5

Bridging Instruction of the Spanish Subjunctive: Exploring Task Types for Heritage and L2 Learners

Sara Fernández Cuenca
Wake Forest University

Abstract

To this day, there are only a few studies that have used a controlled instructional intervention on specific linguistic structures to investigate if language instruction is beneficial for heritage learners (HLs), in the same way it is for second (L2) language learners, and more research in this area is rapidly needed (Bowles, 2018). The results from this small body of research suggest that explicit language instruction can be beneficial for HLs, but overall L2 learners still appear to benefit from language instruction more than their HLs counterparts (Potowski et al., 2009; Torres, 2018, inter alia). The present study seeks to contribute to this growing body of research and further examines if the type of task employed to measure learning gains plays a role in the uneven outcomes heritage and L2 learners evidence after receiving language instruction.

Keywords*: instructed language instruction, assessment tasks, L2 and HL Spanish, subjunctive mood*

Introduction

The field of heritage language acquisition has grown dramatically in the last 20 years. Heritage languages and their speakers have gained more attraction and have become a central focus of different areas of linguistic research. From theoretical accounts that seek to investigate how interrupted input exposure to a first language (upon starting schooling) can result in language loss or attrition (Bayram et al., 2020; Polinsky, 2011), to psycholinguistic accounts that explore if early acquisition (in comparison to later onset of acquisition during puberty) grants heritage speakers (HS) the ability to use abstract grammatical information in real-time (Jegerski, 2018). Nevertheless, less focus has been placed on instructed heritage language acquisition (IHLA), which examines how controlled instructional interventions focusing on specific linguistic structures can help heritage learners (HL) master formal aspects of their heritage language grammar that show signs of differential acquisition (Pascual y Cabo & Rothman, 2012) due to reduced and differential input in a minority context (Bowles, 2022). In fact, as Bayram and colleagues (2016) pointed out, formal and

pedagogical approaches to heritage language acquisition have remained separate and disconnected, but understanding formal linguistic approaches to heritage language acquisition can be extremely beneficial to the development of heritage language pedagogies, as it can help establish how the linguistic system of HS differs from that of monolingual speakers and of second language learners, and can therefore aid establish which and how certain formal aspects of the language could be addressed and taught in the Spanish language curriculum.

According to the most recent American Community Survey data from 2019, the number of Spanish speakers in the United States is 67,802,345 million, making it the most spoken minority language in the United States. Consequently, we have seen that the number of Spanish HS enrolling in Spanish language courses has also increased rapidly (Bowles, 2018). Nevertheless, little is known about the outcomes of classroom teaching of minority languages (Bowles & Torres, 2021; Montrul & Bowles, 2017). Some argue that world languages curricula, designed for foreign language learners, are largely inappropriate for HLs (Carreira & Kagan, 2011; Oh & Au, 2005), and that on some occasions sharing a classroom with Spanish second language (L2) learners can cause HLs to feel insecure about their language ability (Potowski, 2002), affecting their learning outcomes. Today's reality is that most heritage learners (HLs) who choose to study their heritage language in a formal context (school or college) end up in a foreign language classroom with L2 learners, because very few programs have a separate teaching track for HLs that is designed to meet their learning needs (Beaudrie, 2012). Research in mixed classrooms suggests that heritage and L2 learners can work together for mutual benefits, if oral and written tasks are balanced and the two types of learners have a similar proficiency level (Bowles et al. 2014; Henshaw & Hetrovicz, 2021). Unfortunately, this optimal scenario is very rare given that foreign language assessment tools tend to be written non-open-ended tasks (Kang et al., 2019), and the number of standardized tests developed to assess HLs' knowledge to place them in the most appropriate class are scarce (Potowski et al., 2012). If we want education, which is a human right, to be more inclusive, more research efforts should be put towards understanding if and how HLs can develop and/or maintain their heritage language (beyond what they acquire at home) through language and literacy instruction.

The field of heritage language instruction is broad, but the present study aims to focus on the subfield of grammar language instruction which is characterized for employing a traditional approach including the use of pre- and posttests that investigates the effects of one or two pedagogical approaches on one grammatical variable (e.g., Spanish subjunctive in adjectival clauses). Within this subset of studies, results often point to explicit instruction—containing explicit information and feedback—as being particularly beneficial (Beaudrie & Holmes, 2022; Bowles & Torres, 2021; Fernández Cuenca & Bowles, 2022), when compared to implicit approaches. Furthermore, the few studies that investigate how heritage and L2 learners respond to grammar instruction, point to L2 learners benefitting from language instruction more than HLs, at least based on the outcome measures used to assess learning over time (Potowski et al., 2009; Torres, 2018; among others). That is, HLs sometimes do not show improvement or exhibit partial learning gains (Montrul & Bowles, 2010) depending on the assessment task used (Fernández Cuenca &

Bowles, 2022; Montrul & Bowles, 2010; Potowski et al., 2009), whereas L2 learners show significant learning gains across assessment tasks over time and, in some cases, higher effect sizes (Bowles & Torres, 2021; Torres, 2018). Some attribute these differences to heritage and L2 learners approaching the assessment tasks differently–such as L2 learners focusing on form while HLs tend to focus more on meaning and the communicative content of the task at hand–and to HLs not perceiving corrective feedback as corrective (Torres, 2018), which could explain why implicit approaches to language instruction that do not bring awareness to form are the ones leading to the weakest or no learning gains at all for HLs (Beaudrie & Holmes, 2022; Fernández Cuenca & Bowles, 2022). Spanish heritage and L2 learners differ in the context as well as the mode in which they acquire Spanish. L2 learners begin learning a foreign language later in life (i.e., after puberty) with the input coming mostly from a language instructor in a classroom setting, and more prominently in a written than oral form. On the other hand, HLs acquire their heritage language early in life in an oral mode, and do not always receive formal education in their heritage language. Consequently, these differences are likely to influence both the nature of heritage and L2 learners' linguistic knowledge and how they draw on it when approaching language learning or completion of a language task (Bowles, 2011).

As Bowles (2018) points out, the field of instructed heritage language acquisition (IHLA) is just starting to grow, and we need more experimental studies that manipulate aspects of instruction and compare the learning outcomes of different conditions with similar groups of learners to be able to directly assess how effective grammar language instruction is for adult heritage learners. Furthermore, identifying the instructional factors that lead to differential learning outcomes for heritage and L2 learners is paramount since both learners tend to share a language classroom, and understanding how to best adapt language instruction so it becomes equally beneficial for both types of learners is our responsibility as educators (see Carreira & Chik, 2018, to learn more about differentiated instruction). Finally, Bowles (2018) and Bowles and Torres (2022) called for researchers to adopt a systematic empirical approach grounded on conceptual or partial replication, largely based on instructed second language acquisition (ISLA) research, which can confirm or challenge previous findings and that allows us to collect data from larger groups of learners to be able to generalize findings.

The present study seeks to contribute to this growing body of research on the outcomes of heritage language instruction by conducting a partial replication of Fernández Cuenca and Bowles (2022), and by adding a comparison L2 group. The addition of a comparison L2 group has been the norm in previous studies that further examined if the type of outcome measure employed to assess learning gains plays a role in the uneven benefits heritage and L2 learners evidence after receiving language instruction. To our knowledge only two studies, one with L2 learners (Sanz, 1997), and one with HLs (Torres, 2022) have investigated how task modality, and in the case of Sanz (1997) also degree of discreteness (how open-ended the tasks were—), can interact with the observed learning gains these learners exhibit post-instruction.

