Linguistic Approaches to Bilingualism

School, Age, and Exposure Effects in the Child Heritage Language Acquisition of the Spanish Volitional Subjunctive
--Manuscript Draft--

Manuscript Number:	LAB-23057R1
Full Title:	School, Age, and Exposure Effects in the Child Heritage Language Acquisition of the Spanish Volitional Subjunctive
Short Title:	School, Age, and Exposure and the Spanish Subjunctive
Article Type:	Basic science
Keywords:	heritage language acquisition, dual-language education, subjunctive mood
Manuscript Classifications:	10.300: Heritage Language Acquisition; 20.100: Morphology
Abstract:	The present study investigated the production and receptive knowledge of the Spanish subjunctive mood in volitional clauses by 57 English-dominant heritage speakers in fifth, seventh, and eighth grades (ages 10-14), some of whom were enrolled in a dual-language immersion program. Children's self-reported frequency of use of Spanish affected command of this structure, and participants showed more consistent selection of the subjunctive than production of this form. There were no differences in production or selection between children in the dual-language immersion and English-only schools, but older children produced and selected the subjunctive more than younger participants. The lexical frequency of individual subordinate verbs did not affect subjunctive use. These findings argue for the protracted development of this structure. The role for frequency of use and asymmetrical performance between tasks support Putnam and Sánchez's (2013) activation approach to heritage language acquisition. However, the absence of an effect for bilingual schooling or lexical frequency and the increased use of subjunctive mood with age do not strictly align with theories emphasizing exposure effects.
Response to Reviewers:	

Dear Reviewers and Editorial Board,

Thank you for your thoughtful comments on my manuscript. Below please find a section outlining more general changes to the paper, followed by sections to each reviewer's comments. I trust that you will find this version of the manuscript more fluid and coherent, largely thanks to your suggestions, but also due to some changes that I have made to improve flow. I have summarized all of these changes in black. In the following sections, reviewer's comments appear in black, and my responses to these comments appear in red. Finally, changes to the manuscript are highlighted in red. Thank you for the opportunity to collaborate with LAB and I look forward to moving towards publication.

General Revisions

General changes:
- The present project has been adjusted to concentrate more on the effects of lexical frequency and on multiple theory-based accounts. This is because reviewers have commented on these aspects of the project. I have frequency data from this project that were easy to incorporate, so an additional research question was added.
- I have attempted to justify the use of the volitional subjunctive as the structure of interest in this paper, as well as to explain that reassembly is not the only possible result of Putnam and Sánchez's (2013) model. Rather, there could be variability in mapping functions onto form (e.g., the uninterpretable mood feature onto inflectional morphology). These contributions contribute significantly to the coherence of the theoretical arguments advanced in the manuscript.
- I have described how Putnam and Sánchez's (2013) predictions may potentially be adjusted to account for the protracted development of heritage languages. This aligns with recent research that I have published and argues that feature reassembly can be a bidirectional process. To do so, I have cited research on protracted development more prominently throughout the paper.
- The discussion has been restructured to be more concise.

Changes to statistical analysis:
- There is now an analysis section that describes the plan for data analysis.
- I have removed statistical analyses (with the exception of the first two graphs) that concentrate on the Spanish-dominant bilingual adults. I believe that their inclusion may be a distraction from focusing on the factors that accounted for differences between and within individual HSs. While I have provided a more thorough description of who these bilinguals were, I have also refocused the research questions and analyses to concentrate exclusively on heritage speakers. Therefore, the adults are only included as a method for norming the tasks.
- I have included additional figures to visualize each of the predictors. I have described these figures in the manuscript.
- I have placed the individual analyses before the inferential statistics, per a reviewer's suggestion. This seems logical to have all descriptive statistics before inferential analyses.
- I have reduced to a single statistical model that does not require data from Spanish-dominant bilingual adults for the reasons addressed above.
- I have also made some changes to better address the research questions. For instance, properly testing the hypotheses require additional interactions, such as the interaction between frequency of use and task, that were not integrated into the previous modeling. In order to address all of these variables, nested model comparisons were necessary through pairwise comparisons. I have also provided more information on model fitness.
- I have removed the forest plot figures and have reported the statistical models in the prose.

Reviewer #1

Please be sure to cite the Putnam & Sánchez (2013) piece in your references. It's quite surprising that it's not there...

Corrected.

The treatment/discussion of "features" in footnote 1 requires further elaboration, even if the primary focus of this manuscript isn't directly involved with the advancement of any particular theoretical claims. One reference that is worth consulting in this respect is Lohndal & Putnam (2021; hereafter, L&P), who provides a more detailed treatment of the conceptualization of "features" that is largely compatible with P&S's (2013) proposals. As such, the notion of "feature" seems to function as the fundamental unit of linguistic structure in this paper, hence, perhaps it should not appear in a footnote, but rather in the main body of prose.

Thank you for this comment. I have described the role of features and their mapping onto lexical items in greater detail. I have also cited a recent article that I published in LAB supporting this framework with Spanish heritage speakers.

Page 10, lines 22-28: Two points are in order here: First, building upon my previous comment, a more detailed treatment of "features" in this paper would allow A to say something that the features responsible for subjunctive mood (in connection with the morphology that 'realizes' these features). Second, it sounds a bit awkward to say something along the lines of a grammatical feature (or set of features) being "less like to reassemble"; rather, I think that something along the lines of "would be less likely to be impacted", since ultimately what's going on here is a resistance to reassemble (or, perhaps, different associations with allomorphy).

I have rephrased the predictions of the research question to align with the notion of being impacted by reassembly (or not), as the reviewer mentions. Furthermore, I have described how the notion of feature is relevant here for the study of the subjunctive because the uninterpretable mood feature must be mapped onto the appropriate morphology that generates lexical items (e.g., inflected verbs with subjunctive inflections).

A question concerning the interpretation of these experimental results involves whether or not features have been "reassembled", or perhaps, their associate with allomorphy that expresses subjunctive mood has become more variable. Although the P&S (2013) model would support both, an important contribution that A could make here, along the lines of what L&P propose on p. 11 (6) of their article, is that the "loss" or "reconfiguration" of features is one of several outcomes. What some of these groups/individuals in this study may be doing is exhibiting variable feature-morphology associations. Again, although the larger focus and purpose of this paper is not a theoretical contribution per se, it's worth mentioning this, especially since this is in line with the general position offered by A in their interpretation of their own results.

This is an insightful question that deserves greater attention in the manuscript. Features can become reassembled entirely, but difficulty mapping synsem features onto forms is a(n even more) plausible outcome, as indicated in the Lohndal and Putnam (2021) article. I have incorporated this

into the literature review, predictions, and discussion section. In particular, I have attempted to highlight that these two outcomes are situated at different points along a continuum, whereby increasing difficulty of mapping forms to functions is part of the reassembly process (e.g., the early stages of the P&S framework), while "full-fledged" feature reassembly is the outcome of this process (and is actually quite rare in heritage grammars, as Perez-Cortes et al., 2019 discuss).

Page 25, lines 4-24: The question raised by A re: whether or not the P&S (2013) model is adaptable to children, or at the very least requires some deeper thought. I would suggest also taking a look at Putnam et al. (2018) for an expansion of this topic. Surely, this is likely something that will receive short shrift in this piece, but it's worth acknowledging how this could contribute to a larger a discussion on these and related matters.

While I agree that this article *could* be more theoretical in nature, I think that one of the main messages I hope to communicate is that Putnam and Sánchez's (2013) predictions might run in reverse when talking about initial HL acquisition. For instance, children may begin to exhibit more stable receptive knowledge before they develop consistent form-function mappings in production, contra the reassembly process postulated for HL attrition and reassembly. The rate of this process may be driven by patterns of exposure and language dominance, as Putnam et al. (2018) describe. This change features prominently in the discussion.

I would argue more generally that we need follow-up studies with larger samples to confirm all of the findings of the present study.

I have added a sentence to the final paragraph of the discussion section to discuss this.

I recommend including an analysis section in between the methods and results. I am not completely clear on the dependent variables (see specific comment below for p. 18). And if the models are using multiple observations per person (as opposed to a tallied score for one or both tasks), I wonder if would be more appropriate to nest observations (e.g., responses on a task) within individuals. Given the relatively small sample size, an analyses section could also explain the suitability of the statistical models for this number of participants and describe any preliminary tests that may have been conducted to determine the suitability of these models.

Thank you for these comments. I have opted to restructure the statistical analyses completely. I provide more detail about the changes to this analysis in the first section of this review document.

From the abstract, lit review and the research questions, it becomes clear that the study focused on the volitional subjunctive in nominal clauses, which is perfectly reasonable. However, I think that it would help readers—especially those who do not have a background in Spanish—to explain *why* that subjunctive.

This is an insightful comment that I have incorporated into both the introduction of the manuscript and the description of the volitional subjunctive mood. It provides readers with a clearer justification of why I have chosen this element of the subjunctive, defended by citations in recent research (as well as Blake's work). I am confident that this apportion has strengthened the reasoning for using this component of the subjunctive mood in this project.

I understand the choice to hold morphological regularity constant. However, it is important to acknowledge that verbs that have a stem change in the subjunctive are likely to be more salient to students in both oral and written language. For example, the 3s present subjunctive form tenga (indicative counterpart tiene) may stand out more than the 3s present subjunctive form hable (indicative counterpart habla). See Collentine's work, in particular his 1997 piece in Spanish Applied Linguistics. Collentine's participants are traditional FL learners, but the difference in salience is relevant to heritage learners as well.

This is indeed a key finding, and a paper that came out during the process of data collection for this process actually supports this assertion based upon research with heritage speakers (Giancaspro et al., 2022). I have addressed this in the limitations section, as well as through a footnote in the methods section.

Regarding the MLE school, many English-language middle schools in the US offer Spanish as an elective (typically designed for traditional FL students). Did the school that the MLE 7th and 8th graders attended offer such courses? If so, were any of the participants enrolled in those courses?

Along the same lines, I noticed on p. 24 that it says that the DLI group had their input decrease to just 1 class a day. Was this class a Spanish as a foreign/world language class or was it more a of a language arts (reading and composition) class in Spanish? Were they attending schools designated as DLI where the proportion of Spanish instruction shifted, or were these children who attended a DLI elementary but moved into an MLE middle school? These details have important implications for Spanish use at school, so I would like to see more information on this.

Did Gathercole (2002) really use experimental methods? Unless she randomly assigned students to conditions (which seems unlikely), she didn't. Maybe you mean quasi-experimental (e.g., comparison of 2 groups of people that are not randomly assigned to conditions). Maybe you just mean longitudinal and/or comparative.

I tried to look up the specific Gathercole article to verify the methodology used, but the reference is not in the bibliography.

Also, I'm not sure why it's ironic that only Gathercole has done this. Is it that we have a gap in the research that only she has sought to fill? This should be clarified.

p. 9: SDBA= Spanish-Dominant Bilingual Adults, correct? Please spell out the abbreviation upon its first use.

p. 12: MLE-7/8 n = 11, as shown in the table. It says n = 25 in the text, but the table shows that it is the total for MLE at all grade levels.

p. 13: "The children enrolled in the monolingual school were matched for age, socioeconomic status, and family background with those in DLI." How so? This doesn't look like a study with

matched pairs because family background with those in DLI. How so? This doesn't look like a study with matched pairs because Table 1 show us different numbers of participants in each cell. Was it one-to-many matching? Reading the rest of the paragraph, my guess is that the research was conducted at MLE and DLI schools that had similar demographic profiles (i.e., similar socioeconomic status (SES) and family backgrounds), though I'm still unclear on the age part beyond the fact that recruiting from the same grade (e.g., 5th at MLE and DLI schools) will yield participants with similar ages. Choosing schools with similar profiles—while commendable—is not "matching" as it is understood in statistical analyses. I'd like to see a clearer explanation of this matching.

Indeed, this is not matching in the statistical sense, so I have removed this terminology from the manuscript. The manuscript now indicates that the schools were matched based upon their populations, but there is no mention of participant matching.

p. 13: "While only some children's parents spoke English, all spoke Spanish, and all participants were predominantly exposed to Spanish at home (see Table 2)." With regard to "some children's" and "all", are we talking about the children at the school generally, or the participating children? It seems like the latter, but this should be clarified.

Corrected.

p. 13: "The SDBAs... represent a source of input for the HS groups." How so? Where they recruited from the same community where the schools were located? Are some of them parents/family members of child participants? Do some of them teach at the DLI school? Please elaborate.

Yes, all of the SDBAs were from the broader community and most were from the same town; I have clarified this in the manuscript.

p. 13: Please explain how frequency of use of Spanish is being measured. I think that you get to this later, so it may be a matter of moving text up or telling the reader that these measures are described later on.

Corrected.

p. 14: Please explain the proficiency measure used in Table 2. It is not the full DELE, given that it only has a maximum of 18 points. Is this a subset of the DELE? If so, what types of items were included? (It seems like a lot of DELE points for volitional subjunctive only. Is it for all subjunctives or some other combination of items?) Reading on further, it looks like it was the BESA. Again, this is a matter of moving text up or telling the reader that these measures are described later on.

