsPal (Duchon et al., 2013).



### ***Mood Clauses***

Sentences were structurally parallel and semantically plausible; each trial required participants to access pragmatic/semantic information contained in the sentence. Categorical and variable conditions were evenly divided between acceptable and unacceptable utterances. The categorical condition included the following: *sabe que* (knows that) and adverbial *porque* (because) clauses with expected indicative, and volitional *quiere que* (wants) and adverbial *para que* (so that) clauses with expected subjunctive.

The variable condition used copula *ser/estar* + adjective (*es importante que* +sbjv. [it's important that], *está claro que* +ind. [it's clear that]; Brown et al., 2021, pp. 266–7); temporal *cuando* (when) clauses proceeding a main clause for habitual/indicative or future/subjunctive interpretations (pp. 260–1); and adjectival relative clauses (p. 267). To elicit the intended contrast, temporal clauses varied in lexical temporal markers and tense of the matrix verb: *siempre se alegra* (she always becomes happy) for habitual readings with expected indicative and *se va a alegrar mañana* (she is going to be happy tomorrow) for future interpretations with expected subjunctive. Adjectival relative clauses with non-human antecedents (and no differential object marking) differed in their specificity (specific +ind. vs. non-specific +sbjv.; Borgonovo et al., 2015; van Osch et al., 2018). Because relative clauses are pragmatically conditioned, each relative clause type also included a disambiguation clause, similar to stimuli in Montrul (2007), to further prime the specific vs. non-specific interpretation; specifically, expected indicative trials ended with *pero no lo encuentra* (but she can't find it), and expected subjunctive trials with *pero no sabe si hay* (but she doesn't know if there is/are any).

### ***AJT Pilot and Feasibility***

Piloting of these online measures was necessary at various stages. To determine the feasibility of this proposal, measure response time across participants, and guide the predicted result, both groups' responsiveness to an earlier version of the AJT was examined. A small sample of U.S.-residing bilingual children ( $n = 2$ ) and adults ( $n = 2$ ) who varied in their Spanish experience, generation, and country of origin completed an untimed version. Descriptively, these preliminary results showed higher acceptability ratings in adults and obligatory conditions. Afterward, pilot study participants answered questions about the task, which informed future iterations. For example, instructions were added that specifically asked participants to answer as accurately and quickly as possible. For the child group, two practice trials were not enough to learn the task, and 10 practice trials were included in the final version.

### **Sentence Repetition**

A novel researcher-designed sentence repetition (or delayed imitation) task was used to elicit subjunctive production across embedded clause types using the same subset of –AR verbs and same clause types compatible with the subjunctive from the AJT task, thus allowing comparisons across tasks. A total of 10 experimental sentences were used (Table 2), presented in random order. Sentences were constructed to be realistic sentences, with a variety of structures and familiar lexical items. Sentences ranged from 9–15 words, all longer than 6–8 syllables to minimize participants' reliance on short-term memory (i.e., language parroting) without knowledge of the structures (Marinis & Armon-Lotem, 2015; Klem et al., 2015). Sentences used canonical SVO word order and were biclausal, with one embedded clause with a single verb inflected in the subjunctive form. Delayed imitation was chosen for technical purposes (i.e., to avoid unintentionally

cutting off the participant recording) and to make the task slightly more difficult than in younger children. Unlike the mood AJT, all sentences were grammatical and contained the present subjunctive in third-person singular (*gane la copa mundial*. [wins the world cup.]) or plural forms (*se queden en sus sillas* [they stay in their chairs]).

### ***Instructions and Familiarization Procedures***

Participants were instructed to help Bruno repeat sentences they heard. Participants heard a practice sentence that did not contain the subjunctive but was of similar length to the experimental trials (*Simón le pidió un libro a su hermano pero él no se lo prestó* [Simón<sub>i</sub> asked his brother for a book, but he didn't lend it to him<sub>i</sub>]). For all trials, participants heard a sentence, which was followed by a short beep. Afterward, they saw a microphone icon to repeat the trial as accurately as they could. This is similar to previous procedures with children in other Romance languages like French or Catalan (Gavarró, 2017; Tuller et al., 2018). The lead researchers of these projects graciously shared their own sentence repetition protocols through personal communication.

Table 2: Sentence repetition items with subjunctive targets.

Sentence	Target –AR Verb	Mood Type	Clause Type
La maestra quiere que todos sus estudiantes se <b>queden</b> en sus sillas.	quedar	obligatory	quiere-que
Los fans quieren que su equipo <b>gane</b> la copa mundial.	ganar	obligatory	quiere-que
Antes de cenar, el papá quiere que su hijo se <b>lave</b> las manos.	lavar	obligatory	quiere-que
Cuando el estudiante <b>escuche</b> la campana, se puede ir a jugar con su gatito.	escuchar	variable	cuando
Diego va a tener más energía cuando se <b>tome</b> la tacita de café.	tomar	variable	cuando
Daniela le guarda los juegos a su amigo para que se <b>apure</b> .	apurar	obligatory	para-que
La abuela le toca la puerta a Miguel para que se <b>levante</b> de la cama.	levantar	obligatory	para-que
La chica se pone los audífonos para que su familia no <b>escuche</b> su música.	escuchar	obligatory	para-que
Es importante que el niño no <b>camine</b> solo por la calle.	caminar	variable	ser+adj-que
La maestra busca un estudiante que <b>termine</b> rápido el problema.	terminar	variable	relative clause

### Demographics and Language History Questionnaire

Parents completed two surveys about their child's and their own language use and demographic information. A Qualtrics survey in English and Spanish asked about children's demographic data, including lunch status, race/ethnicity, English/Spanish varieties spoken, and possible parental concerns. Parent questions about their current ZIP code (optional), country of origin, and English and Spanish proficiency self-ratings using

a 5-point scale ranging from 1 (*principiante / beginner*) to 5 (*como nativo / native-like*). Parents were automatically redirected to complete the Q-BEx in their preferred language.

Additional child-member information regarding language use and proficiency was collected using the Quantifying Bilingual EXperience survey (Q-BEx; De Cat et al., 2022). This tool was developed by a consortium of bilingual researchers and, in contrast to other measures that are completed by researchers or clinicians, Q-BEx is a self-administered survey to be completed by the child or parent; as well, the researcher contacted Q-BEx primary authors via email to ensure appropriateness for this unmoderated web-based study. Q-BEx is available in English and Spanish and contains seven modules and their sub-modules; except for two required modules—*Background Information* and *Risk Factors*—the researcher may select only those modules most appropriate for their research questions (p. 7). The following six modules were selected: *Background Information*, *Risk Factors*, *Language Exposure and Use*, *Language Proficiency*, *Richness of Linguistic Experience*, and *Attitudes and Satisfaction with Child's Language* (Figure 2). Parents had the option of completing the survey in English or Spanish; no bilingual option was available at the time of completion.

Q-BEx survey response files for the two language versions were downloaded from the website, with three data files for each language version. Because Q-BEx generates a different unique ID, respondent information necessary to link to the rest of the data was manually introduced. Values related to language exposure, use, and richness, were automatically output using Q-BEx's back-end calculator. Language-specific variables (e.g., total Spanish hours) were computed by extracting each respondent's first, second, third, etc. language listed.

Figure 2: Selection of Q-BEx modules used.