The rest of the paper is organized as follows. We start by discussing: (a) the factors that make acquisition of grammatical mood (particularly subjunctive mood) challenging for both heritage and L2 learners, and (b) we synthesize the literature on

Spanish mood acquisition for both types of learners, and how language instruction can help them learn this grammatical feature. In the methods section, we lay out the research questions, provide an in-depth description of the methodology adopted in the study, and report the results from the statistical analyses. We close the paper with a discussion of the importance of the findings, how they compare to previous studies, followed by a list of some pedagogical implications for educators teaching heritage or mixed classes.

Literature Review

Acquisition of Spanish Mood

Spanish grammatical mood is semantically abstract and linguistically complex because it involves, morphology, sentence-level semantics, and morphosyntax. The meaning conveyed by the verb or lexical expression in the matrix clause triggers either the use of subjunctive or indicative in the subordinate clause, resulting in a long-distance dependency in which the use of the subjunctive mood in the subordinate clause depends on the lexical semantics of the verb or lexical expression in the matrix clause (see example 1).

1. Joanne quiere una casa que esté en la ciudad
   Joanne want-PRES3SG a house that be-PST-SBJV in the city
   'Joanne wants a house in the city.'

The subjunctive mood is less frequently used than its indicative counterpart (Biber et al., 2006; Kanwit & Geeslin, 2018). It is also highly variable among monolingually-raised Spanish speakers who still live in a Spanish-speaking country, and this characteristic of the subjunctive holds true for speakers who migrated and reside in the United States (Blas-Arroyo & Porcar Miralles, 1997; Gudmestad, 2006, 2012; Viner, 2018). Moreover, if the verb in the embedded clause is regular, and the difference between the indicative and the subjunctive mood is only marked by a change in thematic vowel (a→e or e→ a, i), the morphological saliency is minimal (Collentine, 1997; Gudmestad, 2006). For this gamut of reasons, it is not surprising that acquisition of mood, especially subjunctive mood, posits a challenge for heritage and L2 learners.

Acquisition of Mood by Heritage and Second Language Learners

Studies that analyzed heritage speaker corpora in the United States have found that there is an intergenerational decline in subjunctive mood use, compared to traditional monolingual standards (Lynch, 1999; Silva-Corvalán, 1994; Viner, 2018). A pattern in which the use of indicative mood grows categorically in the speech from second and third-generation HS in contexts where the subjunctive is normatively expected, and that drops drastically in non-obligatory contexts—a context in which the speaker intentionally uses mood to signal what they mean (e.g., degree of certainty)—, even with first-generation Spanish speakers (Silva-Corvalán, 1994). In the case of the imperfect of subjunctive, which is the target form used in the

present study, Silva- Corvalán (1994) found that less fluent third-generation HS did not use this form at all. Nonetheless, Viners' (2018) most recent study revealed that subjunctive mood was still substantially operational in second-generation HS, but the linguistic constraints conditioning the variation of the two moods (e.g., tense) seemed to be weaker than those observed in first-generation Spanish speakers.

Studies that employed interpretation and production tasks to elicit HS' responses, rather than naturalistic conversation or guided sociolinguistic interviews, also found that Spanish HSs exhibit a gradual loss of mood distinction, which often translated into overproduction of the indicative mood in contexts were the subjunctive is expected (Silva-Corvalan, 1994; Van Osch et al., 2017). Nonetheless, this pattern did not always occur in variable contexts (Perez-Cortes, 2022), supporting the intergenerational differences in grammatical mood competence found by the aforementioned studies (Silva-Corvalán, 1994; Viner, 2018).

It is important to note that HSs' behavior is not uniform, and one can encounter high degrees of variability in interpretation and production tasks that target obligatory contexts (Montrul, 2007; Montrul & Perpiñán, 2011; Pascual y Cabo & Rothman, 2012; Van Osch et al., 2017), but not always so much in contexts where indicative and subjunctive forms can co-exist and using one or the other expresses different meanings that are considered grammatical (Perez-Cortes, 2022). More recent studies have pointed to factors such proficiency and onset of acquisition of the majority language as potential modulators of the variability in subjunctive mood use exhibited by second and third generation HSs. These studies' outcomes suggest that HSs of Spanish with an intermediate and low proficiency did not appear to have a full command of Spanish mood (Giancaspro, 2019b; Montrul, 2007; Montrul & Perpiñán, 2011; Perez-Cortes, 2021) and that HSs who began acquiring the majority language later (from 8 to 12 years age) often displayed higher rates of subjunctive preference and production, in comparison to early acquirers (Giancaspro, 2019b).

Similarly, L2 speakers of Spanish find grammatical mood, particularly subjunctive, to be a difficult construct to learn. However, subjunctive mood can be acquired to some degree, even if such distinction is not present in the learners' native language (Bornogovo et al., 2005). Studies that employed a more explicit methodology, found that L2 speakers with high levels of proficiency can perform close to or in a nativelike fashion with trigger verbs that depict volition, or with negated epistemic and perception predicates in tasks that tap into interpretation and judgment of Spanish mood use (Borgonovo et al., 2005; Iverson et al., 2008; Massery, 2009). In terms of oral production, there is evidence that high proficiency L2 speakers' oral production of mood can be very similar to that of native speakers in terms of frequency and contextual factors that shape variation (e.g., semantic category of the trigger verb), with the only exception being the discourse pragmatic variable of hypotheticality (Gudmestad, 2012). There have also been a few extralinguistic factors that appear to shape mood production among L2 learners. For instance, there is evidence that L2 learners produce more subjunctive forms in written than in oral tasks (Geeslin & Gudmestad, 2008; Montrul 2011) and that more nativelike performance seems to occur with more focus on form tasks, such as a written controlled production tasks or a verb elicitation task, than with a more open-ended free task, such as an oral interview (Collentine, 1995; Terrell et al., 1987).

To summarize, acquisition of the Spanish subjunctive mood posits a problem for both heritage and L2 learners. The Spanish subjunctive (particularly past subjunctive) is a grammatical construct difficult to learn for both heritage and L2 learners because it is late acquired, and because it does not have a direct equivalent in English. As previously mentioned, Spanish L2 and heritage learners often share the same classroom, and even when they do not, the concept of past subjunctive is covered in Spanish language textbooks designed for L2 as well as in textbooks designed primarily for heritage speakers (Potowski, 2010). Thus, the Spanish past subjunctive presents a unique opportunity to explore the effects of controlled language instruction on these two student populations' learning outcomes.

Instruction of the Spanish Subjunctive with L2 Learners

Despite its low frequency in oral and written input compared to its indicative counterpart (Kanwit & Geeslin, 2018), the Spanish subjunctive is always present in Spanish textbooks and taught sometimes as early as in second- and third-year Spanish courses, even though some have argued that L2 learners at this level are not always ready to acquire this linguistic construction (Farley & McCollam, 2004; Massery, 2009). It is probably for this reason that there is a robust body of research that has investigated the most effective pedagogical approaches to teaching Spanish subjunctive.