The DELE is not used to assess proficiency in this study. While I reported the SDBA's DELE results descriptively in the previous draft, I have added clarification that the data represented in Table 2 are from the BESA Spanish morphosyntax sentence completion subsection. The revised manuscript no longer mentions the DELE due to de-emphasizing the SDBAs role in this project.

Also, if Table 2 is reporting the BESA results and 4 items were not reported, shouldn't the maximum score on the table be 14?
Yes; corrected.

p. 16: I'm a bit uneasy about the items with creer being used to test indicative mood, just because we can see some mood variation in negated epistemics like no creer que, which could carry over into the affirmative counterparts if the speaker wants to emphasize doubt. A safer bet would have been testing after *tener que*, *saber que*, or *ver que*. I suggest acknowledging this in the limitations.
I have added a sentence to the limitations section; however, multiple published articles (Giancaspro, 2019, 2020) have used this matrix verb as a control with minimal to no use of subjunctive by any of the target populations.

p. 17: "In the HSs' production data, the children produced the subjunctive in a total of 131/426 instances (30.7%), and alternative forms in the remaining 295 instances (69.2%)." If this is the case, why does every child group in Figure 1 have production percentages over 30.7%? I understand that the 30.7% is for all children and Figure 1 shows the subsets, but I don't see how every group doing better than 30.7% can mathematically also work out to 30.7% for all children.
Thank you for this concern. The 30.7% included the SDBA's data and represented the total count of non-subjunctive forms, rather than the amount of subjunctive forms. I have corrected this section of the manuscript by including only HSs' data in the table and by reporting the correct percentage of subjunctive production (45.4%).

p. 18: "In both models, the suppliance of the expected mood inflection was incorporated as the dependent variable..." I'd like to see the phrase clarified. For any participant, one can calculate a preference (receptive) score and a productive score. So, is the DV a combined receptive + productive score? Or is this more of an odds ratio, that is, the odds that participants will answer correctly? This is one thing that could be included in an analysis section.
This has been taken care of by incorporating an analysis section that describes the process of data coding more precisely, per the recommendations of other reviewers and the editor.

p. 21: "some who showed great variability on the proficiency test" Do you mean that these 6 students showed a wide range of scores on the proficiency test?
I have clarified this through clearer wording.

Footnote 9: Do you mean to refer the reader to footnote 8? Both notes may work better as notes on tables rather than footnotes to the text.
I have placed footnote 8 in the table description per your recommendation, and combined the two footnotes into a single one.

p. 26: Like my comment on Gathercole (2002) above, I wonder if "experiments" is the most accurate term here.

I have changed this to "these tasks."

p. 26: Although every verb in the present study was a regular –ar verb, the authors could look at frequency of those verbs. That is, did students perform better on the items featuring more frequent verbs?

This is an astute observation and lexical frequency data was available. In the first draft, I opted not to analyze this variable because doing so would require reviewing an additional body of work and adding an additional research question. However, I have ultimately opted to include lexical frequency in my analysis and to make it a more salient point throughout the paper. I have described how lexical frequency relates to Putnam and Sánchez's (2013) predictions, and have presented the existing body of work on lexical frequency effects in Spanish as a heritage language. Additionally, I added a research question on lexical frequency, and described my methods for addressing frequency in the tasks. Finally, I incorporated paragraphs on lexical frequency in the discussion section of this paper.

p. 26: Regarding the discrepancy between the findings of the present study and those of Potowski (2007a), how many of the subjunctive items in the latter had stem changes? This could be another factor in addition to the amount of Spanish used for instruction at the DLI school.

This is an excellent question. Potowski's tasks were open-ended, so they may have favored subjunctive use because they could choose their own verbs. However, due to space limitations in the revised manuscript, I opted to remove this paragraph, because it does not contribute to the central arguments of the paper.

p. 28: I agree that we need to consider the importance of output in Spanish (which a DLI setting does not necessarily guarantee) and the role of input quality when considering why we didn't see differences between MLE and DLI students. Something else to consider is the nature of language use questionnaires. Language use at home can vary depending on the interlocutor(s) and the topic of conversation, and it's difficult for a brief questionnaire to capture all of that complexity.

I have added this very accurate observation into the paragraph.

I see other minor errors and recommend a thorough proofreading of the manuscript.

This has been part of the revisions.

Reviewer #3
The reviewer provided multiple corrections on the manuscript itself. Most of these corrections are simple, so I have not included them below. Only those comments that were left on the manuscript and that require further elaboration are included.

Please carefully proofread the paper again; a number of (grammatical) errors and errors in Figures have been pointed out in the PDF. In addition, please introduce all abbreviations at first mention.
Thank you for your comments. The manuscript has been largely rewritten and has been reread multiple times.

Some sentences are very long, run over multiple lines, and are separated by numerous citations within – this makes reading and understanding somewhat cumbersome. Perhaps consider rephrasing some of these to produce as smoother writing style.
Thank you for your recommendations. I have used this writing style in recent publications, including one in LAB. I have attempted to make some sentences shorter, although I do find that there are some areas in which it would not be appropriate to move citations to the end of the sentence. In these instances, the citations refer to the specific part of the sentence. Where possible, I have broken these into separate sentences. The other reviewers indicated that they found the manuscript to be clear, so I trust that this revision strikes a balance between the reviewer's concern and adequately representing the ideas of different authors who are cited.

I think some of the contextualization needs to be a bit more in depth. For example, who are your SDBAs? And please provide more context to the type of schooling – this information was bit scattered here and there and appeared a bit inconsistent. Please check and provide a more coherent picture.
I have added information about the SDBA's community. I have also incorporated information about who these participants were (undergraduate or graduate students, which describes their occupation and level of education).

Whereas your analysis seems solid and the individual steps you took can easily be followed, I encourage the author(s) to also briefly describe tables/figures to point out what the most relevant information is or what the reader is supposed to see from these tables and figures. However, some important information is missing, and this should be added: you mention that you have a random structure in both regression models. However, nothing is said about the impact of the random effects. Please add. In addition, the model fit of both models needs to be reported.
I have added information about Figures 1 and 2 to the text in a descriptive statistics section before the statistical modeling. I also included three additional descriptive figures, and have removed the forest plots (visual model outputs) to avoid confusion. Moreover, I have conducted and reported nested model comparisons and revised my statistical modeling accordingly. These comparisons first compare models with different fixed effects that contain participant and item as random intercepts. The model of best fit included the predictors school, age group, BESA proficiency, frequency of use, and task. Then, I attempted to carry out a second nested model comparison

between two GLMMs that evaluated these same predictors, with one model containing random intercepts only, (1 | Participant) + (1 | Item), for participant and item, and one with random intercepts *and* random slopes, (1 + Item | Participant) + (1 + Participant | Item). The model with random intercepts and random slopes failed to converge, so the final model included for statistical analysis was the one with the lowest AIC in the *anova* comparing the models with random intercepts only. I referred to Cunning's (2012) paper here and have cited that in the manuscript. I trust that this significantly improves the transparency of my data reporting, all of which is traceable on a public repository.

Specific comment to Figure 4/Model 2: Note that if you have an interaction in your model (which you have, namely 3!), you shall interpret the interactions instead of the main effects! This might actually have an effect on your interpretation here, but since this is not given in your results section, it is a bit hard to tell how you interpret the findings. Please add this and also pay attention to how to interpret such regression models. In addition to this, I would suggest considering a step-wise model building process (backward, or forward). You could then, if statistically not significant, remove/or not add predictors/interactions, rather than keeping them. Take a look, for example at Gries 2021. One possible way could be to start out with the maximum model (all main effects, all two-way interactions) and then to step-wise remove non-significant predictors (either main effects, if not in an interaction, or interactions) via drop1 in R.

Thank you for this comment. I have described the extensive changes to the statistical modeling in the first section of this document. Please note that I have used nested model comparisons to construct the model of best fit through pairwise comparisons. Based upon the results of the nested model comparisons, the models of best fit never included any interactions, so this is also now resolved through the revised analysis.

I am a bit skeptical about the discussion, simply because of the interpretation of Model 2 (see my comment above). Perhaps once this has been taken care off, the discussion needs to be adjusted.

I have restructured the statistical analyses, as addressed in the first subsection of this document, and have adjusted the discussion section according to the results obtained.

Perhaps consider having a somewhat more introductory start into the topic by having a shorter introduction and then, as a separate section, to systematically provide an overview of HSs language development.

Corrected.

Is this still up-to-date? 2008, when this was published, is quite a while ago..

It is up to date because the date was mis-listed; I have corrected this to Montrul (2018) and added this to the citations section. Thank you for pointing this out.

How is "morphosyntactic proficiency" a way to operationalize exposure?

Following the researchers cited in the paper, proficiency (both lexical and morphosyntactic) has been operationalized as a proxy to language exposure (see Giancaspro & Sánchez, 2021; López Otero et al., 2023a, 2023b). The argument is that bilinguals who are more exposed to a language develop higher proficiency in it. While I agree that this argument seems to render proficiency a bit redundant with patterns of exposure, most manuscripts continue to analyze this variable. Since this variable was included in Dracos and Requena's (2022) study, which is cited at length throughout my manuscript, I have opted to use a similar proficiency measure to maintain consistency across results.

The setting is not entirely clear - which age cohorts are you referring to here?
Clarified.

So DLI does not only mean that content subjects are taught in Spanish and English, but that English is also taught as a separate subject? It would perhaps be useful to provide more information on DLI programs.
Due to space limitations, I have referred readers to another reference.

At this point, I am wondering, if it made sense to provide some actual sentence-level examples (contrasting subjunctive, with indicative and imperative?) to make it even more transparent for the reader what it is you are investigating (especially since you are not investigating all but only a special case of subjunctive uses?).
This is a logical suggestion, but due the incorporation of additional analyses and the increased word count, I have chosen to leave this as-is due to space limitations.

What does this mean?
These are syntactic categories from Rizzi's cartographic theory that have become standard in literature. I have included a brief description in the manuscript.

Why did you treat 7th and 8h grade as one group. Wouldn't you expect differences also between grade 7 and grade 8 students?
Ideally, I would have incorporated grade as a continuous variable, but there weren't enough participants. This is consistent with Potowski (2007b). I have addressed this as a limitation.

50% vs daily - so in the former, Spanish was not used daily, but in the latter, if daily, what is the % between Spanish and English?
I have clarified that all groups get daily Spanish instruction (e.g., 50% each day versus just one course each day).

How was SES operationalized?

The school reports do not discuss how SES is operationalized, and this is a comment addressed in my dissertation on a similar topic. However, in consistency with the conventions of educational research, the school reports have been anonymized here to protect the identity of the schools and their students. Unfortunately, this is an insurmountable limitation that I have addressed in a footnote.

What is their age?
I have added the average age and SD in the prose of the manuscript.

I find it a bit hard to believe that these children managed to rate their Spanish use on a 5-point Likert scale. Have you piloted and tested, how well this worked with children completing it independently?
I did not pilot this instrument. Firstly, previous research has adopted similar methods with children who were considerably younger than those in my study (Castilla-Earls et al., 2022). Specifically, the authors demonstrate that their instrument, which operationalizes self-reported exposure to Spanish, has high psychometric validity in children between ages four and eight. Furthermore, children are better judges of their own interactions with their peers while at school or in the community, since their parents are not there. I have included the Castilla-Earls et al. (2022) reference in the citations.

To me, it is still unclear what kind of test it was. Oral? written? complete gaps? Could you perhaps specify this?
It was written; this is now included in the manuscript.

You keep calling the second test differently - preference task, selection --- perhaps better use one term only?
I have addressed this throughout the revised manuscript, including in the figures, and now refer to the second task as the "selection task."

What is currently missing is the overall fit of the model, i.e., the predictive power. Please add this for both models as this is important in assessing the quality of the models!
AIC added. If this is not a sufficient measure of model fit, it would be beneficial for the reviewer to provide additional information about what s/he/they identifies as necessary to demonstrate model fitness.

What does "standardized prior to analysis" mean? This should be mentioned and explain in the methods section!
I trust that the inclusion of an analysis section that describes the computation of the variables has addressed this comment sufficiently.

I am surprised - wouldn't this be something you'd check before running the regression models, as part of the data inspection? I'd consider moving this part before the regression models.
Corrected.

I'd even try if a correlation can be constructed here - there seems to be such a trend. Moreover, why are some points not on the lines? How can, for example the lowest red point be inbetween 1 and 2 (x-axis) if this is the absolute number of of subjunctives produced? how can 1.5 sentences be produced? Please explain.
Thank you for the comment. The purposes of individual analyses are to provide descriptive support for correlations, but this would not be appropriate. The graph was generated using the *jitter* layer using *ggplot* in R, which displaces the points slightly along a graph to enhance their visibility. Otherwise, many data points would gather along the same point (e.g., x1, y2) and only one point would be visible. I have two published articles with this format and would prefer to be consistent. To address your concern, I have added a footnote.