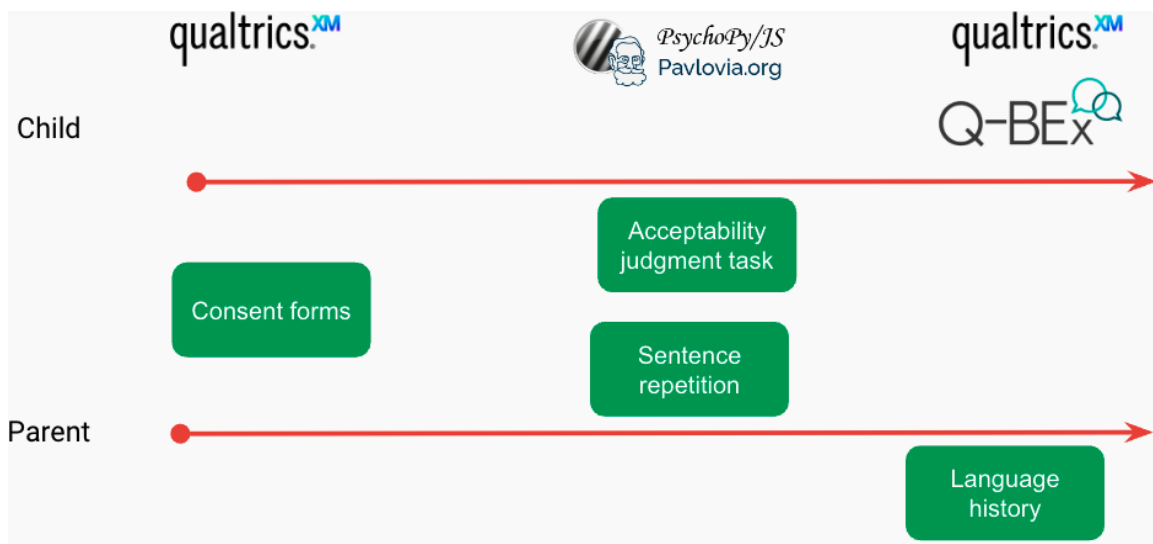
- ☒ **Background information** — *required module*
- ☒ **Risk factors** — *required module — not distributed in child version*
  - ☒ Age of first word — *required sub-module*
  - ☒ Age of first short sentence — *required sub-module*
  - ☒ Concerns — *required sub-module*
- ☒ **Language exposure and use**
  - ☒ Current estimates — *required sub-module*
  - ☐ Cumulative estimates
  - ☒ Weighted estimates
  - ☒ Age and place of first exposure — *required sub-module*
  - ☐ Overheard speech at home
- ☒ **Language proficiency**
  - ☒ Proficiency (no reference group)
  - ☒ Proficiency (reference group) — *not distributed in child version*
- ☒ **Richness of linguistic experience**
  - ☒ Activities
  - ☒ Caregivers' education
  - ☒ Estimated diversity of speakers
- ☒ **Attitudes and satisfaction with child's language**
  - ☐ Attitudes on language mixing
  - ☒ Preferred language
  - ☒ Satisfaction with child's speaking and understanding
  - ☐ Pressure (on caregivers) and (child's) willingness to speak
- ☐ **Language mixing**

## 2.3 PROCEDURES

Data collection was conducted in person and remotely using a personal or researcher-provided device. The researcher wanted to ensure the tasks would be comparable across devices (i.e., phones and tablets), operating systems, and browsers (Google Chrome and Mozilla Firefox). The microphone component was not supported on smartphones during the data collection period, and Safari browser was not well supported. As a result, participants were instructed to use Google Chrome or Mozilla Firefox on their computer, though many families entered the study using their smartphones.

Participants were presented with a Qualtrics survey link containing consent information and were subsequently redirected to the experimental tasks. Both AJT and sentence repetition tasks were delivered in PsychoPy/Pavlovía (Peirce et al., 2019), as shown in Figure 3, and participants' mouse clicks/key presses and reaction time were collected.

Figure 3: Experiment workflow across platforms for child and parent participants.



A challenge for unmoderated studies is cross-platform linking when using various platforms. This dyadic study had the added challenge of requiring tracking both a single participant's data and linking it to a specific dyad partner's data. To ensure appropriate tracking and linking of responses, all participants were automatically assigned a randomly generated unique 4-digit ID within Qualtrics; this ID was subsequently passed to the PsychoPy portion of the experiment using a 'daisy-chaining' approach. In addition, a unique family identifier was created by the parent to allow linking dyads for analysis.

Bot detection was an important consideration given the unmoderated nature of the study, coupled with the compensation incentive and the circulation of flyers containing

the study link in community and online spaces. Before distributing the survey, efforts were made to avoid bots or unintended users by embedding a human verification question (Mitchell et al., 2020); respondents failing the human verification were redirected to an exit message. However, bot respondents and/or speakers who did not meet inclusion criteria were observed at multiple stages of the study. For example, several participants' audio recordings were completed in other languages. In this way, the sentence repetition task itself acted as an additional level of bot detection. In contrast, the AJT task, which did not require speaking and involved 4 keys, could be easily completed by participants who did not qualify. This led to the decision to only include complete cases of both experiment tasks at the expense of some possible data loss.

Language history data was delivered within Qualtrics. Dyad-level data entry was done in Excel, and descriptive data was done in R/RStudio (Ledermann & Kenny, 2015).

### **AJT Acceptable/Unacceptable Ratings**

For the AJT task, dyad members' acceptability ratings on acceptable and unacceptable items were the primary outcome variables. Reaction time data was also collected (Curran, 2016; Montrul et al., 2008). AJT scores were checked for valid responses (i.e., 1, 2, 3, or 4); two- or three-digit entries were considered self-corrections and re-coded with the last integer (e.g., "41" as "1," and "14" as "4"), and extraneous characters were considered unintentional and removed (e.g., "e2" and "2/" as "2"). Missing trials from these participants were excluded ( $n = 50$ ). Acceptability ratings were created separately for expected (AJT-YES) and unexpected (AJT-NO) trials (Spinner & Glass, 2019, p. 117). An *a priori* criterion was used to exclude participants who did not show understanding of the task on filler trials. Specifically, participants who had more than one rating of "1" on expected trials (i.e., strongly rejecting a well-formed sentence)



and/or more than two ratings of “4” on unexpected trials (i.e., strongly accepting an ungrammatical sentence).

### **Sentence Repetition Accuracy**

The sentence repetition task collected participant audio recordings that were automatically transcribed using Google Cloud Speech-to-Text API; off-line second-pass transcription and binary coding (correct/incorrect form) were conducted. PsychoPy transcription features used as a first-pass initial transcript, (Albudoor & Peña, 2022; Fox et al., 2021). Participants’ audio recordings and PsychoPy-generated transcriptions were used to generate a second-pass human transcription, which was entered into an Excel spreadsheet. A form score was calculated for use of subjunctive (1) or any other production (0), without regard to grammaticality. Correct responses included use of subjunctive regardless of verb used. Incorrect responses included indicative and infinitival constructions, regardless of grammaticality. Target verbs whose forms could not be recovered due to low intelligibility or advanced vowel centralization were scored as incorrect. Phonological and grammatical observations such as vowel centralization and article errors were noted.

## **2.4 PLANNED ANALYSES**

Primary research questions regarding group differences were tested using linear mixed-effects models and the actor-partner interdependence model. We describe the rationale for each and outline their respective procedures. Data manipulation, regression analyses, and visualizations were conducted in R/RStudio; the *autoplot* function assessed whether regression assumptions held in the observed data in order to proceed with parametric statistical testing. AJT and production data were assumed to be continuous in

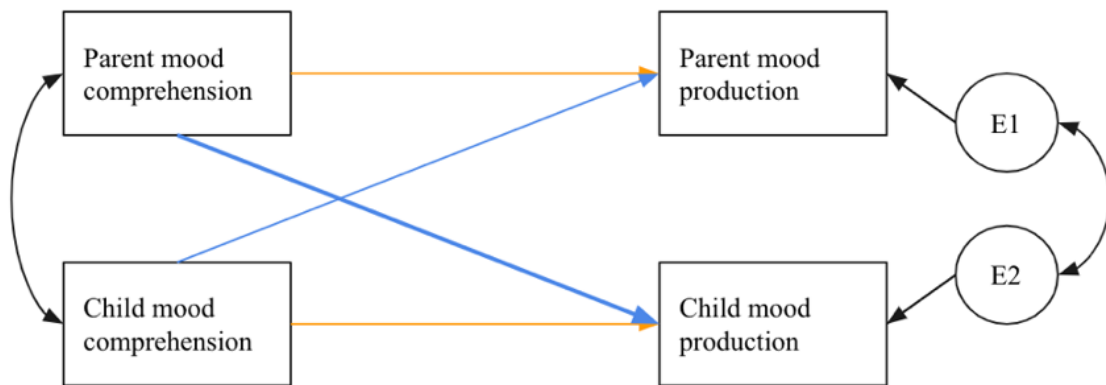
all models, and parametric approaches were chosen (e.g., van Osch et al., 2017). Several studies lend support to the assumption that Likert scales can be treated as continuous and are generally appropriate for parametric approaches when the individual items are not of interest (Carifio & Perla, 2007; but see Ionin & Zyzik, 2014), as was the case here.