One of the instructional interventions that has received more attention is processing instruction (PI) (VanPatten, 1996, 2015), a pedagogical intervention that takes into account some of the processing pitfalls that impede L2 learners' from making accurate form-meaning connections (see VanPatten, 1996, for a more in-depth description of the tenets of Processing Instruction). The effectiveness of PI with L2 learners, in contrast with other pedagogical intervention such as traditional output-oriented instruction or meaning-based output-oriented instruction, has been widely studied (Benati, 2001; Morgan-Short & Bowden, 2006; Potowski et al., 2009; VanPatten & Cadierno, 1993, among others). Within the subset of studies that targeted the Spanish subjunctive, results suggested that output-oriented instruction–as long as it is meaning-oriented–could lead to similar interpretation and production gains as PI does (Collentine, 1998; Farley, 2004a; Morgan-Short & Bowden, 2006; Shintani, 2015). However, most studies found an advantage for learners in the PI group, who displayed significantly greater learning gains, especially when PI was combined with visual input enhancement (Farley, 2001; Kirk, 2013; Russell, 2009) and these learning gains were present across different subjunctive constructions that differed in their degree of markedness, that is, the state of standing out as nontypical or divergent as opposed to regular or more common (Pereira, 1996).

Subsequent studies have moved to examine if the explicit information (EI) component of PI is necessary, or if structured input (SI) activities alone are enough to generate accurate form-meaning connections that lead to positive learning outcomes. Overall, these studies found an advantage for the learners who received a full version of PI (i.e., including EI), as they tended to outperform learners who only completed SI activities (Bowles & Henshaw, 2015; Farley, 2004b; Farley & McCollam, 2004). In addition, Fernández (2008) and Bowles & Henshaw (2015) found that learners in the PI group made the appropriate form-meaning connections faster

than those in the SI group. However, it is important to acknowledge that PI (with or without EI) leads to substantial learning gains in interpretation and production over time, even when L2 learners are categorized as not "ready" to acquire the subjunctive according to processability theory parameters (Farley & McCollam, 2004). Although not exactly PI, Adrada-Rafael (2017) observed that more explicit instructional interventions led to deeper processing of subjunctive form-meaning connections, which in his study correlated with more accurate production. More recently, McNulty-Díaz (2017) and Fernández Cuenca (2019) investigated if PI can aid L2 learners make appropriate subjunctive form-meaning connections in adverbial clauses and their findings confirmed that not only was PI an effective instructional intervention, but its effectiveness was supported even if the order of its components (explicit information and structured input) was inverted (McNulty-Díaz, 2017), leading to positive changes that have been found even in real-time processing (Fernández Cuenca, 2019).

In sum, PI can facilitate L2 learners' acquisition of subjunctive mood and points to metalinguistic knowledge of subjunctive mood as being an asset in the acquisition of Spanish subjunctive. This finding is consistent with Correa (2011) who found that explicit metalinguistic knowledge of Spanish subjunctive positively correlated with subjunctive accuracy use, suggesting that L2 learners with high metalinguistic knowledge of Spanish subjunctives are more accurate in mood selection.

Instruction of the Spanish Subjunctive with Heritage Learners

Research on the instruction of the Spanish subjunctive with heritage learners (HLs) has been less extensive with only a handful of studies that often include a comparison L2 group (Fernández Cuenca & Bowles, 2022; Potowski et al., 2009; Torres, 2018, 2022). Potowski and colleagues (2009) investigated if traditional and processing instruction (TI, PI) facilitated acquisition of past subjunctive mood in adjectival clauses with definite and indefinite referents for Spanish L2 and HLs. In their study, participants were randomly assigned to a PI group that received explicit information including the description of two faulty processing strategies that affect noticing of subjunctive mood, and structure input practice (consisting of five referential and five affective activities). On the other hand, participants assigned to the TI group received output-oriented instruction consisting of explicit information on the past subjunctive and form-focused activities commonly found in language textbooks that include a grammar component. Participants in the control group only completed the pre- and posttest which consisted of an interpretation, production, and grammaticality judgment task. The interpretation task resembled a referential SI activity, the production task was a sentence completion activity, and the grammaticality judgment task (GJT) had a binary choice ("it has no mistakes, sounds good" or "it has a mistake, it does not sound good"). The authors found that HLs in both experimental groups experienced equal (moderate) learning gains in interpretation and production over time, compared to the control group, but no improvements were observed with HLs' judgments of stimulus sentences where the subjunctive in adjectival clauses was used. On the other hand, L2 learners evidenced greater linguistic development with a higher overall accuracy post-instruction and higher effect sizes that were present across all three tasks regardless of the type of instruction.

Following this initial study, Torres (2018, 2022) investigated the effects of task-based language instruction on L2 and HLs' knowledge of subjunctive also in adjectival clauses, but this time in the present tense. As part of the treatment, participants were asked to adopt the role of a director in a university residence, who had to explain disconcerting behavior that was taking place among residents. Participants were assigned to one of two experimental conditions that differed in their cognitive demands (i.e., task complexity, see Torres 2018) but consisted of a series of monologic computerized sentence completion tasks that also delivered written recasts as a form of corrective feedback. A writing and oral sentence completion task (contextualized with an image) were used to assess learning outcomes as a pre-, post-, and delayed posttest 1-2 weeks post-instruction. Results showed, once again, moderate improvements in interpretation and production over time that were greater in the oral than the written assessment task for the HLs. These findings contrasted with those of two comparison L2 groups (+/- complex), which displayed greater overall learning gains and higher effect sizes in the immediate and delayed posttest. Interestingly, the exit questionnaire revealed that L2 learners had taken a more focus-on-form approach to complete the task, whereas HLs were more focused on meaning and content. Torres concluded that these differences in learning gains could be partly explained by the difference in approach that the two types of learners had adopted.

Finally, Fernández Cuenca & Bowles (2022) examined the effects of explicit and implicit language instruction on HLs' knowledge of Spanish past subjunctive in adjectival clauses. Their explicit instruction included EI followed by an explanation of faulty processing strategies that could hinder acquisition of subjunctive morphology, and structured input practice. The implicit treatment, on the other hand, consisted of input flood. Participants in this experimental group did not receive explicit information; instead, they read the sixty-two items that participants in the explicit group read in the SI portion, but in the form of a written story followed by a series of comprehension questions (similar to an input flood treatment). Instruction was delivered in a written mode, and learning gains were assessed with a computerized written AJT (with a 5-point Likert scale) and an oral elicited imitation task (EIT) which encompassed the pre-, post-, and delayed posttest. Results revealed that HLs in the explicit group exhibited significant improvement over time with ungrammatical items in both tasks and with grammatical and ungrammatical items in the EIT task, in both cases, these learning gains were still present in the delayed posttest, two weeks post-instruction. In contrast, HLs assigned to the implicit group only showed learning gains with ungrammatical items from pre- to posttest that disappeared by the time they completed the delayed posttest, a week post-instruction. In a debriefing questionnaire used to subjectively target the source of knowledge that HLs were using to complete the pre- and posttests, a participant in the explicit group reported using the rule for the written test and intuition for the oral task (i.e., the EIT). This finding is in line with Chomón Zamora (2022) who found that HLs predominantly used intuition upon receiving explicit feedback, followed by some use of grammatical knowledge, in contrast with L2 learners, who tend to rely primarily on explicit and comparison grammar.