This is what I've commented on before (and you also mentioned yourself) - this might cast some doubt on the measure of the questionnaire in eliciting this kind of (truthful) information...
Thank you for your comment. I hope that in the revised version of the manuscript, it becomes clear that students often overhear Spanish, but do not consistently respond using this language. This provides a natural explanation for why students reported "0" on the language questionnaire. In fact, it's a more honest measure than what parents may provide since they are not with their children in their schools. There are many studies (Babino & Stewart, 2017; Ballinger & Lyster, 2011; Hamman, 2018; Potowski, 2004, just to name a few) that have found this to be the case. Therefore, this finding is indeed in line with previous studies that have used multiple methodologies and that have identified that oftentimes, children do not use Spanish in immersion programs even though they are exposed to it. I have dedicated considerable attention to this topic in the revised discussion.

Perhaps instead of a footnote, this could be either directly mentioned in the text (also indicating that this concerns both Tables 5 and 6) or added as a note below both tables.
Thank you for your comment. I feel that including this in the description of the tables would be cumbersome, because the description is intended to be brief. I have chosen to leave the footnotes intact as the best possible solution.

But Freq. of use and parental languages does not necessarily suggest this? What is your thought here?
Heritage language research has consistently shown that Spanish HSs often respond in English, even when input is in Spanish (see Babino and Stewart, 2017 or Hamman, 2018 for recent examples in the context of bilingual schools). So, these ratings are in clear alignment with these findings. To expound, simply overhearing Spanish in contexts of heritage language development does not necessarily predict output in that language.

This really makes me wonder again who these SDBAs are – this information is, I think, absent from the paper...
Corrected; this is a very insightful comment and I have made a concerted effort to show that there is no evidence of attrition in the caregiver input.

Where was this shown? I believe in Fig. 4? note my comment above (and in my response letter) about interpreting main effects and interactions.
This is no longer shown in a figure, as I have removed the forest plots (visual model outputs). They seemed to cause more confusion than clarification. Regardless, the model of best fit did not include any interactions, so this comment should now be taken care of. None of the data represented in the first draft or in the current manuscript involve interactions, but I was wondering if the reviewer would be willing to clarify why main effects should be ignored in the event that interactions are significant. I have not seen this in any previous research on Spanish as a second or heritage language, and it is also contra my statistical training. I would value this information for future manuscripts, but I would also be very surprised if main effects should be altogether discarded simply because an interaction involving one of these effects is also significant. For instance, it is completely plausible that, in a comparison of HSs and L2 learners, there is an overall effect for proficiency, as well as a group-by-proficiency interaction (see, for example, Montrul & Perpiñán, 2011). It would not be proper to discard the proficiency effect simply because one group is more affected by proficiency than the other. Rather, while L2 speakers are more effected by proficiency in such an instance than HSs, this is not to say that the overall proficiency effect does not matter (and therefore this main effect need not be discarded).

Same as before, please check your model again!
Thank you for your comment. I am not quite clear on what the reviewer would like for me to change here. The effect for age group indicates that older children were more likely to use the subjunctive than younger children. This does not involve any interaction and is clearly stated in the outputs of the GLMM model.

This substantial information also needs to be added above (see one of my earlier comments).
This information is available in the participants section.

Preposition missing?
The preposition here is "contra," which is standard in academic writing.

Given that you have this information in your regression model, I'd suggest to visually inspect the interaction rather than purely checking % for each group. A multifactorial perspective is to be preferred, I think.

Thank you for this comment. I am not clear as to which interaction the reviewer is referencing here. In the first draft of the manuscript, there were no significant interactions. In the revised version, due to nested model comparisons, none of these interactions were included in the models, so it is unclear how this comment relates to the information presented here.

Def. article missing.
"On one hand" is a standard phrase in American English.

Abstract:

The present study investigated the production and receptive knowledge of the Spanish subjunctive mood in volitional clauses by 57 English-dominant heritage speakers in fifth, seventh, and eighth grades (ages 10-14), some of whom were enrolled in a dual-language immersion program. Children's self-reported frequency of use of Spanish affected command of this structure, and participants showed more consistent selection of the subjunctive than production of this form. There were no differences in production or selection between children in the dual-language immersion and English-only schools, but older children produced and selected the subjunctive more than younger participants. The lexical frequency of individual subordinate verbs did not affect subjunctive use. These findings argue for the protracted development of this structure. The role for frequency of use and asymmetrical performance between tasks support Putnam and Sánchez's (2013) activation approach to heritage language acquisition. However, the absence of an effect for bilingual schooling or lexical frequency and the increased use of subjunctive mood with age do not strictly align with theories emphasizing exposure effects.

Keywords: heritage language acquisition, dual-language education, subjunctive mood

School, Age, and Exposure Effects in the Child Heritage Language Acquisition of the Spanish Volitional Subjunctive

1. Introduction

A central question in bilingual children's development is how patterns of exposure characterize the acquisition of heritage languages (HLs). HLs are spoken in situations of language contact where speakers frequently develop stronger competence in a more socially prevalent language. The most researched example is Spanish heritage speakers (HSs) in the United States, who tend to experience a shift in dominance towards English at the start of schooling (i.e., Castilla-Earls et al., 2019; Hiebert & Rojas, 2021). Research on bilingualism has generally focused on preschool children and adults, so school-aged children are the "missing link" (Montrul, 2018, p. 534) in our understanding of HL development: comparisons of younger versus older children can distinguish between language attrition and protracted development.

Evaluating HL acquisition under different contexts of exposure and by comparing productive and receptive knowledge provides a more holistic view of HSs' developmental trajectory. Putnam and Sánchez (2013) advance a proposal that accounts for addressing receptive and productive knowledge distinctly. These researchers argue that activation of the HL in the memory leads to processing input for intake, which strengthens the associations between features and their morphological realization. Putnam et al. (2019, p. 19) define these features as "indices on lexical items and larger syntactic objects that allow generated structures to be interpreted at external interfaces." Similarly, Lohndal and Putnam (2021) claim that these abstract syntactic and semantic features are mapped onto morphology and are joined together to form lexical items that can be spelled out and interpreted.

Putnam and Sánchez (2013) claim that decreases in HL exposure bring about crosslinguistic influence from the dominant language, which can cause variability in mapping

features onto morphology to generate lexical items, and eventually can result in the reassembly of these features at the underlying level. Following these researchers, the reassembly process begins under the online pressures of language production due to decreases in exposure, such that bilinguals may show stronger receptive knowledge than what they exhibit in production. Moreover, Putnam and Sánchez (2013) predict that it is harder to map features together to create lexical items that are activated less-frequently in the memory. Therefore, HSs are more likely to produce or interpret features that are realized on more-frequent lexical items.

This framework provides specific and testable predictions that can explain different acquisitional outcomes between individual HSs. Nonetheless, this approach applies a process of progressive language restructuring due to crosslinguistic influence, while other studies have shown the opposite, whereby child HSs acquire Spanish in a protracted fashion (Corbet & Domínguez, 2020; Cuza & Miller, 2015; Cuza & Solano-Escobar, 2023; Martinez Nieto & Restrepo, 2022; Montrul & Potowski, 2007; Montrul & Sánchez-Walker, 2013). Therefore, a study testing the influence of both age and patterns of exposure is highly relevant for evaluating if and how these two approaches to HL acquisition may interact in accounting for child HSs' development.

2. Exposure, Heritage Language Development, and Dual Language Immersion

There are multiple ways to empiricalize exposure. This study incorporates overall frequency of HL use, morphosyntactic proficiency, and bilingual education as metrics for exposure to explore HS children's acquisition of subjunctive mood in volitional clauses. Proficiency and frequency of use have been used as proxies for HL exposure in recent research (Dracos & Requena, 2022; Giancaspro & Sánchez, 2019; López-Otero et al., 2023a, 2023b; Perez-Cortes, 2016). However, understanding the impact of exposure through education in the HL is an underexplored approach. To measure input at school, the present study compared Spanish HSs in a traditional English-only

school with age-matched peers in a dual-language immersion (DLI) program where 50% of daily instruction is offered in Spanish during the elementary years.[1]

Evaluating language development in DLI is an underexplored avenue for understanding the impact of exposure in childhood, as HSs enrolled in these programs purportedly receive greater quantity and quality of HL input at a critical developmental time. From the perspective of input quantity, DLI supplements home exposure to Spanish during the school years at a time when comparable children in traditional schools typically begin to exhibit a shift in dominance towards English (Castilla-Earls et al., 2019; Hiebert & Rojas, 2021). Monolingual children master the structure tested here, the volitional subjunctive, around the start of school (Blake, 1983; Dracos et al., 2019). Therefore, HL exposure may be especially impactful during this time in the acquisition of this structure. From the perspective of input quality, DLI schools purportedly offer students with access to vocabulary specific to academic content areas (larger lexicon) and to lengthier and more complex discourse (greater morphological and syntactic diversity).

To date, few studies have employed methods from bilingualism research to study the development of HSs' Spanish language skills in DLI when compared to children of similar characteristics in English-only schools (but see Gathercole, 2002 and AUTHOR, XXXX). While bilingual education has previously been identified as facilitative in the acquisition of morphosyntax in German-dominant HSs, these studies focused on adults who were not actively enrolled in school (see Kupisch and Rothman, 2018). Therefore, exploring the productive and receptive knowledge of Spanish HSs with different quantities of input at home and school *at the time of testing* is important for research on HL development.

---

[1] Due to space limitations, readers are referred to Lindholm-Leary (2018) for further discussion of the premises and research concerning DLI education with Spanish HSs.

The volitional subjunctive is an ideal area of the Spanish inflectional system to investigate in such a context for four reasons. Firstly, it shows minimal to no dialectal variation, unlike some subjunctive contexts (e.g., Faulkner, 2021). Secondly, monolingual children master the volitional subjunctive around the start of schooling (Blake, 1983; Dracos et al., 2019). Since HSs enrolled in DLI programs have more input in Spanish during the time when monolinguals master the subjunctive, this method of education should be especially impactful on the acquisition of this structure. Thirdly, the subjunctive typically occurs in subordinate clauses, so HSs' acquisition of this structure may benefit from academic discourse in DLI programs that features complex, multi-clause sentences. Finally, the subjunctive is realized through verbal morphology, so it is possible to control for lexical frequency on productive and receptive tasks to most fruitfully test Putnam and Sánchez's (2013) predictions. Additional research on the subjunctive is reviewed in the following section. The research questions, hypotheses, and methodology follow. After presenting results, this article concludes with a discussion of findings, their implications, and their limitations.

3. Spanish Subjunctive Mood: Theory and Acquisition

If patterns of exposure are deterministic in HL acquisition, a logical hypothesis is that structures that emerge late in monolingual populations will be particularly susceptible to input effects. The subjunctive is one such structure and is one of three moods in Spanish alongside the indicative and imperative (Seco, 1990). Mood is a morphological realization of modality, which is the evaluation of the truth value of propositions (Bosque, 2012). All Spanish verbal inflections encode one of the three moods, although studies show that as few as 7.2% of inflected verbs are marked with subjunctive morphology (Biber et al., 2006). This points to its infrequence in the input that HSs receive, especially if this exposure is less in quantity compared to monolinguals. As stated

previously, frequent exposure, particularly through DLI, would purportedly be especially relevant for acquiring this structure.

There are two morphological paradigms for the subjunctive, one in the present tense and the other in the imperfective past. Both have forms for subject person and number agreement. This study concerns the present subjunctive, which is formed either through a shift in verbal inflection or through both a morphophonological change in the verb stem and a shift in inflection from the more-frequent indicative. The syntax and semantics of subjunctive mood have been the topic of considerable scholarship. Fábregas (2014) argues that subjunctive inflections comprise a single spell-out of multiple structures that differ in their syntactic and semantic representations. These inflections occur almost exclusively in subordinate clauses whose subject differs from the main clause. Some uses of the subjunctive, such as volitional clauses as in (1), result from a process of obligatory lexical selection, as a lexical item in the matrix clause subcategorizes for the subjunctive in tensed subordinate clauses with a distinct subject.

(1)   *Rosa quiere que Carmen venga a su casa.*
Rosa want-3PS-IND COMP Carmen come-3PS-SUBJ to her house.
Rosa wants Carmen to come to her house.

Following Kempchinsky (2009), subjunctive inflections in volitional clauses such as (1) are the morphological realization of an uninterpretable feature that is checked and deleted in the heads of Force and Fin.[2] Working within feature-oriented frameworks of HL acquisition (Lohndal & Putnam, 2021; Putnam & Sánchez, 2013), the task for the Spanish speaker in acquiring the volitional subjunctive is twofold. Firstly, one must learn which matrix items select the

---

[2]   Following Cinque and Rizzi's (2012) cartographic analysis, both Force and Fin are syntactic categories in the left periphery below the CP layer.

uninterpretable mood feature; secondly, one must map this feature onto (or interpret it through) inflectional morphology to generate lexical items.