Linear mixed-effects models (LMEs) explored linear relationships in parent and child performance. LMEs have become increasingly popular in psycholinguistic (Barr et al., 2013; Matuschek et al., 2017) and applied linguistics inquiry (Cunnings, 2012), particularly for their ability to run one analysis that considers both subject and items as random factors, replacing the need for running two separate F1 by-subject and F2 by-item analysis typical of ANOVAs. Eliminating prior averaging allows modeling of effects that occur throughout an experiment, which is information that can be lost with traditional ANOVA analyses. Another benefit is its ability to handle the cross effects of random factors and fixed factors within the same analysis. Unlike ANOVAs, mixed models are also robust against missing data.

Our fourth research question regarding a parent-to-child partner effect was explored using an actor-partner interdependence model (APIM), a dyadic analysis that empirically tests actor and partner effects (Cook & Kenny, 2005). Failing to recognize nonindependence by using an ANOVA or regression-type model yields biased estimates for significance testing, while running separate parent and child models for these effects results in a loss of power to detect differences (Kenny et al., 2006). For distinguishable dyads that have a readily identifiable within-dyad variable (i.e., family member) as those examined in this study, structural equation modeling (SEM) is a suitable option (Fitzpatrick et al., 2016; Hong & Kim, 2019; Kenny et al., 2006). Data was transformed to dyad-level data for this analysis (Ledermann & Kenny, 2015). The basic APIM using SEM is given in Figure 4, illustrating two actor and two partner effects, as well as

correlated AJT effects and correlated errors or disturbances. SEM models test all possible actor and partner effects, which is best practice even when only one or a subset of these effects are of interest to the researcher. Thus, this model output the effect of parent and child comprehension on their own productions (actor effects), as well as the effect of one member's comprehension on their partner's production (partner effects).

Figure 4: SEM model testing for actor (yellow arrows) and partner effects (blue arrows)



## Chapter 3: Results

### 3.1 DATA OUTPUT AND DATA CLEANING

The following section describes the basic data handling and cleaning procedures used to link across platform-specific data files. Qualtrics, PsychoPy, and Q-BEx results were exported to various .csv data files and cleaned before merging datasets.

The two Qualtrics surveys (i.e., participant consent survey and parent-completed language questionnaire) were downloaded and imported into R/RStudio. The *qualtRics* package was used to remove extraneous header information and retain only the variables of interest (Ginn, 2018). Cases of non-matching email variables ( $n = 3$ ) were flagged and inspected.

PsychoPy provided a .csv output file in long–wide format for each experiment run that yielded information about the presentation of items, the time between trials, as well as the participants’ mouse/keyboard clicks and a transcription of their audio recordings. Since all data is stored in Pavolvia by default, a necessary step was to manually remove all the downloaded files that were generated before the study was made available, before 4/4/2022 ( $n = 208$ ); those that were too small (i.e.,  $\leq 15$  KB) to contain complete data ( $n = 93$ ). Given previously recognized challenges with combining participant files containing different or misaligned columns, a .jar program file, CSVDataMerge, was then used to combine or concatenate these files (Schmidt, 2021); this yielded summary data for each participant that could be imported into R and Mplus.

## 3.2 DESCRIPTIVE STATISTICS

### Acceptable/Unacceptable Ratings

#### *Filler Trials*

On the AJT task, parent and child performance on expected and unexpected filler trials was examined. Given the task instructions and setup, desired mean participant responses would approach 4 for the expected trials and approach 1 for the unexpected trials. On expected trials, the parent group had a mean rating of 3.81, ranging from 3.60 to 4.00. The child group had a mean of 3.34 ( $SD = 0.37$ , range = 2.86–4.00). On unexpected trials, parents had a mean of 1.89 ( $SD = 0.86$ , range = 1.38–3.57), while children had a mean of 3.08 ( $SD = 0.60$ , range = 2.12–3.88). Six children and one parent did not meet the filler trial criterion for the ungrammatical trials, and their AJT data were not analyzed further.

#### *Grammatical/Ungrammatical Trials*

This subset of participants showed lower rates on ungrammatical trials (parent:  $M = 1.90$ ,  $SD = 0.49$ ; child:  $M = 3.28$ ,  $SD = 0.46$ ) than grammatical trials (parent:  $M = 3.68$ ,  $SD = 0.20$ ; child:  $M = 3.65$ ,  $SD = 0.23$ ). Figures 5 and 6 show in both groups' mean grammatical and ungrammatical ratings, respectively.

### Sentence Repetition Subjunctive Accuracy

Subjunctive productions on the sentence repetition task were analyzed for accuracy across 14 parents and 20 children with complete data. Parents repeated the sentences with subjunctive targets close to ceiling ( $M = 0.97$ ,  $SD = 0.07$ ), and children repeated the targets more than half of the time ( $M = 0.56$ ,  $SD = 0.35$ ). Figure 7 shows

both groups' mean performance. The lack of variance in the parent sentence repetition measure precluded models with this outcome variable from being estimated.

Figure 5: Dyad members' mean grammatical ratings by mood variability.

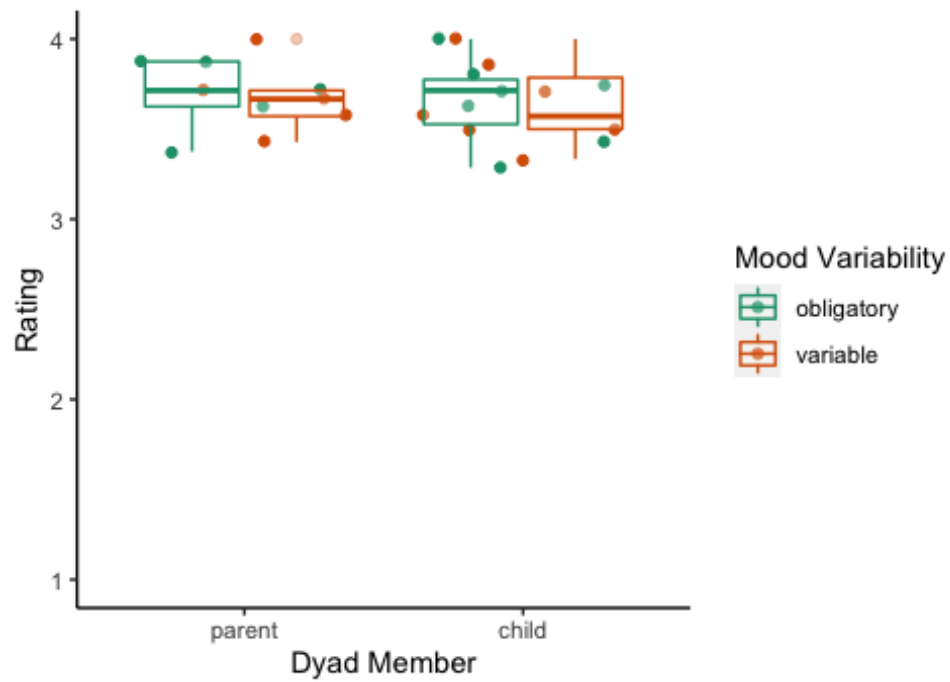


Figure 6: Dyad members' mean ungrammatical ratings by mood variability.