Overall, explicit language instruction seems to have a positive impact on HLs'

grammatical mood development leading to long-lasting learning gains. However, when compared to the learning gains that L2 learners exhibit, HLs appear to benefit less from language instruction. HLs' responses to the exit surveys used by Torres (2018) and Fernández Cuenca and Bowles (2022) appear to indicate that L2 and heritage learners approach the completion of the assessment tasks differently. Thus, it is possible that the assessment task itself plays a role in the uneven learning outcomes observed between these two learner groups.

The Assessment Task

Due to the context and mode of acquisition in which heritage and L2 learners acquire Spanish, being different in terms of mode (oral vs. multimodal) and context (formal vs. informal), it is safe to assume that this does influence both the nature of their linguistic knowledge and how they draw on it when learning language or completing a language task. In fact, this question has been directly addressed in more recent studies that adopted a comparative approach (e.g., Bowles, 2011; Torres, 2018). The few instructional studies that included a heritage and L2 learner group consistently found that L2 learners benefitted from instruction more than HLs and this observation was based on a positive increase in accuracy rates from pre to posttests, often accompanied by higher effect sizes, that were consistent across assessment tasks. A possible explanation for this observed difference could be the outcome measure used to assess learning gains. The way learning gains were assessed in most of these studies that included only an L2 group or both types of learners, points to a bias toward explicit knowledge given that most tasks were designed to elicit a constrained discrete response. This may pose a disadvantage for HLs, who tend to perform worse in more focus-on form-tasks such a timed GJT, than in less explicit-oriented tasks such as an EIT, in comparison to L2 learners who show a reversed pattern (Bowles, 2011; Ellis, 2005). Furthermore, the modality of the task appeared to modulate performance for at least HLs. For instance, HLs in Fernández Cuenca & Bowles (2022) made significant improvements with grammatical and ungrammatical items in the oral assessment task (i.e., EIT), but only with ungrammatical items in the written assessment task (i.e., AJT), a similar pattern to Torres (2022), who found that HLs displayed greater learning gains in an oral than a written assessment task post-instruction. On the contrary, in Sanz's (1997) study, which is–to our knowledge–the only instructional study with L2 learners that directly controlled for task modality and discreetness, L2 learners exhibited significantly superior performance in the written most discrete assessment tasks (i.e., a sentence completion task) vis-a-vis the oral assessment tasks post-instruction. This modality factor affecting heritage and L2 learners' performance differently post-instruction has also been found in non-instructional studies, which consistently show HLs performing better in oral than in written tasks and the opposite pattern for L2 learners (Alarcón, 2011; Montrul & Perpiñán, 2011). Non-instructional studies often find that the task's degree of explicitness and modality seem to play an important role in L2 and heritage learners' performance and this can therefore obscure the observed effects of language instruction. More precisely, with regard to the Spanish subjunctive, in particular, these findings point to more focus on

form/discrete written tasks as being advantageous for L2 learners (e.g., Geeslin & Gudmestad, 2008), which could, in turn, help explain why L2 learners often seem to benefit from language instruction more than HLs do (Potowski et al., 2009), at least with this grammatical construction.

The present study seeks to replicate previous findings supporting the effectiveness of explicit instructional interventions with both learner types (HL and L2) in the learning of the Spanish subjunctive. Furthermore, it aims to tease apart if the modality of the task used to assess longitudinal learning gains may play a role in our understanding of instructional effectiveness. If prior language experience, in terms of context and mode of acquisition, affect these two groups of learners' approach to learning and completion of a language task differently, it is our duty as applied linguists to investigate this empirically with the goal to provide fair assessment of pedagogical interventions' efficacy and overall language assessment moving forward.

Methods

Research Questions

The present study is a partial replication and extension of Fernández Cuenca and Bowles (2022), which aims to fill this gap by examining the effects of explicit language instruction on both heritage and L2 learners' knowledge of Spanish subjunctive using an untimed written acceptability judgment task (AJT) and an oral elicited imitation task (EIT) as outcome measures. We posit the following research questions.

1. Do L2 learners benefit from explicit instruction on the Spanish past subjunctive?
2. Do HLs benefit from explicit instruction on the Spanish past subjunctive?
3. Does the modality of the outcome measure employed modulate heritage and L2 learners observed learning gains differently?

Participants

The first group of participants consisted of 39 undergraduate college-level HLs (31 female and eight male). They were all registered in upper-level Spanish content courses at the time of recruitment, and they were all Spanish major or minors. Thirty-four participants reported having been born in the United States, whereas five were born in a Spanish-speaking country and moved to the United States before the age of six, making them all second-generation HSs. Of the 39 HLs, 22 were simultaneous bilinguals, and the remaining 17 were sequential. All HLs reported having at least one parent who spoke Spanish, but most of them indicated that one of their parents could speak both languages fluently. All HLs had taken at least one Spanish course in college and thirteen listed having attended a Spanish-English bilingual elementary school.

The second group of participants consisted of 40 L2 learners (30 female and 10 male), who were Spanish majors and minors and were enrolled in advanced Spanish courses at the time of recruitment. The majority of these L2 learners were

raised in a monolingual household (except for eight L2 learners that reported an intermediate level of proficiency in a minority language other than Spanish), and they learned Spanish formally in a classroom consistently after puberty. All participants completed a short questionnaire with a 10-point scale to self-report their English and Spanish proficiency and were also asked to complete the modified version of the DELE standardized Spanish proficiency test (Montrul & Slabakova, 2003). See Table 1 for a more detailed description.

Table 1
Participants' Language Background Information

	L2[a]			HL		
	M	SD	range	M	SD	range
Age	20.10	.65	19-21	19.92	1.20	18-23
Age of Acquisition						
English	.35	.60	0-2	3.69	2.40	0-6
Spanish	10.91	2.78	7-15	.59	1.64	0-3
DELE	34.81	2.98	29-39	42.3	6.1	35-45
Self-ratings English						
Comprehension	9.88	.31	9-10	9.56	.91	8-10
Reading	9.86	.42	8-10	9.71	.51	8-10
Writing	9.80	.46	8-10	9.53	.64	8-10
Speaking	9.88	.31	9-10	9.56	.78	6-10
Self-ratings Spanish						
Comprehension	6.77	1.24	3-9	8.53	1.33	4-10
Reading	7.13	1.64	5-10	8.20	1.68	4-10
Writing	6.97	1.20	5-9	7.48	1.93	3-10
Speaking	5.97	1.34	4-8	8.02	1.58	3-10

Note: The maximum score was 50 for the DELE and 10 for the self-rated proficiency.
[a] These are the minority languages that eight L2 learners reported speaking to some degree, respectively: Afrikaans, Arabic, Danish, Greek, Russian (n =2), Tagalog, and Thai. Age was measured in years and participants could report 0 (from birth).