In contrast to Spanish, English has a less-utilized subjunctive. Acceptability judgment data have shown that the English indicative exists as a grammatical alternative to the subjunctive (Iverson et al., 2008; Rojas, 1998). In addition, English lacks a morphological paradigm that is specific to mood, as it is marked through past tense forms. Moreover, English requires control structures in volitional contexts where subjunctive mood is expected in Spanish. For instance, in place of the subordinate clause in sentence (1), English would use the infinitival construction *Rosa wants Carmen to come-INFINITIVE to her house*. Therefore, the volitional subjunctive is a crosslinguistic difference between English and Spanish at both the syntactic and morphological levels. As a result, HSs must acquire and maintain the syntax, semantics, and morphology of the Spanish volitional subjunctive, which emerges late even in monolingual children, on the basis of less input than monolinguals and in the face of crosslinguistic influence from English. Putnam and Sánchez's (2013) approach would predict that decreasing HL activation would lead to increased difficulty mapping subjunctive inflections onto subordinate verbs, especially those that are infrequent and particularly in production. Ultimately, this could cause the complete reassembly of the uninterpretable mood feature at the representational level.

3.1. Monolingual Acquisition

Monolingual Spanish-speaking children begin producing subjunctive inflections before age three, but studies show that monolingual children between ages four and five have not yet reached categorical subjunctive production (Aguirre, 2003; Blake, 1983; Dracos et al., 2019; López-Ornat et al., 1984). The subjunctive emerges in volitional clauses before other syntactic contexts, yet later than other verbal inflections (Mariscal, 2009; Pérez-Leroux, 1998; Rodríguez-Mondoñedo,

2008). Since even monolingual children require extensive exposure to master the subjunctive, more so than other structures, exposure may be particularly impactful when HSs' input spaces are divided between two (or more) languages.

3.2. Heritage Language Acquisition of Subjunctive

The available longitudinal data on English-Spanish bilingual children underscore that exposure influences this group's acquisition of the subjunctive. Two case studies on English-Spanish simultaneous bilingual siblings with different quantities of HL input reveal attrition of subjunctive around the start of schooling by the sibling with less exposure, likely due to increased input in English through school (Anderson, 2001; Silva-Corvalán, 2014). Similarly, Merino (1983) reported that simultaneous bilingual children in an English-only school showed decreased subjunctive production over two years.

Cross-sectional work on older children has also provided evidence for exposure effects. Dracos and Requena (2022) found that proficiency and frequency of exposure, but not age, affected volitional subjunctive production, but this study lacked a large number of older children with which to plot development into adolescence. Potowski (2007b) also found that adolescent HSs in a DLI program produced less subjunctive across six syntactic contexts than Spanish-dominant peers. Finally, Flores et al. (2017) found that German-dominant HSs of Portuguese who had higher exposure to their HL showed growth in subjunctive mood production between ages 8 and 12, while children with less exposure achieved comparable growth after age 13 only. The latter study suggests a positive relationship between age and exposure. While these findings reinforce the central role of exposure in Putnam and Sánchez's (2013) approach, they imply that exposure modulates the rate of protracted development rather than the rate of increasing optionality and feature reassembly. Across these studies, there is not yet an experiment comparing productive and

receptive knowledge in bilingual children, nor one that controls for both home and educational exposure. Furthermore, exploring lexical frequency may yield more nuanced results that can explain variability at the within-speaker level.

Like children, adult HSs are more precise in their command of the subjunctive at higher proficiency levels and with more frequent HL use (Giancaspro, 2019b; Montrul, 2009; Perez-Cortes, 2016). Furthermore, factors such as the morphological regularity and lexical frequency of the subordinate verb affect adult HSs' production and interpretation of subjunctive mood (Giancaspro, 2020; Giancaspro et al., 2022; Perez-Cortes, 2022), but neither has been explored with bilingual children. In fact, previous studies within Spanish as a HL have revealed frequency effects with adults (AUTHOR, XXXX; Giancaspro, 2020; Hur, 2020; Hur et al., 2020; López Otero, 2023; Perez-Cortes, 2022). Only previous study has explored frequency effects with child HSs, but did so with only two verbs (Goldin et al., 2023). Therefore, exploring the influence of lexical frequency with bilingual children not only provides a nuanced account of subjunctive development, but also has broader implications for HL acquisition.

4. The Study

To contribute to the growing body of research that emphasizes the importance of HL exposure, a study that measures proficiency, use, *and* method of schooling with multiple age groups of HSs and that incorporates productive and receptive tasks testing lexical frequency is necessary. Based on the research reviewed thus far, five research questions (RQs) were proposed:

1. How do HSs in a DLI school compare age-matched HSs without bilingual education in the production and selection of the volitional subjunctive?

Putnam and Sánchez's (2013) framework would predict that HSs with more-frequent exposure to and processing of Spanish would show greater consistency in mapping the subjunctive onto lexical

items. Therefore, it was predicted that HSs in DLI who have greater exposure to Spanish would produce and select more volitional subjunctive than age-matched peers in an English-only school.

2. Do proficiency in and frequency of use of Spanish affect individual HSs' rates of volitional subjunctive production and selection?

Putnam and Sánchez's (2013) approach to HL acquisition would posit that proficiency and frequency of use affect HSs' subjunctive mood knowledge. Previous studies on child and adult HSs have found that both variables are predictive of rates of subjunctive production, acceptability judgment, and interpretation (Dracos & Requena, 2019; Flores et al., 2017; Montrul, 2009; Montrul & Perpiñán, 2011; Perez-Cortes, 2016). Therefore, it was predicted that HSs who obtain higher proficiency scores and who report using Spanish more frequently would produce and select the volitional subjunctive more often.

3. Do older HSs show increased production and selection of the volitional subjunctive?

Since exposure to the HL decreases as children become older due to the prevalence of English in school, Putnam and Sánchez (2013) would predict that school-aged HSs would experience greater difficulty mapping features onto inflectional morphology as they get older. While subjunctive loss has been found in HSs in early childhood (Anderson, 2001; Merino, 1983; Silva-Corvalán, 2014), previous studies have lacked a substantial group of late-childhood and adolescent participants. However, the closest available data with older HSs of Portuguese from Flores et al.'s (2017) study show protracted development of subjunctive mood, rather than attrition. Therefore, it was predicted that the older participants (7th/8th grade; ages 12-14) would produce the volitional subjunctive and select it more frequently than younger participants (5th grade; ages 10-11).

4. Do HS children exhibit asymmetries between productive and receptive knowledge of volitional subjunctive?

Putnam and Sánchez's (2013) framework argues that individuals with less-frequent use of Spanish possess stronger receptive knowledge than they exhibit in production. Previous research on the subjunctive supports this position (Perez-Cortes et al., 2019). Therefore, it was predicted that HSs in the DLI program, as well as those who reported frequent use of Spanish overall, would not show differences between rates of volitional subjunctive production and selection. In contrast, those who reported lower activation and who were in an English-only school would select more volitional subjunctive than they would produce it.

5. Does the lexical frequency of the subordinate verb affect child HSs' production and selection of volitional subjunctive?

Putnam and Sánchez (2013) argue that HSs are more likely to produce and select inflectional morphology on more-frequent lexical items. Furthermore, research on adult HSs has found that the lexical frequency of the subordinate verb affects productive and receptive knowledge (Giancaspro, 2020). For this reason, it was predicted that HSs would produce the volitional subjunctive and select it more frequently when the subordinate verb onto which the mood feature is mapped is more frequent.

4.1. Participants

57 English-dominant Spanish HS children completed this experiment. Four groups of HS children participated: HSs in DLI in 5th grade (DLI-5; $n = 19$) and 7th/8th grades (DLI-7/8; $n = 13$), as well as HSs in a monolingual English school in 5th grade (MLE-5; $n = 14$) and 7th/8th grades (MLE-7/8; $n = 11$). Participants in 5th grade were 10-11 years old, and those in the 7th/8th grade group were 12-14 years old. The comparison of HSs to one another can isolate how certain factors related to

language experience, such as bilingual schooling and age, affect the acquisition of Spanish morphosyntax, and moves away from conventional HS-versus-monolingual comparisons (Rothman et al., 2023).

School	5th	7th/8th	Total by school
DLI	19	13	**32**
Monolingual English	14	11	**25**
Total by grade group	**33**	**24**	**57**

Table 1. Division of participants by grade and school.

In the DLI school, HSs received 50% of instruction in Spanish each day from kindergarten through fifth grade. After sixth grade, they continued to receive a daily Spanish for HSs course at the same school, but no longer received content courses in Spanish. To participate in this study, children needed to have attended the DLI school for at least half of elementary school (second grade or earlier). The 5th graders represent the "finished product" of elementary DLI programs; the 7th/8th grade groups address the absence of bilingual children of this age range in previous studies. The age-matched peers in the MLE school had not received any bilingual or world language instruction during their schooling. Both schools' demographic reports indicated matching percentages (within 1%) of Latinx families and students with low socioeconomic status (measured through family income). According to questionnaire data, while some children's parents spoke English, all spoke Spanish, and all participants were predominantly exposed to Spanish at home (see Table 2). Children in both schools were primarily of Mexican descent, although there were also participants who spoke Dominican and additional Central American varieties of Spanish.

18 Spanish-dominant bilingual adults (SDBAs; average age: 33.6, SD = 10.1) also completed this experiment. These adults were included only to norm the tasks and confirm that they elicited the expected subjunctive morphology as anticipated. These bilinguals were from seven Spanish-speaking countries and averaged 9.5 years in the United States. They primarily

worked and resided in the same region where the two schools were located, and most were students at a large regional university in the same town as the monolingual school. These bilinguals' data are presented in Figures 1 and 2 to verify that they used the subjunctive categorically in this experiment, but they were not submitted for further analysis since this project concentrates on HSs (see Rothman et al., 2023). Table 2 provides a summary of each group's proficiency using a subset of the Bilingual English-Spanish Assessment (Peña et al., 2014), frequency of use of Spanish (as calculated across six contexts using a language questionnaire described below), and number of monolingual Spanish-speaking parents.

Variable	SDBA (*n* = 18)		DLI-7/8 (*n* = 11)		MLE-7/8 (*n* = 13)		DLI-5 (*n* = 19)		MLE-5 (*n* = 14)	
	μ	SD	μ	SD	μ	SD	μ	SD	μ	SD
Frequency of use of Spanish (max. 30 points)	15.2	5.9	15.7	4.7	14.0	4.6	15.5	6.2	13.7	4.2
Proficiency score (max. 14 points)	12.1	2.0	12.0	2.3	11.5	2.2	9.3	3.3	9.0	2.4
Number of monolingual Spanish-speaking parents	1.9	0.2	1.0	0.9	1.4	0.8	0.9	0.9	1.4	0.9

Table 2. Participant group averages with standard deviations.

4.2. Method and Tasks

All tasks were administered using Qualtrics software. A software known as Phonic was embedded into Qualtrics to record participants' responses on the production task. Children carried out the experiment in their schools with the researcher present; the SDBAs carried out the study asynchronously online.

4.2.1 Language background tasks. Before completing the experiment, participants completed a language questionnaire. Since the youngest children in this study were 10 years old, it was determined that these individuals would be able to complete this questionnaire independently, which provides a more direct measurement of their Spanish usage away from home. This approach has been shown to be effective with high psychometric validity with children

younger than those in the present study (Castilla-Earls et al., 2022). The questionnaire asked participants to indicate when they began attending their school and to rate how frequently they used Spanish in six contexts: with parents, with family members, at school, with friends, in public, and while watching television. Participants indicated how often they used Spanish in each context using 1-5 Likert scales with fixed descriptors (0: never, 1: hardly ever, 2: 3-4 times per week, 3: 1-2 times per day, 4: every day/almost always, 5: always). The sum of these Likert scales comprised each participant's 30-point overall frequency of use score.

Subsequently, participants completed part of the Bilingual English-Spanish Assessment (BESA; Peña et al., 2014) to measure Spanish morphosyntactic proficiency. This task included eighteen stimuli testing inflectional morphology, although four questions were not reported here because they evaluated subjunctive mood, whose inclusion would be tautologous. The remaining questions targeted gender/number agreement with articles ($k = 4$), verbal agreement ($k = 4$), preterit aspect morphology ($k = 2$), and clitic gender ($k = 4$). The number of correct responses comprised each participant's proficiency score. Participants responded to each prompt by writing a full sentence.

4.2.2. Experimental tasks. In addition to the questionnaire and proficiency test, there were two tasks centered around the same communicative context, in which a mother shared how she wanted her twin daughters to care for their younger brother, Juanito, while away at sleepaway camp. The same eight verbs listed in Table 3 were used once per task. All verbs pertained to the most-frequent first conjugation class (*–ar*) and were disyllabic and transitive.[3] The lemmatic

---

[3] Recent research has shown that HSs are more likely to produce the subjunctive with irregular verbs (Giancaspro et al., 2022); the use of regular verbs is addressed as a limitation in the discussion of results in this paper.

frequency[4] of each verb was retrieved from the Davies (2016) *Corpus del español,* a large database used in previous research on the subjunctive (e.g., Giancaspro, 2020; Perez-Cortes, 2022). Since this corpus has been shown to correlate with HSs' self-reported lexical frequency ratings (AUTHOR, XXXX), it was determined to be an appropriate source for analyzing the impact of lexical frequency on participants' mood selection tendencies. The lemmatic frequencies for each verb also appear in Table 3.