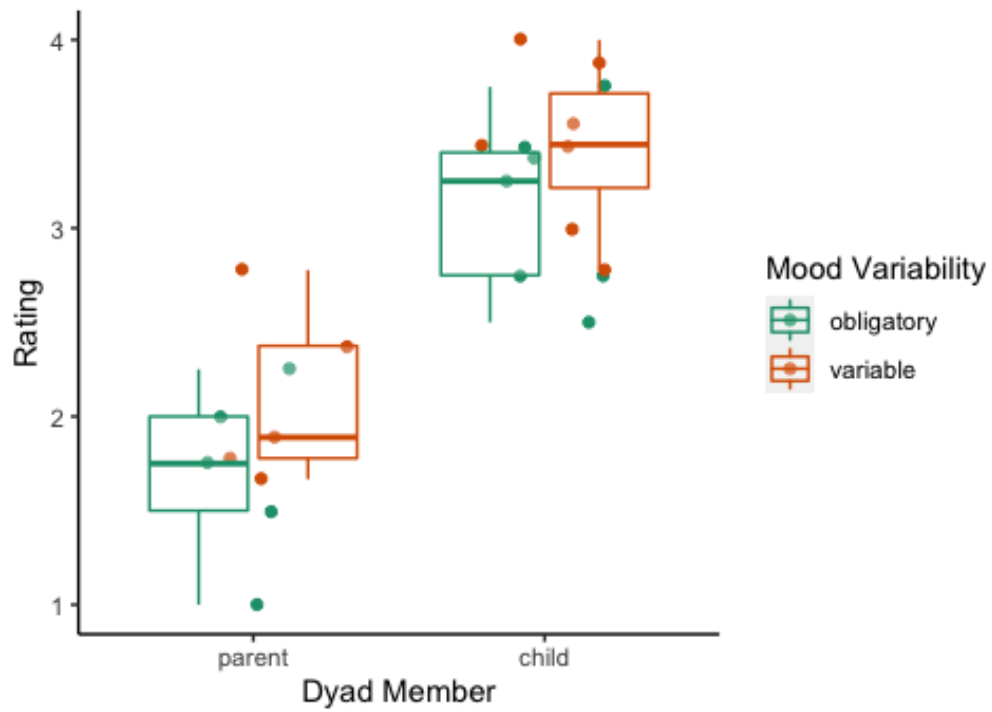
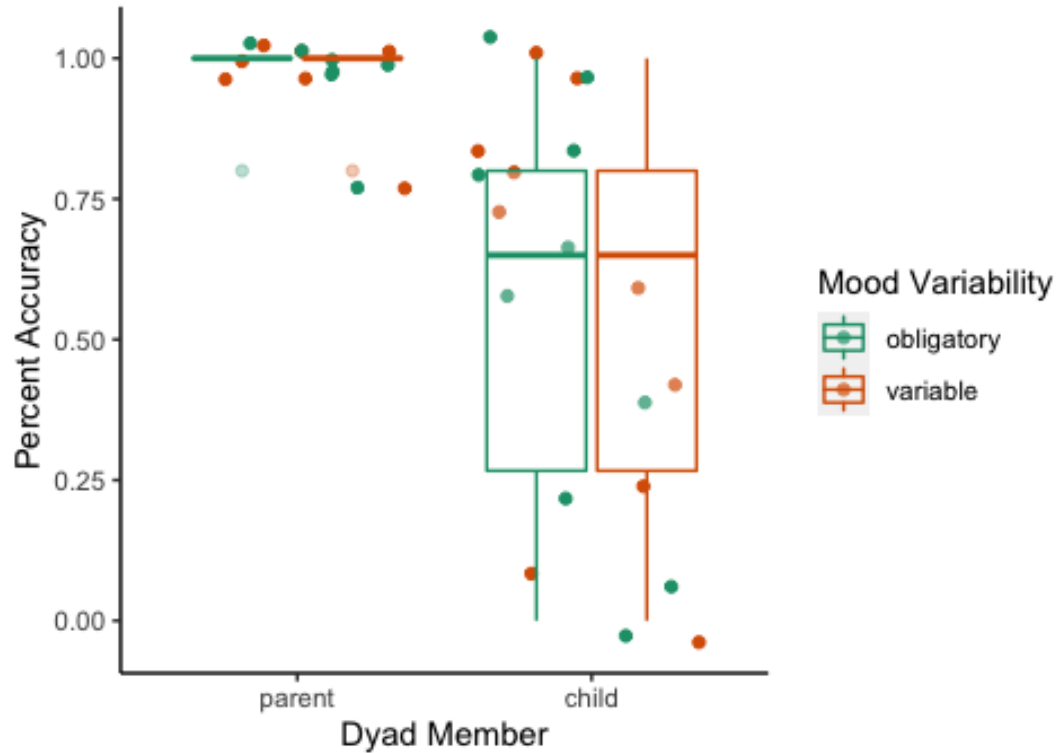


Figure 7: Dyad members' mean sentence repetition accuracy by mood variability.



### 3.3 DYAD MEMBER AND MOOD VARIABILITY EFFECTS

To test whether mood performance differed by dyad member (research question 1) and mood variability (research question 2), linear mixed-effects models were fitted using the *lme4* R package (Winter, 2013), with the following outcome variables: mean acceptability for expected trials (AJT-YES), mean acceptability for unexpected trials (AJT-NO), and sentence target accuracy. Dyad member (i.e., parent or child) and mood variability (i.e., obligatory or variable) were entered as fixed effects; dyads and participants (nested within dyad) were entered as random effects. Given the small number of dyads, no interaction term was entered into the model. Assumptions of homoscedasticity and normality appeared to be met based on visual inspection of data.



The AJT-YES model, which examined predictors of participants' correct acceptance of expected structures, was significant ( $p < .001$ ), but none of the fixed effects were significant, indicating that neither family member nor mood clause type significantly influenced participants' ratings.

The AJT-NO model, which examined predictors of participants' correct rejection of expected structures, was also significant,  $t(17.97) = 19.49$ ,  $p < .001$ . Both family member ( $t(8.60) = -5.46$ ,  $p < .001$ ), and mood variability ( $t(225) = 2.73$ ,  $p = .007$ ) significantly influenced AJT-NO ratings.

The sentence repetition model could not be estimated due to the virtual zero-variance in the parent group. However, descriptively parents clearly outperformed their children, while neither group appeared to be sensitive to mood variability in this task.

### **3.4 CHILD-LEVEL PREDICTORS OF MOOD SELECTION**

To examine whether children's language history was predictive of their mood selection performance (research question 3), a multiple regression model was fitted with the available child data for AJT ( $n = 14$ ) and sentence repetition ( $n = 20$ ). Spanish use and age were entered as predictors. The Spanish use and age zero-order correlation was not significant,  $r = .14$ ,  $p = .07$ . Across all three models, results showed nonsignificant effects of both Spanish use and age ( $ps > .05$ ).

### **3.5 PARTNER EFFECTS**

To examine whether parents' acceptability ratings positively influenced their children's productions (research question 4), an APIM model was run in Mplus. Given the zero-variance in the adult sentence repetition this model was not able to be estimated.

## **Chapter 4: Discussion**

This dissertation contributes to the existing body of work examining input–performance relationships associated with bilingual speakers’ mood selection use. Specifically, we examined parent–child mood selection performance across obligatory and variable clause types common to many spoken varieties. A secondary aim was to demonstrate the use of a dyadic research design and advanced statistical techniques in heritage language research.

### **4.1 GROUP-LEVEL DIFFERENCES**

The first research question tested whether bilingual children and their parents exhibit differences in mood performance across tasks. We predicted that parents would be outperform their children across tasks, which was supported in this sample: the parent group was significantly more accurate at rejecting ungrammatical constructions and in repeating subjunctive forms. Although these findings follow work showing generational differences, this was, to my knowledge the first study establishing these group effects in parent–child dyads. Notably, the child group accepted most grammatical sentences but was much less likely to reject ungrammatical mood sentences. While this may be in line with earlier findings of a general yes-bias (i.e., over-accepting structures) in heritage adults (Zyzik & Sanchez, 2019), this is unlikely given that at least a subset of children were able to reliably reject clearly ungrammatical filler sentences. It is also possible that there are nonlinguistic differences in attention affecting performance in this unmoderated format.

## **4.2 MOOD VARIABILITY PERFORMANCE ACROSS GROUPS**

Our second research question explored the effect of mood variability, and we predicted that obligatory mood would show greater knowledge and use across groups, in line with previous work on acquisition and adult performance (Dracos & Requena, 2022). This was indeed the case for ungrammatical trials. However, no such difference was found in the grammatical trials or the sentence repetition task for either parents or children. There are two plausible explanations for these null findings. First, it may be that specific clause types would reveal more clear patterns (López-Beltrán, 2021), although we were underpowered to conduct such analyses given the sample reported on here. The second explanation concerns specific task demands; the sentence repetition task in particular may not have been challenging enough to reveal these competence—performance differences.

## **4.3 AGE AND SPANISH USE ON CHILD PERFORMANCE**

The third research question explored the contribution of two external factors—children’s age and current Spanish use—on children’s ratings and subjunctive accuracy. Across models, no significant effect of age or current Spanish use was found to be predictive of children’s performance; thus, our predictions were not supported. This was a surprising finding given the robust effects that have been reported at younger ages across various tasks (Bedore et al., 2012; Castilla-Earls et al., 2018). One explanation for these results is that we were underpowered to find a true effect in this group of children. For the Spanish use measure, it is possible that other Q-BEx experiential measures would reveal a significant relationship. An alternative explanation is that these effects are indeed attenuated in older children, and that other social factors such as language attitudes become more important.