The Target Form

The present study targets the imperfect of subjunctive in subordinate adjectival clauses with specific and non-specific referents, to be consistent with previous research that had explored instructional effects with both types of learners (Potowski et al, 2009; Fernández Cuenca & Bowles, 2022), and because past subjunctive posits an acquisitional challenge for both heritage and L2 learners (Silva-Corvalán, 1994). As such, this target form allows for a fairly similar baseline to assess learning gains for L2 and HLs over time. In the case of the subjunctive with (in)existential clauses, the subjunctive can be found in the embedded clause when the referent in the matrix clause is unknown or nonspecific (see example 2b). However, the indicative mood is expected if the entity in the matrix clause is something specific or known (example 2a). The use of the indicative mood in a relative clause when the referent in the main clause is nonspecific or unknown is considered non-standard (example 2c).

2a. Michelle encontró a los turistas que hablaban español. Grammatical

"Michelle found tourists that could speak$_{IND}$ Spanish"

2b. Michelle no encontró a ningún turista que hablara español. Grammatical

2c. Michelle no encontró a ningún turista que hablaba español. Ungrammatical

"Michelle didn't find any tourist that could speak$_{SUBJ/IND}$ Spanish"

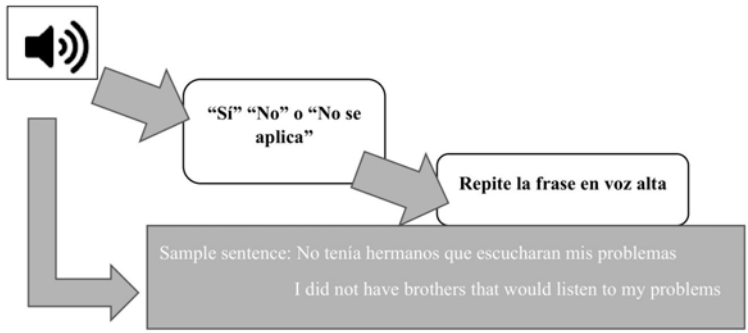
Due to the non-categorical nature of the subjunctive mood, being used slightly differently even by monolingually raised speakers (Waltermire, 2017), the stimuli employed in the assessment tasks and instructional module were normed with a group of 20 monolingually raised Spanish speakers from different Spanish speaking countries. Their responses showed that when the antecedent was nonspecific or unknown (often marked by negation), sentences were rated as "completely acceptable" 96% of the time, and as "unacceptable" 4% of the time, in this later case, due to reasons that did not involve the use of subjunctive mood, but rather the syntax of the sentence. On the other hand, when the sentence had a nonspecific or unknown referent in the main clause and the verb in the embedded clause was in the indicative mood (as in example 2c), they rated them as "completely unacceptable" 74% of the time, and as "acceptable" 26 % of the time. Overall, these findings mirrored those of previous studies that found subjunctive mood use to be variable among Spanish monolingual and monolingually raised speakers (Blas-Arroyo & Porcar Miralles, 1997; Murillo-Mendrano, 1999) as well as with second-generation heritage speakers (Viner, 2018).

Materials

Participants in the present study completed a pretest, immediate posttest, and a two-week delayed posttest consisting of an elicited imitation task (EIT), which was followed by an untimed acceptability judgment task (AJT) (identical to the ones employed in Fernández Cuenca and Bowles, 2022).

In the EIT task, which was administered via PowerPoint, participants were asked to (a) listen to a statement in Spanish, (b) state whether this statement applied to them by saying "sí", "no", "no se aplica" ["yes", "no", "does not apply"] and (c) to repeat the statement they had heard correctly (if it contained a mistake). See Figure 1 below for a visual representation of a sample trial. In keeping with previous EIT research, we followed a series of parameters to ensure that learners had to reconstruct the sentences they heard in the EIT and could not just simply repeat them (Erlam, 2006; Yan et al., 2016). First, sentences were long and included a relative clause on the likelihood that it would exceed the participants' working memory span. Second, after listening to each utterance, participants had to make a judgment regarding whether the content of the sentence applied to them or not, which pushed them to focus on meaning—rather than form—, before they proceeded to reconstruct the sentence to the best of their ability. Third and last, there was a delay between presentation of the stimuli and repetition to avoid rote repetition. The EIT contained eight experimental items, four grammatical and four ungrammatical, in addition to eight distractors that targeted other grammatical constructions. The full task can be found in Appendix A.

Figure 1 EIT Task Procedure



After completing the EIT, learners moved to the AJT which was administered using Survey Gizmo. In this task, participants were asked to rate the acceptability of a series of statements that were presented in writing and orally, using a 5-point Likert scale with 1 being totally unacceptable and 5 being completely acceptable. In addition to rating these sentences, participants had to correct sentences (with a rating under five) to make them completely acceptable using a textbox that appeared after each stimulus sentence. The presentation of the stimuli in this task was bimodal to accommodate HLs who might prefer aural to written stimuli. Participants read or listened to a total of eight experimental and eight distractor items (half of them grammatical and half ungrammatical). This additional step enabled us to examine if the corrections implemented involved the target form addressed in the study. The whole task can be found in Appendix B. Three different versions of these two tests (EIT and AJT) were created and counterbalanced so they would be presented at the pre-, post-, and delayed posttest the same number of times with heritage and L2 learners. Finally, these two tasks were also employed as a screening test at the pretest, to ensure that participants were not familiar with the target form, and as assessment tests to capture potential learning gains over time. Only learners who scored less than 60% in both tasks at the pretest stage, which is the traditional baseline in PI studies–, were allowed to continue onto the second phase of the study (instruction/immediate posttest), which took place within a week. The delayed posttest took place two weeks post instruction. Overall, participants completed the whole study within 3-4 weeks. Reliability for all three versions of these two tests were measured by Cronbach alpha (see Table 2), which reflected medium to high instrument validity (Plonsky & Derrick, 2016).

Table 2
Cronbach Alpha Values for Task Reliability

	Version A	Version B	Version C
AJT	.72	.81	.85
EIT	.70	.93	.89

Pedagogical Intervention

Heritage and L2 learners were randomly assigned to an experimental and a control group, which led to a final distribution of 19 HLs in the explicit group and 20 HLs in the control group, and to 20 L2 learners for both experimental and control group. The instructional intervention took place in a laboratory setting, where students completed a self-paced instructional module in a desktop computer, which lasted approximately 30-40 minutes. This instructional module was designed following PI principles (VanPatten, 1996) that consisted of explicit information (EI) on the target form and accurate processing strategies, followed by structured input (SI) practice. In addition to reading the explicit information, participants were prompted to answer a few multiple-choice questions targeting the explicit information they had just read to ensure that they had understood this information, before engaging with the SI practice. SI items were designed so that learners first read an incomplete relative clause and were asked to choose which matrix clause out of four available options was the most appropriate to start the sentence (see example 3). Both the incomplete written sentence and possible endings appeared in the same screen. Participants were asked to press the letter key that matched their chosen response and after they pressed the corresponding key, they moved to a new screen where they received corrective feedback ("correct" or "incorrect") before repeating the process all over again with each of the total sixty-two items that encompassed the SI practice module (see Figure 2 below). The whole instructional module including: (a) the grammatical explanations, (b) the brief multiple-choice items used to verify that participants understood the explicit terminology explained, and (c) SI practice. All three modules were administered using the psychology software Paradigm. The stimuli in the instructional module were only presented in a written modality. Learners assigned to the control group completed the EIT and AJT tasks at the pre- and posttests, but they did not receive instruction.