Spanish	English	Davies (2016) frequency
amar	to love	60,260
cuidar	to care for	208,649
llamar	to call	303,936
llevar	to take/carry	1,192,620
mirar	to look at	186,441
peinar	to comb	6,151
pintar	to paint	45,606
tratar	to treat	590,952

Table 3. List of verbs used across tasks with their lexical frequency.

The production task contained a trial and 16 experimental items and followed a sentence completion format. There was a brief written prompt, and participants needed to complete the final sentence orally using the verb in parentheses, as well as any other necessary words. 8 items targeted the subjunctive mood following the matrix verb *querer* ('to want'), as in (2). In addition, 4 stimuli tested knowledge of the indicative mood following the matrix verb *creer* ('to believe'), and 4 were fillers. Only the eight subjunctive items were analyzed, but the additional sentences were included to prevent responses from becoming predictable.

(2) *A veces Juanito se pone triste si sus hermanas dicen que no quieren hablar con él. ¿Qué quiere la mamá? Quiere que las hermanas _____ (LLAMAR) Juanito cada noche.*[5]

---

[4] Vannest et al. (2011, p. 2) define lemmatic/base frequency as "the frequency of the base morpheme, equivalent to the total frequency of all the words containing this morpheme." The term *lexical frequency* is interchangeable with lemmatic frequency in this paper.

[5] Note that in these sentences, the differential object marker *a* was omitted intentionally.

Sometimes Juanito gets sad if his sisters say that they don't want to talk with him. What does the mother want? She wants that the sisters _____ (CALL) Juanito every night.

For the selection task, participants again read prompts related to the mother's desires for her children. There was a total of 23 items in the task. As in the production task, 8 items targeted the subjunctive in volitional clauses following *querer*, as in (3), 3 targeted the indicative mood with *creer*, and the remaining 12 were fillers. Only the subjunctive items were analyzed. In this task, participants read two sentences that described the prompt, and needed to select which one looked best. These sentences differed only in the use of the subjunctive or indicative. Therefore, this task tapped participants' underlying knowledge of mood, as it required that they discern between the indicative and subjunctive without requiring HL production.

(3) *La mamá sabe que las hermanas no siempre hablan con Juanito. ¿Qué quiere la mamá?*
      a. **Quiere que las hermanas lo miran cuando hablan con él.*
      b. *Quiere que las hermanas lo miren cuando hablan con él.*

The mother knows that the sisters don't always talk with Juanito. What does she want?
      a. She wants the sisters to look at (*indicative) him when they talk with him.
      b. She wants the sisters to look at (subjunctive) him when they talk to him.

5. Results

5.1. Analysis

After administering the tasks, all data were manually coded, and a binary score was assigned to each response based upon the suppliance of the anticipated mood. Only the stimuli testing the subjunctive were included for analysis. All instances in which participants produced or selected the expected subjunctive form received a score of *1*, and all other forms were assigned a score of *0*. In the production task, forms with correct mood inflections but with tense or agreement morphology that did not match the prompt were accepted with a score of *1* because this study did not concentrate on these morphological categories. This comprised the binary dependent variable for both the descriptive and inferential statistics.

There were 30 instances in the HSs' data from the production task where there were errors with participants' audio recording or in which responses were not related to the stimuli. These observations were discarded, leaving 426/456 (93.4%) available for analysis. The children produced the subjunctive in a total of 193/426 instances (45.3%), and alternative forms in the remaining 233 instances (54.6%). While indicative substitution was by far the most frequent alternative, followed by infinitival forms, there were also alternative structures used in the HSs' data, as summarized in Table 4.

Structure	Observations
Indicative mood	177 (75.9%)
Infinitive	35 (15.0%)
Ir a (*going to*) periphrastic future	13 (5.5%)
Uncategorizable forms	4 (1.7%)
Modal verbs in indicative	3 (1.2%)
Preterit	1 (0.4%)

Table 4. List of alternative structures (total 298) to subjunctive mood in HSs' data. [6]

Additional data were coded to generate independent variables. Firstly, each participant's group (DLI-7/8, MLE-7/8, DLI-5, MLE-5, SDBA) was coded to allow for a comparison in the descriptive statistics and individual analyses. Moreover, each HS participant's grade group (5th grade versus 7th/8th) and school (DLI versus MLE) were included as categorical variables so that these two factors could be evaluated independently. Additional data from the BESA proficiency task and language questionnaire were analyzed to generate a pair of continuous variables. Each participant's proficiency score was calculated as the number of expected answers on the BESA morphosyntax subsection (maximum 14; see Table 2), and frequency of use was operationalized by adding together the 1-5 Likert scales for the six contexts of language use from the questionnaire

---

[6] "Uncategorizable forms" were innovative inflections that are not a part of the Spanish inflectional system, and therefore, could not be placed into named categories.

(maximum 30; see Table 2). Finally, lexical frequency was operationalized as the lemmatic frequency from the Davies (2016) corpus for each verb (see Table 3). All data were anonymized and uploaded to a GitHub repository (BLINDED). Data analysis was carried out in RStudio (R Core Team, 2022) using the *lme4* (Bates et al., 2015), *lmerTest* (Kuznetsova et al., 2017), and *tidyverse* (Wickham et al., 2019) packages.

5.2. Descriptive Analysis

HSs' and SDBAs' rates of subjunctive production and selection are summarized by group in Figures 1 and 2 below. As anticipated, the SDBAs used the subjunctive invariably, confirming that the tasks elicited subjunctive morphology as expected. Since the HSs were the focal point of the present project, the SDBAs' data are not analyzed further. These figures also show that the 7th/8th grade groups produced and selected more subjunctive than the 5th grade groups, implying that children develop knowledge of the volitional subjunctive with age. However, there were no discernible differences between children in the DLI and English-only schools. All HS groups selected the subjunctive more frequently on the selection task than they produced this structure.
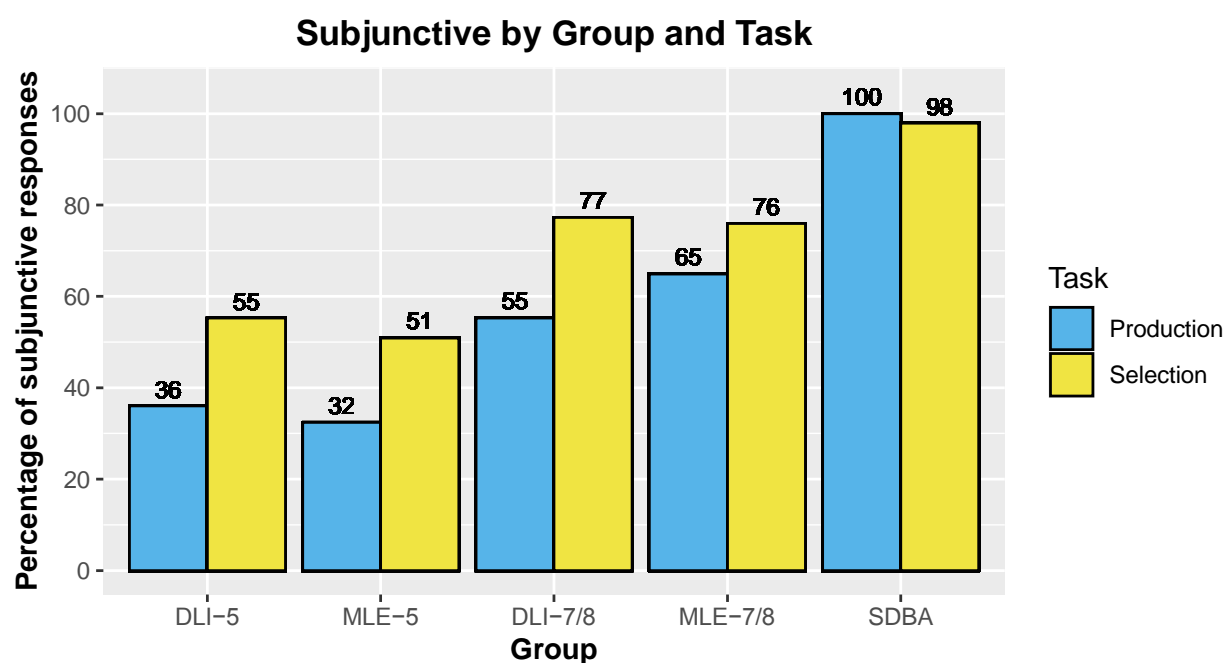

Subjunctive by Group and Task

Percentage of subjunctive responses by Group (DLI−5, MLE−5, DLI−7/8, MLE−7/8, SDBA) and Task (Production, Selection):

Group	Production	Selection
DLI−5	36	55
MLE−5	32	51
DLI−7/8	55	77
MLE−7/8	65	76
SDBA	100	98

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

19

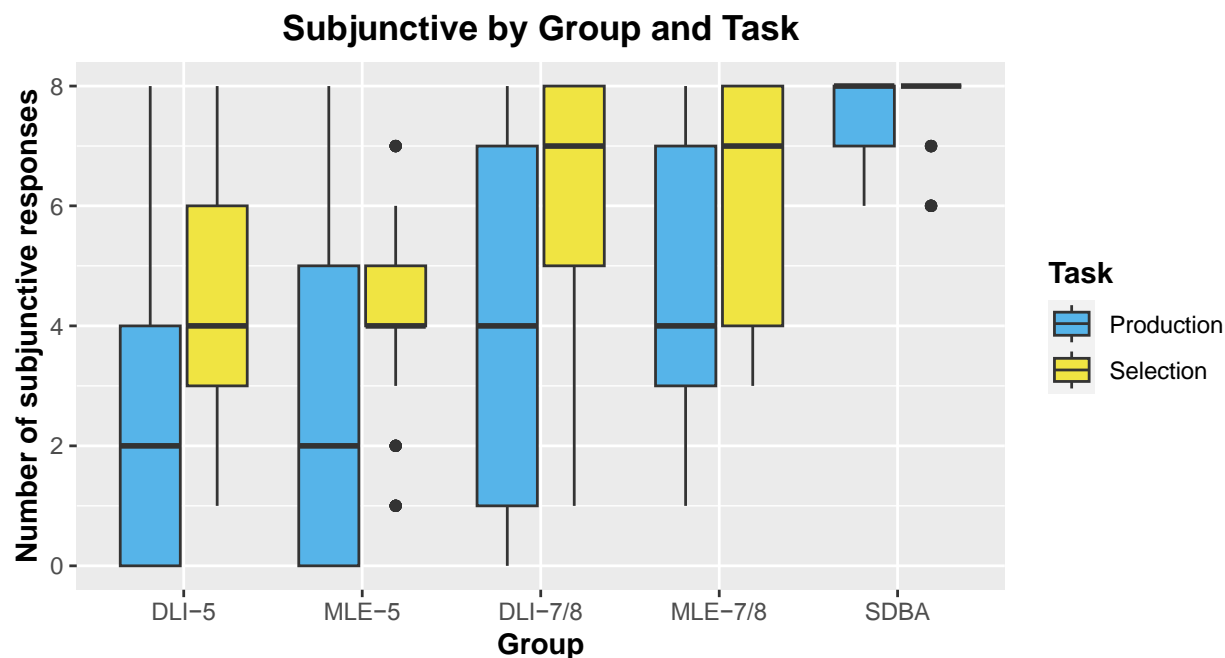**Figure 1.** Percentages of subjunctive production and selection by group and task.

Figure 2. Summary of number of subjunctive sentences produced or selected by group and task.