The fourth research question could not be answered by the current data. It remains an open question the extent to which form-specific input, compared to broader input measures discussed above, influences children's use of those same forms.

#### **4.4 CONSIDERATIONS FOR DIVERSITY RESEARCH**

In addition to the specific research questions tested, this study had a focus on recruiting a diverse sample of U.S. bilinguals. This was accomplished through the research questions, task design, recruitment procedures, and anticipation of potential barriers. A few comments related to this process are noted here. In terms of task design, designing, recording, and using auditory stimuli is important to move the field forward (Sedarous & Namboodiripad, 2020). Because this may be more time-intensive and costly for the research team, this should be included in research planning. Moreover, auditory stimuli and auditory responses present a challenge, as they are more prone to technical issues, particularly in web-based designs, than written tasks.

In terms of recruitment strategies, participant compensation and active recruitment in local community centers across the country, as well as social media, allowed for a diverse set of participants (Sugden & Moulson, 2015). While the unmoderated nature did not allow for a systematic investigation of specific barriers and potential sampling bias, various barriers to participation were informally observed during the recruitment process. For example, the high number of user views and initial interest did not translate to participation. Some reasons for this high bottleneck effect may include not feeling that they were able to reach out for technical help; one family mentioned there was no "help" button. Another possible reason may have been the requirement for computer access; for example, several families stated they did not have a computer readily available or only a school-provided computer that did not allow access

for non-school sites. In fact, although the unmoderated nature of this study was intended to reduce potential barriers related to scheduling, it is an open question whether a moderated version (i.e., in person or through a Zoom call) may have led to greater participation. As developmental research moves to conducting more web-based studies, researchers committed to increasing diversity of samples should consider these factors in relation to the specific communities, prioritizing smartphone compatibility and possibly allowing for both in-person and unmoderated sessions.

#### **4.5 LIMITATIONS**

The findings and discussion regarding language use and their possible implications should be considered with several limitations in mind. Since our intended sample size was not reached, there is the open possibility that we failed to detect true effects. The parent group's at-ceiling performance on repeating subjunctive forms precluded modeling the APIM; this may likely change with additional recruitment of parents, particularly those who are second-generation. To this end, participant recruitment will continue through additional federal funding to better understand these linguistic and experiential variables in bilingual acquisition.

Given a usage-based understanding of language as changing over the lifespan, it is possible that data from more spontaneous real-world tasks might serve to better approximate these parent-child relationships. Bearing these limitations in mind, this design still offers an innovative approach to a quantitative understanding of input-performance relationships in bilingual speakers. The experimental task completed by child and parent participants offered several advantages over those presented in the literature, including greater experimental control (with pseudo-randomization of items,

sufficient practice and filler trials, and verb controls) and task procedures that were considered more appropriate for the target populations.

#### **4.6 CONCLUSION**

Examining bilingual acquisition in the HL context offers social and practical implications about the largest group of bilingual speakers in the country, as well as theoretical implications about relevant input factors (Valdés, 2005). With a high number of bilingual children in the U.S. (i.e., 1 in 3 children, Child Trends, 2019), there is a clear need to understand how input factors affect a large group of language users and the role that different language experiences and familial factors play. This study examined parent–child links in these grammars and offered novel methodological and analytical contributions to the study of heritage languages. Continued data collection will shed more light on the research questions. Future studies may adopt similar dyadic designs to test existing HL theories. It is also hoped by the researcher that continued methodological/statistical advancements alongside imperatives to understand the linguistic communities studied in order to conduct responsible research that can positively shape language acquisition research related to language minoritized groups.

## References

- Aaron, J. E., & Hernández, J. E. (2007). Quantitative evidence for contact-induced accommodation: Shifts in /s/ reduction patterns in Salvadoran Spanish in Houston. In K. Potowski & R. Cameron (Eds.), *Spanish in Contact: Policy, social and linguistic inquiries* (pp. 327–341). John Benjamins Publishing.
- Albudoor, N., & Peña, E. D. (2022). Identifying language disorder in bilingual children using automatic speech recognition. *Journal of Speech, Language, and Hearing Research*, 65(7), 2648–2661. [https://doi.org/10.1044/2022\\_JSLHR-21-00667](https://doi.org/10.1044/2022_JSLHR-21-00667)
- Ambridge, B., Pine, J. M., Rowland, C. F., & Young, C. R. (2008). The effect of verb semantic class and verb frequency (entrenchment) on children’s and adults’ graded judgements of argument-structure overgeneralization errors. *Cognition*, 106(1), 87–129. <https://doi.org/10.1016/j.cognition.2006.12.015>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bedore, L. M., Peña, E. D., Summers, C. L., Boerger, K. M., Resendiz, M. D., Greene, K., ... & Gillam, R. B. (2012). The measure matters: Language dominance profiles across measures in Spanish–English bilingual children. *Bilingualism: Language and Cognition*, 15(3), 616–629. <https://doi.org/10.1017/S1366728912000090>
- Betancourt, H., & López, S. R. (1993). The study of culture, ethnicity, and race in American psychology. *American Psychologist*, 48(6), 629–637. <https://doi.org/10.1037/0003-066X.48.6.629>
- Blake, R. (1983). Mood selection among Spanish-speaking children, ages 4 to 12. *Bilingual Review/La Revista Bilingüe*, 10(1), 21–32.
- Blything, R. P., Ambridge, B., & Lieven, E. V. (2014). Children use statistics and semantics in the retreat from overgeneralization. *PloS one*, 9(10). <https://doi.org/10.1371/journal.pone.0110009>
- Borgonovo, C., De Garavito, J. B., & Prévost, P. (2015). Mood selection in relative clauses: Interfaces and variability. *Studies in Second Language Acquisition*, 37(1), 33–69. <https://doi.org/10.1017/S0272263114000321>
- Bosque, I. (2012). Mood: Indicative vs. subjunctive. In J. I. Hualde, A. Olarra, & E. O’Rourke (Eds.), *The handbook of Hispanic Linguistics* (pp. 373–394). Oxford, Blackwell.

- Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. (2020). The timing mega-study: comparing a range of experiment generators, both lab-based and online. *PeerJ*. <https://doi.org/10.7717/peerj.9414>
- Brown, A.V., Paz, Y.B., Brown, E.K., & Muñoz-Basols, J. (2021). El léxico-gramática del español: Una aproximación mediante la lingüística de corpus [Spanish Lexical grammar: A corpus-based approach] (1st ed.). Routledge. <https://doi.org/10.4324/9780429026058>
- Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, 1(1), 9. <http://doi.org/10.5334/joc.10>
- Carifio, J., & Perla, R. J. (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Sciences*, 3(3), 106–116.
- Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon’s MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, 29(6), 2156–2160. <https://doi.org/10.1016/j.chb.2013.05.009>
- Castilla-Earls, A., Francis, D., Iglesias, A., & Davidson, K. (2019). The impact of the Spanish-to-English proficiency shift on the grammaticality of English learners. *Journal of Speech, Language, and Hearing Research*, 62(6), 1739–1754. [https://doi.org/10.1044/2018\\_JSLHR-L-18-0324](https://doi.org/10.1044/2018_JSLHR-L-18-0324)
- Castilla-Earls, A., Pérez-Leroux, A. T., Restrepo, M. A., Gaile, D., & Chen, Z. (2018). The complexity of the Spanish subjunctive in bilingual children with SLI. *Language Acquisition*, 25(1), 72–84. <https://doi.org/10.1080/10489223.2016.1192636>
- Child Trends. (2019, March 7). *Dual language learners*. <https://www.childtrends.org/indicators/dual-language-learners>
- Colantoni, L., Martínez, R., Mazzaro, N., Pérez-Leroux, A. T., & Rinaldi, N. (2020). A phonetic account of Spanish-English bilinguals’ divergence with agreement. *Languages*, 5(4), 58. <https://doi.org/10.3390/languages5040058>
- Cook, W. L., & Kenny, D. A. (2005). The actor–partner interdependence model: A model of bidirectional effects in developmental studies. *International Journal of Behavioral Development*, 29(2), 101–109. <https://doi.org/10.1080/01650250444000405>
- Correa, M. (2011). Heritage language learners of Spanish: What role does metalinguistic knowledge play in their acquisition of the subjunctive. In *Selected proceedings of the 13th Hispanic linguistics symposium* (pp. 128-138). Cascadilla.



- Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research*, 28(3), 369–382. <https://doi.org/10.1177%2F0267658312443651>
- Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology*, 66, 4–19. <https://doi.org/10.1016/j.jesp.2015.07.006>
- Cuza, A., & Miller, L. (2015). The protracted acquisition of past tense aspectual values in child heritage Spanish. In R. Klassen, J. M. Liceras, & E. Valenzuela (Eds.), *Hispanic linguistics at the crossroad: Theoretical linguistics, language acquisition, and language contact* (pp. 211–30).
- Cuza, A., Pérez-Tattam, R., Barajas, E., Miller, L., & Sadowski, C. (2013). The development of tense and aspect morphology in child and adult heritage speakers. In J.W. Schwieter (Ed.), *Innovative research and practices in second language acquisition and bilingualism*, 38, (pp. 193–220).
- De Cat, C., Kaščelan, D., Prévost, P., Serratrice, L., Tuller, L., & Unsworth, S. (2022). *Quantifying Bilingual EXperience (Q-BEx): Questionnaire manual and documentation*. <https://doi.org/10.17605/OSF.IO/V7EC8>
- Dracos, M., & Requena, P. E. (2022). Child heritage speakers' acquisition of the Spanish subjunctive in volitional and adverbial clauses. *Language Acquisition*, 1–28. <https://doi.org/10.1080/10489223.2022.2071156>
- Dracos, M., Requena, P., & Miller, K. (2019). Acquisition of mood selection in Spanish-speaking children. *Language Acquisition*, 26(1), 106–118. <https://doi.org/10.1080/10489223.2018.1464006>
- Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods*, 45(4), 1246–1258. <https://doi.org/10.3758/s13428-013-0326-1>
- Duncan, T. S., & Paradis, J. (2020). Home language environment and children's second language acquisition: The special status of input from older siblings. *Journal of Child Language*, 47(5), 1–24. <https://doi.org/10.1017/S0305000919000977>
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in second language acquisition*, 27(2), 141–172. <https://doi.org/10.1017/S0272263105050096>
- Erker, D. (2018). Spanish dialectal contact in the United States. In K. Potowski (Ed.), *The Routledge handbook of Spanish as a heritage language* (pp. 269–283). Routledge.
- Erker, D., & Otheguy, R. (2016). Contact and coherence: Dialectal leveling and structural convergence in NYC Spanish. *Lingua*, 172, 131–146. <https://doi.org/10.1016/j.lingua.2015.10.011>

- Faulkner, T. (2021). Prescriptively or descriptively speaking?: How ‘information-quality’ influences mood variation in Spanish emotive-factive clauses. *Pragmatics*, 31(3), 357–381. <https://doi.org/10.1075/prag.19044.fau>
- Fitzpatrick, J., Gareau, A., Lafontaine, M. F., & Gaudreau, P. (2016). How to use the actor-partner interdependence model (APIM) to estimate different dyadic patterns in Mplus: A step-by-step tutorial. *The Quantitative Methods for Psychology*, 12(1), 74–86. <https://doi.org/10.20982/tqmp.12.1.p074>
- Flores, C., Santos, A. L., Jesus, A., & Marques, R. (2017). Age and input effects in the acquisition of mood in Heritage Portuguese. *Journal of Child Language*, 44(4), 795–828. <https://doi.org/10.1017/S0305000916000222>
- Fox, C. B., Israelsen-Augenstein, M., Jones, S., & Gillam, S. L. (2021). An evaluation of expedited transcription methods for school-age children’s narrative language: Automatic speech recognition and real-time transcription. *Journal of Speech, Language, and Hearing Research*, 64(9), 3533–3548. [https://doi.org/10.1044/2021\\_jslhr-21-00096](https://doi.org/10.1044/2021_jslhr-21-00096)
- Gallego, M., & Alonso-Marks, E. (2014). Degrees of subjunctive vitality among monolingual speakers of Peninsular and Argentinian Spanish. *Borealis—An International Journal of Hispanic Linguistics*, 3(2), 95-104. <https://doi.org/10.7557/1.3.2.3129>
- Gathercole, V. C. M. (2007). Miami and North Wales, so far and yet so near: A constructivist account of morphosyntactic development in bilingual children. *International Journal of Bilingual Education and Bilingualism*, 10(3), 224–247. <https://doi.org/10.2167/beb442.0>
- Gathercole, V. C. M., Sebastián, E., & Soto, P. (1999). The early acquisition of Spanish verbal morphology: Across-the-board or piecemeal knowledge?. *International Journal of Bilingualism*, 3(2-3), 133-182. <https://doi.org/10.1177%2F13670069990030020401>
- Gavarró, A. (2017). A sentence repetition task for Catalan-speaking typically-developing children and children with specific language impairment. *Frontiers in Psychology*, 8, 1865. <https://doi.org/10.3389/fpsyg.2017.01865>
- George, S., Duran, N., & Norris, K. (2014). A systematic review of barriers and facilitators to minority research participation among African Americans, Latinos, Asian Americans, and Pacific Islanders. *American Journal of Public Health*, 104(2), e16-e31. <https://doi.org/10.2105/AJPH.2013.301706>
- Giancaspro, D. (2017). Heritage speakers' production and comprehension of lexically-and contextually-selected subjunctive mood morphology (Doctoral dissertation, Rutgers University-School of Graduate Studies).
- Giancaspro, D. (2019). Over, under and around: Spanish heritage speakers' production (and avoidance) of subjunctive mood. *Heritage Language Journal*, 16(1), 44.

- Giancaspro, D., Perez-Cortes, S., & Higdon, J. (2022). (Ir)regular mood swings: Lexical variability in heritage speakers' oral production of subjunctive mood. *Language Learning*, 72(2), 456–496. <https://doi.org/10.1111/lang.12489>
- Ginn, J. (2018). qualtrics: Retrieve survey data using the Qualtrics API. *Journal of Open Source Software*, 3(24), 690. <https://doi.org/10.21105/joss.00690>
- Green, P., & MacLeod, C. J. (2016). SIMR: an R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4), 493–498. <https://doi.org/10.1111/2041-210X.12504>
- Hong, S., & Kim, S. (2019). Comparisons of multilevel modeling and structural equation modeling approaches to actor–partner interdependence model. *Psychological Reports*, 122(2), 558–574. <https://doi.org/10.1177%2F0033294118766608>
- Hurtado, I., & Montrul, S. (2020). Examining the effect of structural priming on three different populations: Spanish native speakers, Spanish L2 learners, and Spanish heritage speakers. In M. M. Brown & A. Kohut (Eds.), *Proceedings of the 44th Boston University Conference on Language Development* (pp. 322–35).
- Ionin, T., & Zyzik, E. (2014). Judgment and interpretation tasks in second language research. *Annual Review of Applied Linguistics*, 34, 37–64. <https://doi.org/10.1017/S0267190514000026>
- Jacobson, P. F., & Cairns, H. S. (2008). Considering linguistic input in a bilingual situation: Implications for acquisition. *Language, Speech, and Hearing Services in Schools*, 39(3), 352–364. [https://doi.org/10.1044/0161-1461\(2008/033\)](https://doi.org/10.1044/0161-1461(2008/033))
- Jasso, J., Bedore, L. M., & Peña, E. D. (2020). What the sentence repetition task tells us about predictors in bilingual children's Spanish performance. [Manuscript in preparation]. Department of Speech, Language, and Hearing Science, University of Texas at Austin.
- Jegerski, J. (2018). Psycholinguistic perspectives on heritage Spanish. In K. Potowski (Ed.), *The Routledge handbook of Spanish as a heritage language* (pp. 221–234). Routledge.
- Jiang, N. (2012). Conducting reaction time research in second language studies. Routledge. <https://doi.org/10.4324/9780203146255>
- Kaščelan, D., Prévost, P., Serratrice, L., Tuller, L., Unsworth, S., & De Cat, C. (2021). A review of questionnaires quantifying bilingual experience in children: Do they document the same constructs?. *Bilingualism: Language and Cognition*, 1–13. <https://doi.org/10.1017/S1366728921000390>
- Keating, G. D., & Jegerski, J. (2015). Experimental designs in sentence processing research: A methodological review and user's guide. *Studies in Second Language Acquisition*, 37(1), 1–32. <https://doi.org/10.1017/S0272263114000187>