3. tuviera acento de Chicago.	had$_{\text{PAST-SUBJ}}$ an accent from Chicago
A. Había mucha gente que …	A. There were people that …
B. No había mucha gente que…	B. There weren't many people that…
C. Hay mucha gente que…	C. There are people that…
D. Ninguna	D. None of the above

Figure 2

Sample Trial Structured Input Practice

Data Coding and Analysis

The present study examined accuracy scores for the experimental sentences in the EIT and AJT at three points in time. The scoring procedure for the EIT task was conducted as follows: grammatical sentences that were repeated grammatically were assigned a 1, but those that were considered ungrammatical because their repetition did not include the standard use of the past subjunctive, when the referent was unknown, received a 0. We acknowledge that we considered instances of other forms that were not imperfect subjunctive, in this context, as "inappropriate," even though these are completely acceptable forms in non-standard varieties of Spanish. However, this approach had to be adopted if we hoped to establish a clear baseline and potential changes in grammatical mood judgment and use over time. In addition to these two scenarios, on some occasions, learners also altered the structure of the sentence by making the statement positive or by simplifying the syntax and eliminating the relative clause all together, which disallowed the use of past subjunctive. Sometimes, the verb was not uttered or was inaudible due to poor quality recording, leading to no codable data. This resulted in 8% of the EIT data, that were therefore excluded from the final data pool consisting of 1,754 items.

Learners' individual AJT responses were entered based on participants' numeric selection of ratings from the 5-value Likert scale employed (see more details in Appendix B) and were later analyzed separately for grammatical and ungrammatical items. The text box included after each item allowed us to examine if ratings were assigned based on the appropriate or inappropriate use of the past subjunctive, rather than by other morphosyntactic or lexical components of the sentence. This technique was particularly helpful with five items in which learners had assigned an ungrammatical sentence a low value in the scale due to something that was not related to subjunctive/indicative mood use. These five items were excluded from the AJT final data pool, which ended up being 1,915 items.

Learners' individual EIT and AJT responses were coded by the main researcher and another Spanish native speaker fellow researcher, who coded 25% of the data separately to establish inter-rater reliability. Inter-rater agreement was high at 97%, and the few disagreements encountered were discussed and agreed upon by both researchers in a short discussion session.

The EIT data was fitted to a mixed-effects logistic regression and the AJT data was fitted to a mixed effects linear regression model for heritage and L2 learners separately. For all primary analyses, the fixed effects were instruction (explicit, control), and time (pre, post, delayed posttest); participants and items were included as random effect (intercepts and random slopes). Grammatical and ungrammatical items within the AJT and EIT data were analyzed separately. The control group and pretest session data were used as reference levels. Statistical analysis was carried out using R (R Development Core Team, 2023) with the *lme4* package (Bates et al., 2015) and keeping the maximal random effect structure whenever possible (following Barr, 2013). Pairwise comparisons were obtained using the *emmeans* package (Lenth et al., 2018). Data processing and visualization were conducted using the *tidyverse* package (Wickham et al. 2019). Finally, further analysis was run on mean scores to determine the effect size (Cohen's d) for the AJT and EIT tasks, for the experimental and control group, with grammatical and ungrammatical items.

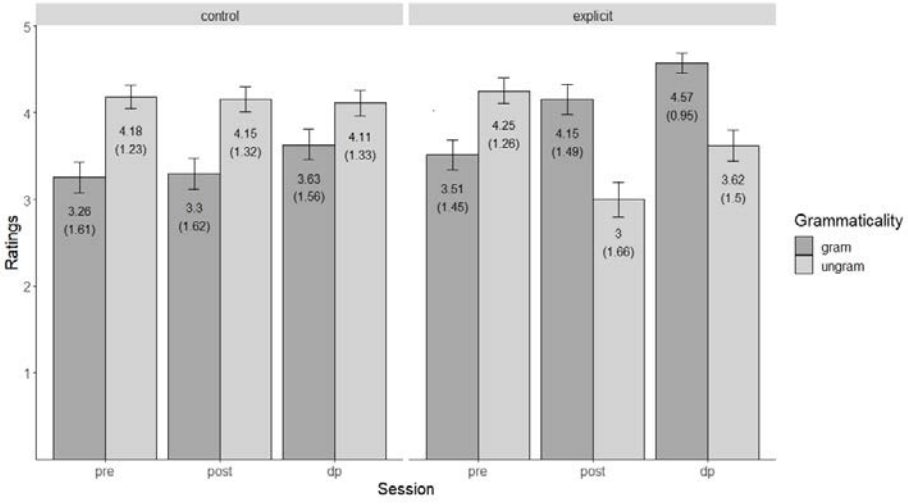
Results

L2 Learners' AJT Results

To explore if explicit language instruction led to positive learning gains for L2 learners as measured by a highly explicit written assessment task, we calculated the average ratings for grammatical and ungrammatical items and compared them by group (experimental vs. control) and session (pre-, post-, and delayed posttest). Average L2 AJT ratings per session and group for both grammatical and ungrammatical items can be found in Figure 3.

The output of the mixed effects linear regression model containing L2 learners' AJT ratings for grammatical items (see Appendix C) yielded a significant interaction of time and instruction at the immediate posttest, $p = 0.05$, $d = 0.42$, as well as a significant interaction of instruction with time at the delayed posttest, $p = 0.03$, $d = 0.33$. Pairwise comparisons run to explore these significant interactions yielded no significant effect of time for the control group, from pre- to posttest, $p = 0.98$, $d = 0.01$, from post to delayed posttest, $p = 0.31$, $d = 0.16$ or from pre- to delayed posttest, $p = 0.21$, $d = 0.10$. With the instructed L2 group, a significant effect of time was found from pre- to posttest, $p = 0.01$, $d = 0.34$, and from pre- to delayed posttest, $p = 0.00$, $d = 0.70$, but there was no significant effect from post- to delayed posttest, $p = 0.19$, $d = 0.27$. Overall, these results evidenced that the ratings for grammatical sentences for L2 learners in the instructed group improved significantly as the result of instruction and were maintained over time. Participants in the control group did not show improvement over time as expected.

Figure 3
Mean L2 AJT ratings for grammatical and ungrammatical items for L2 instructed and control group (SDs in parenthesis)



The output of the mixed effects linear regression model containing L2 learners' AJT ratings for ungrammatical items (see Appendix C) also yielded a significant