Figure 3 summarizes HSs' production and selection of volitional subjunctive by BESA proficiency, Figure 4 summarizes production and selection by frequency of use of Spanish, and Figure 5 outlines the role of lexical frequency on subjunctive production and selection. These figures point towards modest roles for proficiency and frequency of use of Spanish, but argue against any lexical frequency effects.
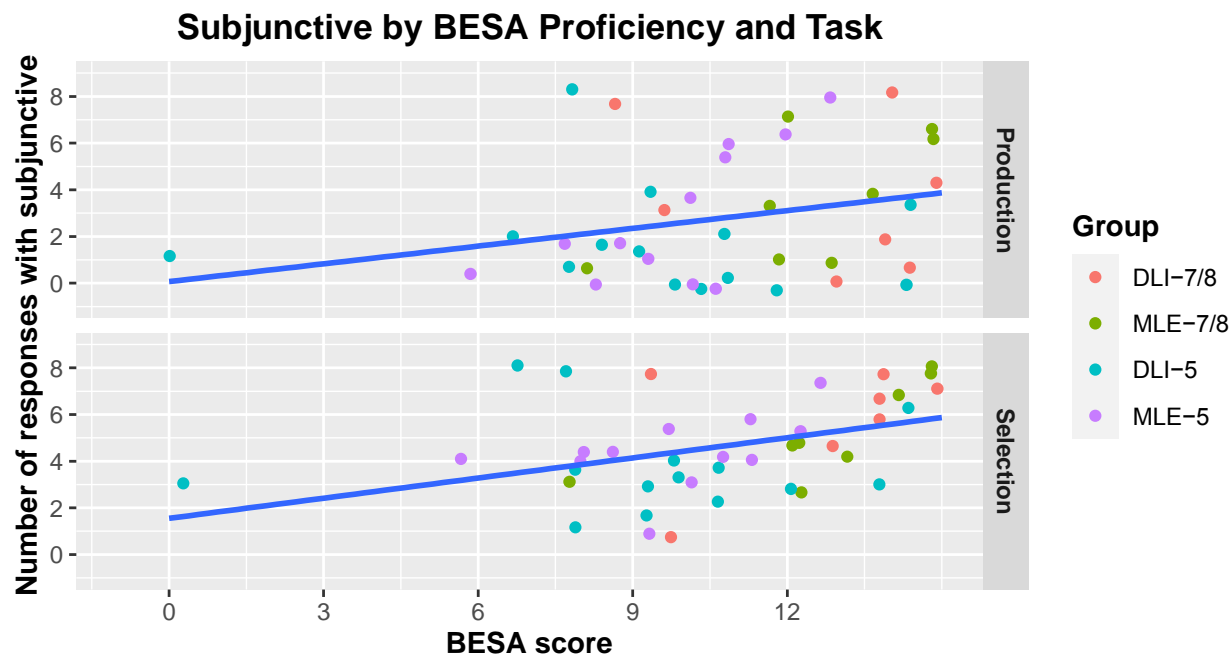
Figure 3. Production and selection of volitional subjunctive by BESA proficiency.
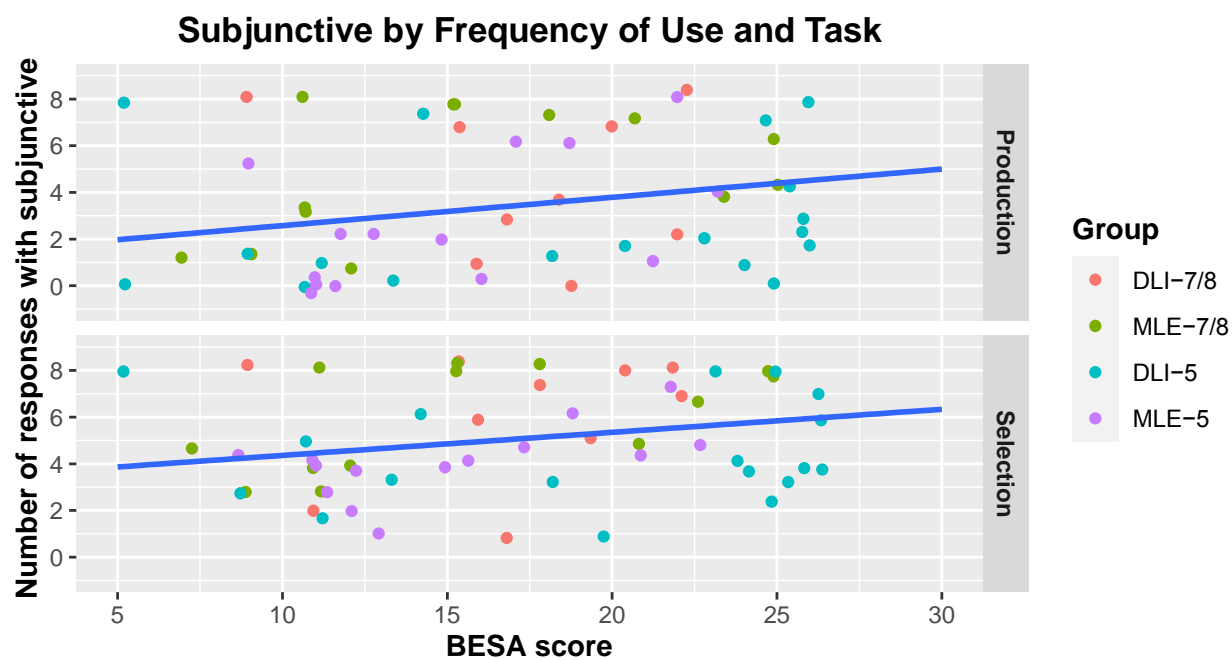

Figure 4. Production and selection of volitional subjunctive by frequency of use of Spanish.
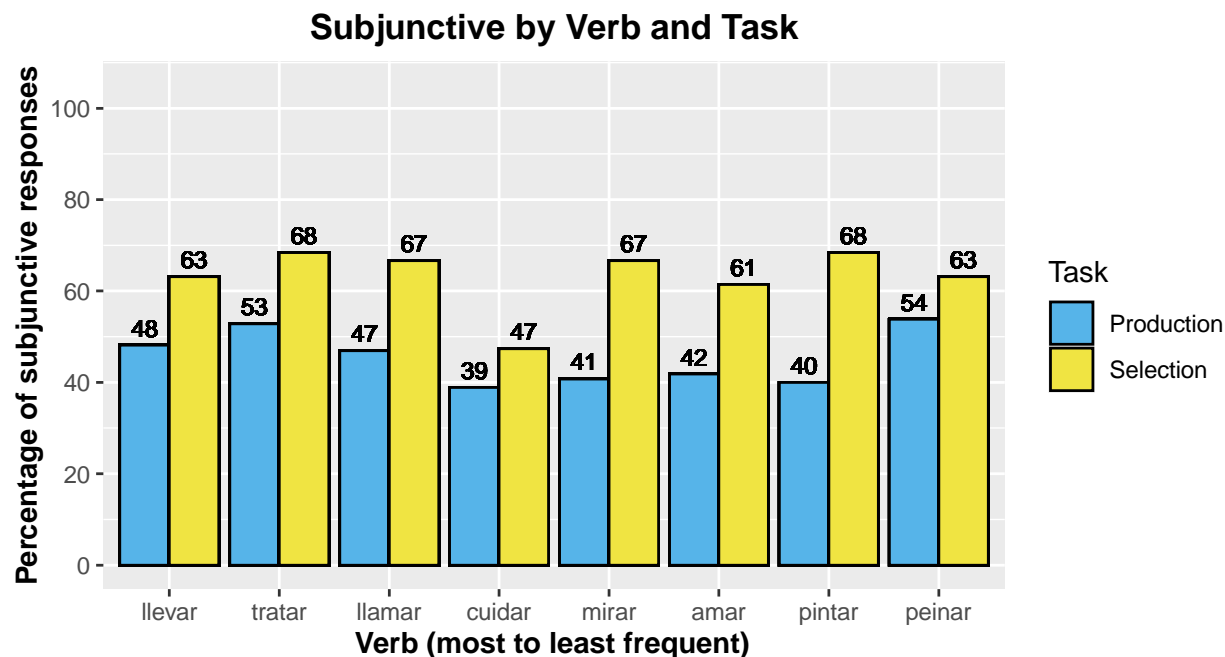
Figure 5. Production and selection of volitional subjunctive by verb, organized from most to least frequent as rated in the Davies (2016) corpus.

5.3. Individual Analyses

Additionally, individual analyses were conducted by calculating the total number of subjunctive forms that each participant produced and selected. Figure 6 illustrates that all HSs produced or selected the subjunctive in at least two contexts. This argues against the altogether absence of this structure from HSs' grammar, and instead that these bilinguals map this feature onto morphology variably, as feature-oriented accounts predict (Lohndal & Putnam, 2021; Putnam & Sánchez, 2013).
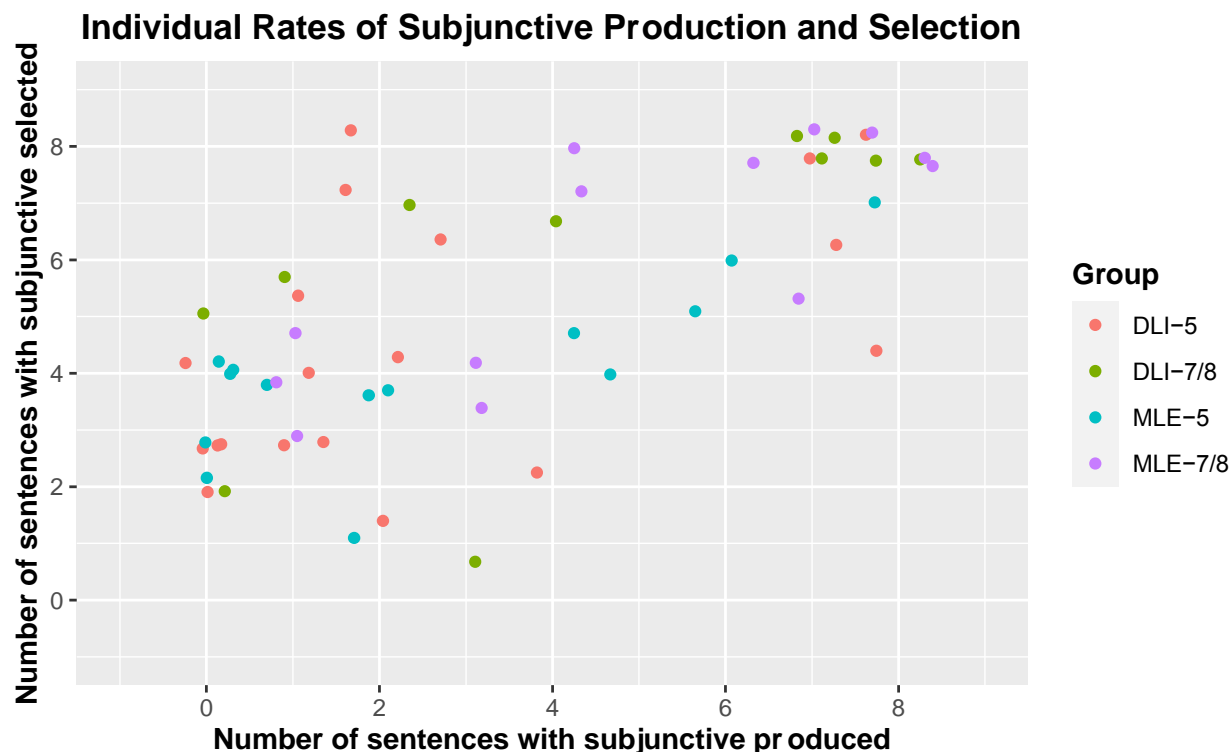
Figure 6. Individual rates of production and selection of subjunctive by participant.[7]

Six participants produced and selected the subjunctive in all 16 contexts. These participants' characteristics are listed in Table 5. That 5/6 of the participants were in the 7th/8th grade group further implies an effect for age, but there were no clear trends involving school, proficiency, frequency of use of Spanish, or parental bilingualism. In fact, three of these participants, one of whom attended the DLI school, reported never using Spanish at school.[8]

Part.	Group	Freq. of use	School use	BESA	Parental languages
H5B10	DLI-5	5/25	0/5	7/14	1 Spanish only, 1 bilingual

[7] The *geom_jitter* layer was applied to this graph to slightly displace points along the *x* and *y* axes to enhance visibility of each participant's data. Alternatively, if there were multiple participants with the same production and selection rates, only one point would be visible on the graph. Therefore, each of these points should be approximated to the nearest available integer.

[8] This may appear counterintuitive; however, students' frequent use of English during Spanish instructional time is well-documented (Ballinger & Lyster, 2011; Hamman, 2018; Potowski, 2004).

H7B03	DLI-7/8	8/25	1/5	8/14	1 Spanish only, 1 bilingual
H8B04	DLI-7/8	19/25	3/5	11/14	Both Spanish only
H8M02	MLE-7/8	11/25	0/5	14/14	1 Spanish only, 1 bilingual
H8M06	MLE-7/8	15/25	0/5	14/14	Both Spanish only
H8M13	MLE-7/8	13/25	2/5	12/14	Both bilingual

Table 5. Characteristics of HS children who produced and selected subjunctive mood at ceiling.[9]

In comparison, there were three children, whose characteristics are summarized in Table 6, who selected the subjunctive in two instances and did not produce it. All three of these children were in the 5th grade age group. Furthermore, these participants reported using Spanish outside of school in less than half of their interactions, which also aligns with the prediction that low exposure to Spanish modulates subjunctive production and selection.

Part.	Group	Freq. of use	School use	BESA	Parental languages
H5B18	DLI-5	9/25	2/5	9/14	Both Spanish only
H5M01	MLE-5	10/25	2/5	14/14	Both Spanish only
H5M02	MLE-5	12/25	1/5	7/14	Both Spanish only

Table 6. Characteristics of HS children who produced and selected the fewest instances of subjunctive mood.

5.4. Inferential Statistics

To further evaluate these data, a Generalized Linear Mixed Methods (GLMM) binomial logistic regression model was necessary. The binary dependent variable was the suppliance of subjunctive mood as described in section 5.1. HSs' data from both the production and selection task were incorporated in the same model. The predictors for the GLMM model were school group (DLI versus MLE), age group (5th grade versus 7th/8th grade), BESA proficiency score, frequency of use, task (production versus selection), lexical frequency, and three two-way interactions: school group by age group, school group by task, and frequency of use by task. The latter three variables

---

[9] In Tables 5 and 6, "freq. of use" refers to the five contexts of language use targeted on the language questionnaire except for "at school," which is represented separately to the right.