- Kenji, H., & d'Andrea, D. (1992). Some properties of bilingual maintenance and loss in Mexican background high-school students. *Applied Linguistics*, 13(1), 72–99. <https://doi.org/10.1093/applin/13.1.72>
- Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. The Guilford Press.
- Klee, C. A., & Lynch, A. (2009). Contacto del español con el inglés de los Estados Unidos. In *El español en contacto con otras lenguas [Spanish contact with U.S. English]*. Georgetown University Press.
- Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S. A. H., Gustafsson, J. E., & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science*, 18(1), 146–154. <https://doi.org/10.1111/desc.12202>
- Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford publications.
- Komeili, M., & Marshall, C. R. (2013). Sentence repetition as a measure of morphosyntax in monolingual and bilingual children. *Clinical Linguistics & Phonetics*, 27(2), 152–162. <https://doi.org/10.3109/02699206.2012.751625>
- Kupisch, T., & Rothman, J. (2018). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, 22(5), 564–582. <https://doi.org/10.1177%2F1367006916654355>
- Lastra, Y., & Butragueño, P. M. (2012). Aproximación al uso del modo subjuntivo en el corpus sociolingüístico de la ciudad de México [Approximating the use of subjunctive mood in the Sociolinguistic Corpus of Mexico City]. *Boletín de filología*, 47(2), 101–131. <https://doi.org/10.4067/S0718-93032012000200004>
- Ledermann, T., & Kenny, D. A. (2015). A toolbox with programs to restructure and describe dyadic data. *Journal of Social and Personal Relationships*, 32(8), 997–1011. <https://doi.org/10.1177%2F0265407514555273>
- Li, P., Zhang, F., Yu, A., & Zhao, X. (2020). Language History Questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition*, 23(5), 938–944. <https://doi.org/10.1017/S1366728918001153>
- Lipski, J. M. (2008). *Varieties of Spanish in the United States*. Georgetown University Press
- López, B. G., Luque, A., & Piña-Watson, B. (2021). Context, intersectionality, and resilience: Moving toward a more holistic study of bilingualism in cognitive science. *Cultural Diversity and Ethnic Minority Psychology*. <https://doi.org/10.1037/cdp0000472>

- López-Beltrán, P. (2021). *Heritage speakers' online processing of the Spanish subjunctive: A comprehensive usage-based study* (Doctoral dissertation, The Pennsylvania State University).
- López-Beltrán, P., & Carlson, M. T. (2020). How usage-based approaches to language can contribute to a unified theory of heritage grammars. *Linguistics Vanguard*, 6(1). <https://doi.org/10.1515/lingvan-2019-0072>
- Lynch, A. (1999). *The subjunctive in Miami Cuban Spanish: Bilingualism, contact, and language variability* (Doctoral dissertation, University of Minnesota).
- Maniaci, M. R., & Rogge, R. D. (2014). Caring about carelessness: Participant inattention and its effects on research. *Journal of Research in Personality*, 48, 61–83. <https://doi.org/10.1016/j.jrp.2013.09.008>
- Marinis, T., & Armon-Lotem, S. (2015). Sentence repetition. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 95–124) Multilingual Matters.
- Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, 94, 305–315. <https://doi.org/10.1016/j.jml.2017.01.001>
- Mikulski, A. M. (2010). Receptive volitional subjunctive abilities in heritage and traditional foreign language learners of Spanish. *The Modern Language Journal*, 94(2), 217–233. <https://doi.org/10.1111/j.1540-4781.2010.01018.x>
- Miller, K. L., & Schmitt, C. (2012). Variable input and the acquisition of plural morphology. *Language Acquisition*, 19(3), 223–261. <https://doi.org/10.1080/10489223.2012.685026>
- Mitchell, J. W., Chavanduka, T. M., Sullivan, S., & Stephenson, R. (2020). Recommendations from a descriptive evaluation to improve screening procedures for web-based studies with couples: cross-sectional study. *JMIR Public Health and Surveillance*, 6(2), e15079. <https://doi.org/10.2196/15079>
- Montrul, S. (2007). Interpreting mood distinctions in Spanish as a heritage language. In K. Potowski & R. Cameron (Eds.), *Spanish in contact: Policy, social and linguistic inquiries* (pp. 23–40). John Benjamins Publishing. <https://doi.org/10.1075/impact.22.04mon>
- Montrul, S. (2009). Knowledge of tense-aspect and mood in Spanish heritage speakers. *International Journal of Bilingualism*, 13(2), 239–269. <https://doi.org/10.1177%2F1367006909339816>
- Montrul, S. (2016). Dominance and proficiency in early and late bilingualism. In C. Silva-Corvalán & J. Treffers-Daller (Eds.), *Language dominance in bilinguals: Issues of measurement and operationalization* (pp. 15–35). Cambridge University Press. <https://doi.org/10.1017/CBO9781107375345.002>

- Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, 58(3), 503–553. <https://doi.org/10.1111/j.1467-9922.2008.00449.x>
- Montrul, S., & Perpiñán, S. (2011). Assessing differences and similarities between instructed heritage language learners and L2 learners in their knowledge of Spanish tense-aspect and mood (TAM) morphology. *Heritage Language Journal*, 8(1), 90–133. <https://doi.org/10.46538/hlj.8.1.5>
- Montrul, S. A., & Polinsky, M. (Eds.) (Accepted/In press). *The Cambridge Handbook of Heritage Languages and Linguistics*. (Cambridge Handbooks in Language and Linguistics). Cambridge University Press. <https://doi.org/10.1017/9781108766340>
- Muthén, L.K. and Muthén, B.O. (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén.
- Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G., & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry*, 57(11), 1247–1257. <https://doi.org/10.1111/jcpp.12573>
- Obregon, P. A. (2010). *Usage and experiential factors as predictors of Spanish morphosyntactic competence in US heritage speakers* (Doctoral dissertation, The Ohio State University).
- Oller, D. K., & Eilers, R. E. (Eds.). (2002). *Language and literacy in bilingual children* (Vol. 2). Multilingual Matters.
- Otheguy, R. (2013). The linguistic competence of second-generation bilinguals: A critique of 'incomplete acquisition'. In C. Tortora, M. den Dikken, I. L. Montoya, T. O'Neill (Eds.), *Romance Linguistics 2013: Selected Papers from the 43rd Linguistic Symposium on Romance Languages*, (pp. 301–320). John Benjamins Publishing Company.
- Otheguy, R., & Zentella, A. C. (2012). *Spanish in New York: Language contact, dialectal leveling, and structural continuity*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199737406.001.0001>
- Parodi, C. (2008). Stigmatized Spanish inside the classroom and out. In D. M. Brinton, O. Kagan, & S. Bauckus (Eds.), *Heritage language education: A new field emerging* (pp. 199-214). Routledge.
- Pascual y Cabo, D., Lingwall, A., & Rothman, J. (2012) *Applying the interface hypothesis to heritage speaker acquisition: Evidence from Spanish mood*. In:



- BUCLD 36: Proceedings of the 36th annual Boston University Conference on Language Development. Cascadilla Press, Somerville, MA, pp. 437–438.
- Pascual y Cabo, D., & Rothman, J. (2012). The (il) logical problem of heritage speaker bilingualism and incomplete acquisition. *Applied Linguistics*, 33(4), 450–455. <https://doi.org/10.1093/applin/ams037>
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior research methods*, 51(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- Peña, E. D. (Principal Investigator). (2010–2014). *Cross-language outcomes of bilingual experience in typical and atypical development* (Project No. R01DC010366) [Grant]. National Institute of Deafness and other Communication Disorders. [https://projectreporter.nih.gov/project\\_info\\_description.cfm?aid=8662743](https://projectreporter.nih.gov/project_info_description.cfm?aid=8662743)
- Peña, E. D., Gutiérrez-Clellen, V. F., Iglesias, A., Goldstein, B., & Bedore, L. M. (2018). *Bilingual English Spanish Assessment (BESA)*. Brookes.
- Perez-Cortes, S. (2016). *Acquiring obligatory and variable mood selection: Spanish heritage speakers and L2 learners' performance in desideratives and reported speech contexts* [Doctoral dissertation, Rutgers University-Graduate School-New Brunswick]. <https://doi.org/doi:10.7282/T3B27XG0>
- Perez-Cortes, S. (2020). Lexical frequency and morphological regularity as sources of heritage speaker variability in the acquisition of mood. *Second Language Research*. <https://doi.org/10.1177/0267658320918620>
- Pérez-Leroux, A. T. (1998). The acquisition of mood selection in Spanish relative clauses. *Journal of Child Language*, 25(3), 585–604. <https://doi.org/10.1017/S0305000998003614>
- Pham, G., & Tipton, T. (2018). Internal and external factors that support children's minority first language and English. *Language, Speech, and Hearing Services in Schools*, 49(3), 595–606. [https://doi.org/10.1044/2018\\_LSHSS-17-0086](https://doi.org/10.1044/2018_LSHSS-17-0086)
- Pires, A., & Rothman, J. (2009). Disentangling sources of incomplete acquisition: An explanation for competence divergence across heritage grammars. *International Journal of Bilingualism*, 13(2), 211–238. <https://doi.org/10.1177/1367006909339806>
- Polinsky, M., & Scontras, G. (2020). Understanding heritage languages. *Bilingualism: Language and Cognition*, 23(1), 4–20. <https://doi.org/10.1017/S1366728919000245>
- Putnam, M. T., & Sánchez, L. (2013). What's so incomplete about incomplete acquisition?: A prolegomenon to modeling heritage language grammars. *Linguistic Approaches to Bilingualism*, 3(4), 478–508. <https://doi.org/10.1075/lab.3.4.04put>

- Requena, P. E. (2022). Variation versus deviation: Early bilingual acquisition of Spanish differential object marking. *Linguistic Approaches to Bilingualism*.
- Rhodes, M., Rizzo, M., Foster-Hanson, E., Moty, K., Leshin, R., Wang, M. M., Benitez, J., & Ocampo, J. D. (2020). Advancing developmental science via unmoderated remote research with children. *Journal of Cognition and Development*, 21(4), 477–493. <https://doi.org/10.1080/15248372.2020.1797751>
- Ronquest, R. (2013). An acoustic examination of unstressed vowel reduction in heritage Spanish. In *Selected proceedings of the 15th Hispanic linguistics symposium* (pp. 151–171). Cascadilla Proceedings Project.
- Rothman, J. (2008). Linguistic epistemology and the notion of monolingualism. *Sociolinguistic Studies*, 2(3), 441–458. <https://doi.org/10.1558/sols.v2i3.441>
- Rozzi, G. C. (2021). zipcodeR: Advancing the analysis of spatial data at the ZIP code level in R. *Software Impacts*, 9, 100099. <https://doi.org/10.1016/j.simpa.2021.100099>
- Schmidt, J. R. (2021). CSVDataMerge: A simple and free program for concatenating experimental data files. *Journal of Open Research Software*, 9(1)
- Sedarous, Y., & Namboodiripad, S. (2020). Using audio stimuli in acceptability judgment experiments. *Language and Linguistics Compass*, 14(8), e12377. <https://doi.org/10.1111/lnc3.12377>
- Shin, N. (2018). Child heritage speakers' morphosyntax: Rate of acquisition and crosslinguistic influence. In K. Potowski (Ed.), *The Routledge handbook of Spanish as a heritage language* (pp. 235–253). Routledge.
- Silva-Corvalán, C. (1994). The gradual loss of mood distinctions in Los Angeles Spanish. *Language Variation and Change*, 6(3), 255–272. <https://doi.org/10.1017/S095439450000168X>
- Silva-Corvalán, C. (2014). *Bilingual language acquisition: Spanish and English in the first six years*. Cambridge University Press.
- Silva-Corvalán, C., & Enrique-Arias, A. (2017). *Sociolingüística y pragmática del español: segunda edición*. Georgetown University Press.
- Soto-Corominas, A., Paradis, J., Rusk, B. V., Marinova-Todd, S., & Zhang, X. (2020). Oral language profiles of English second language learners in adolescence: Cognitive and input factors influence how they compare to their monolingual peers. *Studies in Second Language Acquisition*, 42(4), 697–720. <https://doi.org/10.1017/S0272263119000767>
- Spinner, P., & Gass, S. M. (2019). *Using judgments in second language acquisition research*. Routledge. <https://doi.org/10.4324/9781315463377>



- Sugden, N. A., & Moulson, M. C. (2015). Recruitment strategies should not be randomly selected: empirically improving recruitment success and diversity in developmental psychology research. *Frontiers in Psychology*, 6, 523. <https://doi.org/10.3389/fpsyg.2015.00523>
- Suñer, M. (1986). Lexical subjects of infinitives in Caribbean Spanish. In *Studies in Romance linguistics* (pp. 189-204). De Gruyter Mouton.
- Surrain, S., & Luk, G. (2019). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005-2015. *Bilingualism: Language and Cognition*, 22(2), 401–415. <https://doi.org/10.1017/S1366728917000682>
- Tiv, M., Gullifer, J. W., Feng, R. Y., & Titone, D. (2020). Using network science to map what Montréal bilinguals talk about across languages and communicative contexts. *Journal of Neurolinguistics*, 56. <https://doi.org/10.1016/j.jneuroling.2020.100913>
- Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, 40(6), 1245–1260. <https://doi.org/10.1044/jslhr.4006.1245>
- Tuller, L. (2015). Clinical use of parental questionnaires in multilingual contexts. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 301–330) Multilingual Matters.
- Tuller, L., Hamann, C., Chilla, S., Ferré, S., Morin, E., Prévost, P., dos Santos, C., Abed Ibrahim, L., & Zebib, R. (2018). Identifying language impairment in bilingual children in France and in Germany. *International Journal of Language & Communication Disorders*, 53(4), 888–904. <https://doi.org/10.1111/1460-6984.12397>
- Valdés, G. (2000). Spanish for native speakers: AATSP professional development series handbook for teachers K-16 (Vol. 1).
- Valdés, G. (2005). Bilingualism, heritage language learners, and SLA research: Opportunities lost or seized?. *The Modern Language Journal*, 89(3), 410–426. <https://doi.org/10.1111/j.1540-4781.2005.00314.x>
- van Osch, B., Hulk, A., Aalberse, S., & Sleeman, P. (2018). Implicit and explicit knowledge of a multiple interface phenomenon: Differential task effects in heritage speakers and L2 speakers of Spanish in the Netherlands. *Languages*, 3(3), 25. <https://doi.org/10.3390/languages3030025>
- van Osch, B., Hulk, A., Sleeman, P., & Aalberse, S. (2017). Knowledge of mood in internal and external interface contexts in Spanish heritage speakers in the Netherlands. In K. Bellamy, M. W. Child, P. González, A. Muntendam, & M. C.

- Parafita Couto (Eds.), *Multidisciplinary Approaches to Bilingualism in the Hispanic and Lusophone World* (pp. 67–94). John Benjamins.
- Viner, K. M. (2018). Conditional morphology in New York heritage Spanish: General and variable usage patterns. *Studies in Hispanic and Lusophone Linguistics*, 11(2), 429–457. <https://doi.org/10.1515/shll-2018-0016>
- Willard, J. A., Agache, A., Jäkel, J., Glück, C. W., & Leyendecker, B. (2015). Family factors predicting vocabulary in Turkish as a heritage language. *Applied Psycholinguistics*, 36(4), 875–898. <https://doi.org/10.1017/S0142716413000544>
- Winter, B. (2019). *Statistics for linguists: An introduction using R*. Routledge.
- Wittenborn, A. K., Dolbin-MacNab, M. L., & Keiley, M. K. (2013). Dyadic research in marriage and family therapy: Methodological considerations. *Journal of Marital and Family Therapy*, 39(1), 5–16. <https://doi.org/10.1111/j.1752-0606.2012.00306.x>
- Wood, C., & Hoge, R. (2019). Average change in sentence repetition by Spanish-English speaking children: kindergarten to first grade. *International Journal of Bilingual Education and Bilingualism*, 22(6), 725–740. <https://doi.org/10.1080/13670050.2017.1308310>
- Zyzik, E. (2016). Toward a prototype model of the heritage language learner. *Innovative strategies for heritage language teaching: A practical guide for the classroom*, 19–38.