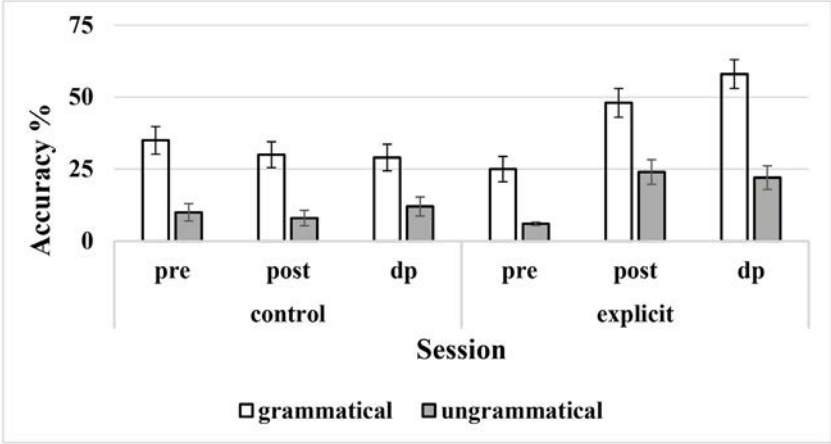
interaction of instruction with time at the delayed posttest, $p = 0.03$, $d = 0.51$, and of instruction with time at the immediate posttest, $p = 0.00$, $d = 0.61$. Pairwise comparisons of group type (experimental vs. control) by time revealed no significant effect of time for the control group, from pre- to posttest, $p = 0.97$, $d = 0.01$, from post- to delayed posttest, $p = 0.79$, $d = 0.02$, or from pre- to delayed posttest, $p = 0.90$, $d = 0.04$, evidencing no improvement with ungrammatical times over time for L2 learners in the control group. On the other hand, L2 learners in the instructed group displayed a significant effect of time from pre- to posttest, $d = 0.68$, from post- to delayed posttest, $d = 0.36$, and also from pre- to delayed posttest, $p = 0.04$, $d = 0.31$, showing that instructed L2 learners' ratings for ungrammatical items decreased upon receiving instruction, despite ratings starting to pick back up two weeks post-instruction.

L2 Learners' EIT Results

To examine if explicit language instruction led to positive learning gains for L2 learners as measured by a structured oral assessment task, we coded L2 learners' responses to EIT items for accuracy with grammatical and ungrammatical items and compared them by group (experimental vs. control) and session (pre-, post-, and delayed posttest).

L2 EIT accuracy percentages per session and group for both grammatical and ungrammatical items can be found in Figure 4. The output of the mixed effects binomial logistic regression model containing L2 learners' EIT accuracy scores for grammatical items (see Appendix C) yielded a significant interaction of instruction with time at the delayed posttest, $p = 0.00$, $d = 0.58$, and of instruction with time at the immediate posttest, $p = 0.01$, $d = 0.68$. Pairwise comparisons of instruction by time yielded no significant effect of time for the control group, from pre- to posttest, $p = 0.75$, $d = 0.08$, from post- to delayed posttest, $p = 0.99$, $d = 0.01$, or from pre- to delayed posttest, $p = 0.67$, $d = 0.10$, evidencing no improvement on accuracy with grammatical items over time for L2 learners in the control group. On the contrary, the instructed L2 learners showed a positive improvement in EIT accuracy with grammatical items, as evidenced by the significant effect of time for the experimental group found, from pre- to posttest, $p = 0.02$, $d = 0.39$, and from pre- to delayed posttest, $p = 0.00$, $d = 0.56$, but the effect from post- to delayed posttest was not significant, estimate = 0.38, $SE = 0.33$, $z = 1.15$, $p = 0.47$ $d = 0.16$. These results confirmed once more that instructed L2 learners' ratings for grammatical sentences improved upon receiving instruction.

Figure 4
Accuracy percentages for grammatical and ungrammatical items for the L2 instructed and control group



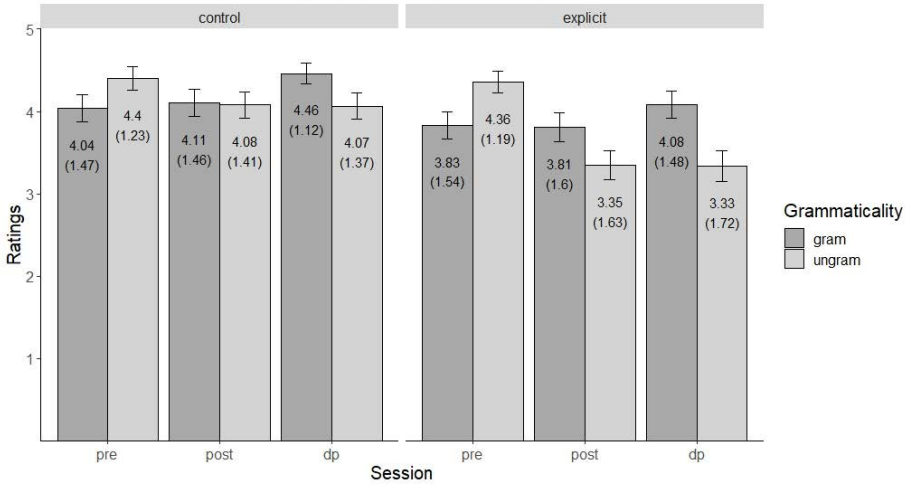
The output of the mixed effects binomial logistic regression model containing L2 learners' EIT accuracy scores for ungrammatical items (see Appendix C) yielded only a significant interaction of instruction with time at the immediate posttest, $p = 0.01$, $d = 0.63$. Pairwise comparisons revealed no significant effect of time for the control group, from pre- to posttest, $p = 0.75$, $d = 0.05$, from post- to delayed posttest, $p = 0.99$, $d = 0.10$, or from pre- to delayed posttest, $p = 0.67$, $d = 0.05$, showing that accuracy on EIT grammatical items did not improve over time for L2 learners in the control group. On the contrary, instructed L2 learners exhibited positive gains in accuracy with ungrammatical items as shown by a significant effect of time for the experimental group, from pre- to posttest, $p = 0.01$, $d = 0.43$, from pre- to delayed posttest, $p = 0.02$, $d = 0.40$. No significant effect of time was found from post- to delayed, $p = 0.94$, $d = 0.03$.

Heritage Learners' AJT Results

Similar to the approach employed with L2 learners, potential learning gains for HLs–as measured by a written assessment task–were examined using the average ratings for grammatical and ungrammatical items that we later compared by group (experimental vs. control) and session (pre-, post-, and delayed posttest). HL AJT average ratings per session and group for both grammatical and ungrammatical items can be found in Figure 5.

Figure 5

Mean HL AJT ratings for grammatical and ungrammatical items for HL instructed and control group (SDs in parenthesis)



The output of the mixed effects linear regression model containing HLs' AJT ratings for grammatical items (see Appendix D) did not yield any significant main effects or interactions, showing that HLs' AJT ratings for grammatical items did not change over time.

The output of the mixed effects linear regression model containing HLs learners' AJT ratings for ungrammatical items (see Appendix D) yielded a significant main effect of instruction, $p = 0.00$, $d = 0.51$, showing that overall ratings for HLs in the explicit group were lower, when compared to those of HLs in the control group. In addition, the model yielded a significant interaction of instruction with time at the pretest stage, $p = 0.01$, $d = 0.33$. Pairwise comparisons revealed a significant main effect of time for instructed HLs from pre to posttest, $p = 0.00$, $d = 0.37$ and from pre- to delayed posttest, $p = 0.00$, $d = 0.33$, which was not present from post- to delayed posttest, $p = 0.99$, $d = 0.02$, evidencing that instructed HLs' ratings for ungrammatical sentences decreased overtime. On the other hand, HLs in the control group did not show significant changes from pre- to posttest, $p = 0.17$, $d = 0.22$, from pre-to delayed posttest, $p = 0.53$, $d = 0.08$, or from post- to delayed posttest, $p = 0.78$, $d = 0.32$. In other words, the ratings from HLs in the control group did not change over time, suggesting that it was instruction that helped instructed HLs.