(proficiency, frequency of use, and lexical frequency) were continuous, so they were standardized to refit them along the same scale in the statistical modeling.

Given there were numerous predictors and interactions necessary to address the RQs, nested model comparisons were carried out through pairwise comparisons to determine the model of best fit. Pairwise comparisons were conducted manually by adding one predictor into each model and comparing these models using the *anova* function in RStudio. In the pairwise comparisons, the model with random intercepts for participant and item and fixed effects for school, grade group, proficiency, frequency of use, and task accounted for the greatest amount of variance (AIC = 949.18, χ^2 = 17.87, p < .0001). Finally, this model, which contained random intercepts only, was compared to a similar model with the same fixed effects but with participant and item as random intercepts and random slopes, following Cunnings (2012). The second model failed to converge, so the first model containing participant and item as random intercepts only was retained.

The final GLMM model (AIC = 949.2) contained school, grade group, frequency of use, BESA proficiency, and task as predictors with participant and item as random intercepts. The 5th grade group and the production task were set as reference levels. The model revealed main effects for the 7th/8th grade group (β = 1.47, SE = 0.46, p = .0013), frequency of use (β = 0.52, SE = 0.23, p = .0220), and the selection task (β = 1.13, SE = 0.20, p < .0001). These results imply that older HSs show stronger knowledge of subjunctive mood, that frequency of use modulates subjunctive knowledge, and that participants were more likely to select the subjunctive than they were to produce it. These findings are consistent with the data reported in Figures 1, 2, and 4.

6. Discussion

The present study is the first to compare productive and receptive knowledge of subjunctive mood across multiple age groups of school-aged Spanish HSs, and adopts a novel approach by evaluating HL exposure through different methods of schooling. Moreover, it evaluated lexical frequency effects with bilingual children, which is largely novel (but see Goldin et al., 2023 for an analysis with two lexical items). There were five RQs, the first of which investigated whether HSs in a DLI program differed from other HSs in an English-only school in their command of the volitional subjunctive. It was predicted that participants in the DLI school would recognize and produce this structure more consistently than monolingually-educated peers. Contra the predictions, there were no differences observed between the HSs in the DLI program and those in the English-only school. On the surface, this finding does not support theories of language acquisition that posit a central role for exposure and activation (e.g., Putnam & Sánchez, 2013), a result that merits further discussion below.

The second RQ evaluated whether proficiency and frequency of use would affect rates of subjunctive production and selection. It was predicted that both variables would affect the production and selection of the volitional subjunctive, although only frequency of use accounted for speakers' mood tendencies. Therefore, participants' overall patterns of language use predicted their command of the volitional subjunctive, but proficiency did not, which partially supports the hypothesis. The finding that frequency of use was predictive of individual differences in subjunctive production and selection aligns with Putnam and Sánchez's (2013) model.

The third RQ evaluated the role of age in the HL acquisition of the volitional subjunctive. It was predicted that older children would have greater volitional subjunctive production and selection, which is consistent with results. These findings align with a growing number of studies on other areas of the Spanish inflectional system revealing protracted HL development in

childhood (i.e., AUTHOR, XXXX, XXXX; Corbet & Domínguez, 2020; Cuza & Miller, 2015; Cuza & Solano-Escobar, 2023; Martinez Nieto & Restrepo, 2022; Montrul & Potowski, 2007; Montrul & Sánchez-Walker, 2013; see also Flores et al., 2017 concerning the Portuguese subjunctive).

The fourth RQ evaluated whether HS children would show asymmetrical productive and receptive knowledge of the volitional subjunctive. Putnam and Sánchez's (2013) theory posits stronger receptive than productive knowledge for speakers with less-frequent exposure, so it was predicted that HSs with lower reported use of Spanish and who attended the English-only school would show stronger performance on the selection task than in production. However, HSs as a whole selected the subjunctive more frequently than they produced it, regardless of exposure patterns. These task asymmetries are also consistent with Perez-Cortes' (2016) study on the subjunctive mood; however, unlike in Perez-Cortes (2016), frequency of use did not modulate these asymmetries.

Finally, RQ5 addressed the role of lexical frequency in the production and selection of subjunctive mood in volitional clauses. Based upon previous studies on the subjunctive, as well as on adult and child HSs, it was predicted that participants in this study would produce and select the subjunctive more consistently with more-frequent lexical items. However, this result was not upheld, and there was no discernible role of lexical frequency. This is contra previous research as well as Putnam and Sánchez's (2013) predictions.

Together, these findings partially support Putnam and Sánchez's (2013) account of HL acquisition and maintenance in the context of child HL acquisition. The effects for frequency of use of Spanish and for asymmetries between production and selection support the central tenet of these researchers' framework, in that exposure modulates HL acquisition. However, the positive

role of age suggests that HSs do not reassemble the mood feature or experience greater difficulty mapping it onto morphological forms over time due to increased exposure to English. Rather, the data here point towards the protracted development of subjunctive mood. This finding is particularly surprising when considering that the DLI-7/8 group had experienced a drop in exposure to Spanish at school, yet still produced and selected more subjunctive mood when compared to the DLI-5 group that still received immersion in Spanish. This is inconsistent with what has been documented with younger HSs (Anderson, 2001; Merino, 1983; Silva-Corvalán, 2014). These findings do not preclude early attrition and subsequent reacquisition of subjunctive mood, but testing this possibility would require younger participants who may have found these tasks to be unwieldly lengthy. Furthermore, there was no role of lexical frequency in the present study, which does not align with Putnam and Sánchez's (2013) predictions.

Considering these results, one possibility is that Putnam and Sánchez's (2013) predictions run in reverse in the initial yet protracted development of Spanish morphosyntax. Specifically, HSs may develop stronger receptive knowledge before they consistently map syntactic forms onto morphology in language production. This differs subtly from Putnam and Sánchez's (2013) predictions, as these researchers argue that HSs begin to experience optionality in production, which then extends to underlying syntactic knowledge. The rate of child HSs' development is likely modulated by patterns of exposure, as found in the present study and by Flores et al. (2017) in Portuguese. Recent research on bilingual children's and adults' acquisition of the differential object marker in Spanish (AUTHOR, XXXX, XXXX) has also advanced similar claims. Such a finding would align with Putnam et al. (2018), who assert that language dominance affects the degree of crosslinguistic influence in HL development. Individuals who have greater exposure consequently experience less crosslinguistic influence from English, so they may map the

subjunctive mood feature onto subordinate verbs earlier and with greater consistency than children with lower exposure. Studies pointing towards protracted HL development continue to amass, so this explanation provides a possibly promising account for these findings. More research comparing children's productive *and* receptive knowledge across a broader age range would be crucial to test this possibility.

This explanation does not incorporate lexical frequency, which is an important component of Putnam and Sánchez's (2013) model. The present study differs from previous work on adult HSs that has reported frequency effects in HSs' command of subjunctive mood, grammatical gender, differential object marking, preterit aspect, verbal agreement morphology, and imperative commands (e.g., AUTHOR, XXXX; Giancaspro, 2020; Goldin et al., 2023; Hur, 2020; Hur et al., 2020; López Otero, 2023; Perez-Cortes, 2022). It is difficult to interpret why lexical frequency did not affect participants' mood selection in the present study, but one possibility is that a subset of only eight verbs is not sufficient to observe this effect. Therefore, future studies would be well advised to incorporate more lexical items; however, doing so makes the experimental tasks lengthy, which is often problematic in child language acquisition research.

In addition, it should be noted that while Putnam and Sánchez's (2013) framework correctly predicted differences between individual HSs, the groups as a whole exhibited optionality in both tasks. This supports theories of HL acquisition that argue that as a group, HSs develop morphosyntactic systems that exhibit greater variability than would be found across speakers dominant in the same language (i.e., Montrul, 2008, 2009, 2013). However, the individual analyses show that all HSs produced and recognized the subjunctive in at least two contexts, which suggests that this structure is not completely absent from any speaker's grammar. Rather, most HSs

experience persistently variable form-function mappings, which Lohndal and Putnam (2021) argue is typical in HL acquisition.

Before addressing limitations, it is important to contemplate the implications of the absence of the predicted effect for DLI on subjunctive mood knowledge. A widespread assumption is that DLI assures continued HL acquisition (e.g., Lindholm-Leary and Genesee, 2014, p. 169; Potowski, 2007a, p. 188), but this effect is not observed in this study. It is likely that factors beyond input quantity are necessary to support HL acquisition and maintenance. Firstly, it is likely that output is equally crucial, which aligns with recent research that has found that rates of production affect HS children's morphosyntactic development (Goldin, 2021; AUTHOR, XXXX). This is particularly plausible given that some DLI students reported never using Spanish at school, a finding that is consistent with previous research that has shown that children in bilingual schools frequently avoid Spanish even when receiving instruction in that language (Ballinger & Lyster, 2011; Hamman, 2018; Potowski, 2004). Moreover, Montrul and Perpiñán (2011) argue that explicit instruction provided advanced proficiency adult second language learners with an advantage in subjunctive knowledge over proficiency-matched HSs. Together, these findings along with those of the present study imply that for HSs, simply overhearing Spanish is insufficient for acquiring less-frequent structures such as the subjunctive, even in the case of DLI. Rather, output is a necessary ingredient of HL acquisition and maintenance, as in second language acquisition (Swain, 1993), and explicit instruction may be necessary to acquire some structures in bilingual contexts. The DLI environment could be adapted to be more conducive to output and explicit instruction.

An additional consideration is the role of input providers in DLI programs. Teachers in some immersion programs (but not those with whom the DLI students in this study worked) are HSs or learners of Spanish themselves. Since these populations have been demonstrated to show

variable knowledge of subjunctive mood, as reviewed previously, they could transmit a qualitatively different mood system to bilingual children. In future research on HL acquisition in DLI schools, it would be beneficial for teachers to complete the same tasks as a point of comparison.

Before concluding, it is critical to recognize key limitations. Firstly, the present study only used only canonical –ar verbs, yet Giancaspro et al. (2022) found that adult HSs are more likely to use the subjunctive with verbs with irregular subjunctive stems. While it is prudent to control morphological regularity for methodological reasons, such a decision may have undervalued HSs' production and selection of the subjunctive in this study. Furthermore, while not the focal point of analysis, the strong intensional verb *creer* (*to believe*) could still imply contexts of doubt and consequently may have also inadvertently triggered subjunctive use in unexpected contexts.

Additionally, it may have been helpful to incorporate a group of bilingual children dominant in Spanish, as well as other age groups of bilingual children, to most fruitfully compare these populations. Another possibility is that the present study did not fully control for all possible contexts of HL use; this is particularly plausible considering the brevity of the language questionnaire. Most importantly, this study had a limited sample of bilingual children from only two schools, which required grouping seventh and eighth grade students into a single category. The absence of a meaningful effect for school type based upon observations in a single DLI program need not be taken as an argument against the effectiveness of bilingual education.

7. Conclusion

The present study concentrated on Spanish HSs' acquisition of the subjunctive mood in volitional clauses using production and selection tasks. The finding that children improve in their command of subjunctive knowledge with age during the secondary school years and into adolescence is novel

in research on the acquisition of this structure. The influence of exposure over subjunctive production and asymmetrical productive versus receptive knowledge that HSs exhibited align with Putnam and Sánchez's (2013) activation approach to HL acquisition. However, the finding that DLI did not have an impact on children's acquisition of the Spanish subjunctive raises questions about the role of input and output quantity and quality in HL acquisition. Moreover, lexical frequency did not align with findings, contra Putnam and Sánchez's (2013) predictions. Since children's subjunctive knowledge increased with age, there was no evidence of feature reassembly or increasing optionality due to sustained activation of English. Rather, Putnam and Sánchez's (2013) feature-oriented model that emphasizes patterns of exposure may also be relevant in the context of protracted development, a possibility that deserves testing through future studies.

References

** Note that articles containing the author's name have been removed from this section **

Aguirre, C. (2003). Early verb development in one Spanish-speaking child. In D. Bittner (Ed.),

   Development of verb inflection in first language acquisition: A cross-linguistic

   perspective (pp. 1–25). De Gruyter.

Ballinger, S., & Lyster, R. (2011). Student and teacher oral language use in a two-way

   Spanish/English immersion school. *Language Teaching Research*, *15*(3), 289–306.

   https://doi.org/10.1177/1362168811401151

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models

   using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

   https://doi.org/10.18637/jss.v067.i01

Biber, D., Davies, M., Jones, J. K., & Tracy-Ventura, N. (2006). Spoken and written register

   variation in Spanish: A multi-dimensional analysis. *Corpora*, *1*(1), 1–37.

   https://doi.org/10.3366/cor.2006.1.1.1

Blake, R. (1983). Mood selection among Spanish-speaking children, ages 4 to 12. *The Bilingual*

   Language Review, *10*, 21–32.

Bosque, I. (2012). Mood: Indicative vs. subjunctive. In J. I. Hualde, A. Olarrea, & E. O'Rourke

   (Eds.), *The handbook of Hispanic linguistics* (pp. 373–394). John Wiley & Sons, Ltd.

   https://doi.org/10.1002/9781118228098.ch19

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

33

Castilla-Earls, A., Ronderos, J., & Fitton, L. (2022). Can bilingual children self-report their

bilingual experience and proficiency? The Houston Questionnaire. *Journal of Speech,*

Language, and Hearing Research, *65*(10), 3835–3853.

https://doi.org/10.1044/2022_JSLHR-21-00675

Cinque, G., & Rizzi, L. (2012). The cartography of syntactic structures. In B. Heine & H. Narrog

(Eds.), *The Oxford Handbook of Linguistic Analysis* (1st ed., pp. 51–66). Oxford

University Press. https://doi.org/10.1093/oxfordhb/9780199544004.013.0003

Corbet, J., & Domínguez, L. (2020). The comprehension of tense–aspect morphology by Spanish

heritage speakers in the United Kingdom. *Languages*, *5*(46), 1–20.

https://doi.org/10.3390/languages5040046

Cunnings, I. (2012). An overview of mixed-effects statistical models for second language

researchers. *Second Language Research*, *28*(3), 369–382.

https://doi.org/10.1177/0267658312443651

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.), *Issues in Hispanic*

and Lusophone linguistics (Vol. 4, pp. 211–230). John Benjamins.

https://doi.org/10.1075/ihll.4.11cuz

Cuza, A., & Solano-Escobar, L. (2023). Protracted development in child heritage Spanish:

Evidence from inalienable possession. *Second Language Research*, 02676583231202608.

https://doi.org/10.1177/02676583231202608

Davies, M. (2016). *Corpus del español*. Available online. http://www.corpusdelespanol.org/web-

dial/

Dracos, M., & Requena, P. E. (2022). Child heritage speakers' acquisition of the Spanish

    subjunctive in volitional and adverbial clauses. *Language Acquisition*, *30*(1), 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

Fábregas, A. (2014). A guide to subjunctive and modals in Spanish: Questions and analyses.

    Borealis – An International Journal of Hispanic Linguistics, *3*(2), 1–94.

    https://doi.org/10.7557/1.3.2.3064

Faulkner, T. (2021). *A systematic investigation of the Spanish subjunctive: Mood variation in*

    subjunctive clauses [Doctoral dissertation]. Georgetown University.

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

Gathercole, V. C. (2002). Grammatical gender in bilingual and monolingual children: A Spanish

    morphosyntactic distinction. In K. Oller & R. Eilers (Eds.), *Language and literacy in*

    bilingual children (pp. 207–219). Multilingual Matters.

    https://doi.org/10.21832/9781853595721-010

Giancaspro, D. (2019a). Over, under and around: Spanish heritage speakers' production (and

    avoidance) of subjunctive mood. *Heritage Language Journal*, *16*(1), 44–70.

    https://doi.org/10.46538/hlj.16.1.3

Giancaspro, D. (2019b). The late(r) bird gets the verb? Effects of age of acquisition of English

on adult heritage speakers' knowledge of subjunctive mood in Spanish. *Languages*,

4(69), 1–34. https://doi.org/10.3390/languages4030069

Giancaspro, D. (2020). Not in the mood: Frequency effects in heritage speakers' subjunctive

knowledge. In B. Brehmer & J. Treffers-Daller (Eds.), *Studies in bilingualism* (Vol. 59,

pp. 72–97). John Benjamins. https://doi.org/10.1075/sibil.59.03gia

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

Giancaspro, D., & Sánchez, L. (2021). Me, mi, my: Innovation and variability in heritage

speakers' knowledge of inalienable possession. *Glossa: A Journal of General Linguistics*,

6(1). https://doi.org/10.5334/gjgl.1240

Goldin, M. (2021). Language activation in dual language schools: The development of subject-

verb agreement in the English and Spanish of heritage speaker children. *International

Journal of Bilingual Education and Bilingualism*, *25*(8), 1–22.

https://doi.org/10.1080/13670050.2021.2005529

Goldin, M., López Otero, J. C., & Hur, E. (2023). How frequent are these verbs?: An exploration

of lexical frequency in bilingual children's acquisition of subject-verb agreement

morphology. *Isogloss. Open Journal of Romance Linguistics*, *9*(2), 1–25.

https://doi.org/10.5565/rev/isogloss.194

Hamman, L. (2018). Translanguaging and positioning in two-way dual language classrooms: A

case for criticality. *Language and Education*, *32*(1), 21–42.

https://doi.org/10.1080/09500782.2017.1384006

Hiebert, L., & Rojas, R. (2021). A longitudinal study of Spanish language growth and loss in

    young Spanish-English bilingual children. *Journal of Communication Disorders*, *92*, 1–

    15. https://doi.org/10.1016/j.jcomdis.2021.106110

Hur, E. (2020). Verbal lexical frequency and DOM in heritage speakers of Spanish. In A.

    Mardale & S. Montrul (Eds.), *Trends in language acquisition research* (Vol. 26, pp. 207–

    235). John Benjamins. https://doi.org/10.1075/tilar.26.hur08

Hur, E., López Otero, J. C., & Sánchez, L. (2020). Gender agreement and assignment in Spanish

    heritage speakers: Does frequency matter? Languages, 5(48).

    https://doi.org/10.3390/languages5040048

Iverson, M., Kempchinsky, P., & Rothman, J. (2008). Interface vulnerability and knowledge of

    the subjunctive/indicative distinction with negated epistemic predicates in L2 Spanish.

    EUROSLA Yearbook, *8*, 135–163. https://doi.org/10.1075/eurosla.8.09ive

Kempchinsky, P. (2009). What can the subjunctive disjoint reference effect tell us about the

    subjunctive? *Lingua*, *119*(12), 1788–1810. https://doi.org/10.1016/j.lingua.2008.11.009

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/10.1177/1367006916654355

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). **lmerTest** package: Tests in

    linear mixed effects models. *Journal of Statistical Software*, *82*(13).

    https://doi.org/10.18637/jss.v082.i13

Lindholm-Leary, K. (2018). Developing Spanish in dual language programs: Preschool through

    twelfth grade. In K. Potowski (Ed.), *The Routledge handbook of Spanish as a heritage*

    language (pp. 433–444). Routledge.

Lindholm-Leary, K., & Genesee, F. (2014). Student outcomes in one-way, two-way, and indigenous language immersion education. *Journal of Immersion and Content-Based Language Education*, *2*(2), 165–180. https://doi.org/10.1075/jicb.2.2.01lin

Lohndal, T., & Putnam, M. T. (2021). The tale of two lexicons: Decomposing complexity across a distributed lexicon. *Heritage Language Journal*, *18*(2), 1–29. https://doi.org/10.1163/15507076-12340010

López Ornat, S., Fernández, A., Gallo, P., & Mariscal, S. (1994). *La adquisición de la lengua española* (1. ed). Siglo Veintiuno de España Editores.

López Otero, J. C. (2023). Imperatives in heritage Spanish: Lexical access and lexical frequency effects. *Languages*, *8*(218), 1–22. https://doi.org/10.3390/languages8030218

López Otero, J. C., Cuza, A., & Jiao, J. (2023). Object clitic use and intuition in the Spanish of heritage speakers from Brazil. *Second Language Research*, 026765832110176. https://doi.org/10.1177/02676583211017603

López Otero, J. C., Hur, E., & Goldin, M. (2023). Syntactic optionality in heritage Spanish: How patterns of exposure and use affect clitic climbing. *International Journal of Bilingualism*, 136700692311706. https://doi.org/10.1177/13670069231170691

Mariscal, S. (2009). Early acquisition of gender agreement in the Spanish noun phrase: Starting small. *Journal of Child Language*, *36*(1), 143–171. https://doi.org/10.1017/S0305000908008908

Martillo Viner, K. (2016). Second-generation NYC bilinguals' use of the Spanish subjunctive in obligatory contexts. *Spanish in Context*, *13*(3), 343–370. https://doi.org/10.1075/sic.13.3.02vin

Martinez-Nieto, L., & Restrepo, M. A. (2022). Production and comprehension of grammatical

gender by Spanish heritage speakers: Evidence from accusative clitic pronouns.

International Journal of Bilingualism, 136700692110573.

https://doi.org/10.1177/13670069211057318

Merino, B. J. (1983). Language loss in bilingual Chicano children. *Journal of Applied

Developmental Psychology*, *4*(3), 277–294. https://doi.org/10.1016/0193-3973(83)90023-

0

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

39). John Benjamins Publishing Company. https://doi.org/10.1075/sibil.39

Montrul, S. (2009). Knowledge of tense-aspect and mood in Spanish heritage speakers.

International Journal of Bilingualism, *13*(2), 239–269.

https://doi.org/10.1177/1367006909339816

Montrul, S. (2013). Incomplete L1 acquisition. In J. Herschensohn & M. Young-Scholten (Eds.),

The Cambridge handbook of second language acquisition (pp. 353–371). Cambridge

University Press. https://doi.org/10.1017/CBO9781139051729.022

Montrul, S. (2018). Heritage language development: Connecting the dots. *International Journal

of Bilingualism*, *22*(5), 530–546. https://doi.org/10.1177/1367006916654368

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., & Potowski, K. (2007). Command of gender agreement in school-age Spanish-

    English bilingual children. *International Journal of Bilingualism*, *11*(3), 301–328.

    https://doi.org/10.1177/13670069070110030301

Montrul, S., & Sánchez-Walker, N. (2013). Differential object marking in child and adult

    Spanish heritage speakers. *Language Acquisition*, *20*(2), 109–132.

    https://doi.org/10.1080/10489223.2013.766741

Peña, E. D., Gutiérrez-Clellen, V. F., Iglesias, A., Goldstein, B. A., & Bedore, L. M. (2014).

    Bilingual English-Spanish Assessment (BESA). BROOKES Publishing CO.

    https://brookespublishing.com/wp-content/uploads/2021/01/BESA-excerpt.pdf

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 – New Brunswick.

Perez-Cortes, S. (2022). Lexical frequency and morphological regularity as sources of heritage

    speaker variability in the acquisition of mood. *Second Language Research*, *38*(1), 149–

    171. https://doi.org/10.1177/0267658320918620

Perez-Cortes, S., Putnam, M., & Sánchez, L. (2019). Differential access: Asymmetries in

    accessing features and building representations in heritage language grammars.

    Languages, *4*(81), 1–27. https://doi.org/10.3390/languages4040081

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

Potowski, K. (2004). Student Spanish use and investment in a dual immersion classroom:

    Implications for second language acquisition and heritage language maintenance. *The*

Modern Language Journal, *88*(1), 75–101. https://doi.org/10.1111/j.0026-7902.2004.00219.x

Potowski, K. (2007a). Characteristics of the Spanish grammar and sociolinguistic proficiency of dual immersion graduates. *Spanish in Context*, *4*(2), 187–216. https://doi.org/10.1075/sic.4.2.04pot

Potowski, K. (2007b). *Language and identity in a dual immersion school*. Multilingual Matters Limited.

Putnam, M., Kupisch, T., & Pascual y Cabo, D. (2018). Different situations, similar outcomes: Heritage grammars across the lifespan. In D. Miller, F. Bayram, J. Rothman, & L. Serratrice (Eds.), *Bilingual cognition and language: The state of the science across its subfields* (pp. 251–280). John Benjamins Publishing Company.

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

Putnam, M. T., Sánchez, L., & Perez-Cortes, S. (2019). Language attrition and the Feature Reassembly Hypothesis. In M. S. Schmid & B. Kopke (Eds.), *Oxford handbook of language attrition* (pp. 18–24). Oxford University Press.

R Core Team. (2022). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Rodríguez-Mondoñedo, M. (2008). The acquisition of Differential Object Marking in Spanish. *Probus*, *20*(1). https://doi.org/10.1515/PROBUS.2008.004

Rojas, D. M. (1998). The situation of the subjunctive: Perception and comprehension of subjunctive forms in English. *The SECOL Review*, *22*(2), 85.

Seco, R. (1990). *Manual de gramática española* (11th ed.). Aguilar.

Sherkina-Lieber, M. (2015). Tense, aspect, and agreement in heritage Labrador Inuttitut: Do

receptive bilinguals understand functional morphology? *Linguistic Approaches to*

Bilingualism, *5*(1), 30–61. https://doi.org/10.1075/lab.5.1.02she

Swain, M. (1993). The Output Hypothesis: Just speaking and writing aren't enough. *The*

Canadian Modern Language Review, *50*(1), 158–164.

https://doi.org/10.3138/cmlr.50.1.158

Vannest, J., Newport, E. L., Newman, A. J., & Bavelier, D. (2011). Interplay between

morphology and frequency in lexical access: The case of the base frequency effect. *Brain*

Research, *1373*, 144–159. https://doi.org/10.1016/j.brainres.2010.12.022

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G.,

Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller,

K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the

Tidyverse. *Journal of Open Source Software*, *4*(43), 1686.

https://doi.org/10.21105/joss.01686

Competing interests: The author(s) declare none.