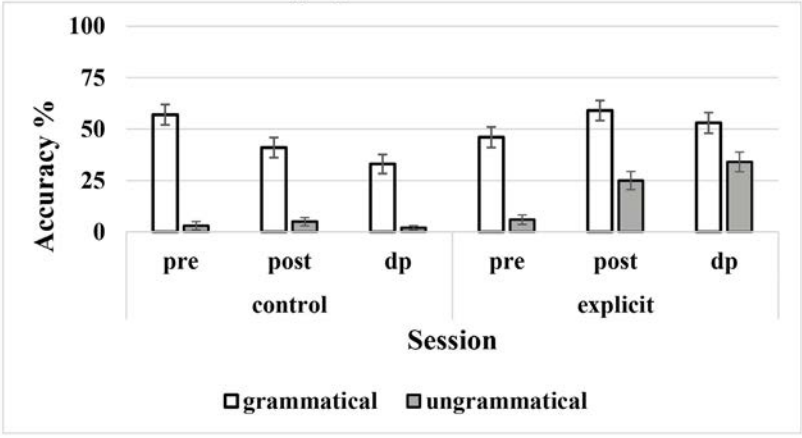
Heritage Learners' EIT Results

To explore potential learning gains for HLs–as measured by a structured oral assessment task–HLs' responses to EIT items were coded for accuracy with grammatical and ungrammatical items and compared by group (experimental vs. control) and session (pre-, post-, and delayed posttest). See descriptive statistics in Figure 6.

The output of the mixed effects binomial logistic regression model containing HLs' EIT accuracy scores for grammatical items (see Appendix D) yielded a

significant main effect of time at the posttest, $p = 0.05$, $d = 0.31$, and delayed posttest, $p = 0.00$, $d = 0.29$, showing that accuracy was lower at the posttest and delayed posttest stage when compared to the pretest. Furthermore, the model yielded a significant interaction of instruction with time at the posttest, $p = 0.01$, $d = 0.32$, and at the delayed posttest, $p = 0.00$, $d = 0.27$. Pairwise comparisons by time and instruction revealed that for HLs in the control group, accuracy for grammatical items decreased significantly from pre- to delayed posttest, $p = 0.01$, $d = 0.35$, but no significant changes were observed from pre- to posttest, $p = 0.13$, $d = 0.39$, or from post- to delayed posttest, $p = 0.57$, $d = 0.13$ Instructed HLs did not show any significant changes in accuracy with grammatical items from pre- to posttest, $p = 0.24$, $d = 0.42$, from pre- to delayed posttest, $p = 0.62$, $d = 0.22$, or from post- to delayed posttest, $p = 0.73$, $d = 0.09$. Interestingly, and against any predictions, HLs in the control group showed a decrease in accuracy with grammatical sentences over time, and no positive effects were found for instructed HLs.

Figure 6
Accuracy percentages for grammatical and ungrammatical items for the HL instructed and control group



The output of the mixed effects binomial logistic regression model containing HLs' EIT accuracy scores for ungrammatical items (see Appendix D) yielded a significant main interaction of instruction with time at the delayed posttest, $p = 0.03$, $d = 0.57$. Pairwise comparisons of instruction by time revealed no significant effect of time for the control group, from pre- to posttest, $p = 0.89$, $d = 0.08$, from post- to delayed posttest, $p = 0.73$, $d = 0.12$, or from pre- to delayed posttest, $p = 0.90$, $d = 0.04$. On the other hand, instructed HLs showed a significant increase in accuracy with ungrammatical items from pre- to posttest, $p = 0.00$, $d = 0.45$, and from pre- to delayed posttest, $p = 0.02$, $d = 0.63$. There was no significant effect of time from post- to delayed posttest, $p = 0.45$, $d = 0.15$. This time, in line with our expectations, instructed HLs showed learning gains with ungrammatical sentences over time and HLs in the control did not.

Brief Summary of Results

AJT results suggest that L2 participants in the control group did not evidence learning gains at any point in time, validating the current study's methodology, whereas instructed L2 learners showed an improvement in ratings with grammatical and ungrammatical items from pre- to immediate as well as from pre- to delayed posttest with grammatical sentences, and also from post- to delayed posttest with ungrammatical sentences. These findings were similar to those in the EIT, where instructed L2 learners showed a significant increase in accuracy from pre- to immediate and from pre- to delayed posttest–but not from immediate posttest to delayed posttest–with grammatical and ungrammatical items. Once again, the control group did not display learning gains over time.

Instructed heritage learners' results were also consistent by task but differed from L2 learners' results based on the grammaticality condition. Instructed HLs improved on their ratings of ungrammatical sentences from pre- to immediate posttest and from pre- to delayed posttest only with ungrammatical sentences. EIT results mirrored these findings with a significant increase in accuracy from pre- to immediate posttest, and from pre- to delayed posttest also with only ungrammatical items. No improvement was observed for instructed HLs with grammatical items. Surprisingly, a rather unpredictable pattern emerged with HLs in the control group, who showed a significant decrease in accuracy from pre- to delayed posttest only in the EIT only with grammatical items.

Discussion

The first two research questions in the present study were set to determine if explicit language instruction in the form of PI is beneficial for both types of learners (heritage and L2), considering potential learning gains in both assessment tasks. At the pretest stage, L2 learners showed more acceptability of the use of the imperfect of indicative in sentences containing a matrix clause with a nonspecific referent (M = 4.2), than of the past subjunctive (M = 3.3). L2 learners also produced either the imperfect of indicative or the preterite as alternatives to the expected imperfect subjunctive in the EIT, when repeating sentences with a non-specific referent in the matrix clause. As predicted, learners who received instruction exhibited learning gains in both assessment tasks showing significant improvement immediately after receiving instruction, and in a two-week delayed posttest, in contrast with the control group who did not show any signs of learning. This positive outcome occurred with both grammatical and ungrammatical sentences and is the result of a one-time short instructional intervention. These findings are also consistent with previous research that examined the benefits of explicit language instruction, particularly with PI studies that focused on the Spanish subjunctive (Farley & McCollum, 2004; Fernández, 2008; Kirk, 2013, among others), which as previously stated, is difficult to acquire in a naturalistic manner due to a series of intra- and extra-linguistic factors. The present study provides further evidence that full PI (EI + SI) is an effective instructional intervention for L2 learners that can lead to long-lasting learning gains in metalinguistic written and productive knowledge of the imperfect subjunctive in non-existential clauses.

Instruction was also beneficial for HLs. At the pre-test stage, HLs also showed strong acceptance of the imperfect indicative (M = 4.3) with sentences that contained a matrix clause with a specific and non-specific referent, but also displayed high acceptance of the imperfect subjunctive (M = 4 out of 5) in sentences where its used was expected. This pattern was complemented with a 42% accurate use of subjunctive when repeating sentences that contained a non-specific referent in the matrix clause. This high acceptance of both alternatives could suggest that use of subjunctive mood in this context was not categorical for HLs, supporting the results from studies that find a reduced use of subjunctive mood among second generation Spanish heritage speakers (Viner, 2018). Nonetheless, and unlike L2 learners, HLs used almost categorically the imperfect indicative–when they did not opt for the imperfect subjunctive–with sentences that had a matrix clause with a nonspecific referent in both tasks. This could be interpreted as HLs having developed a rule that calls for the use of either the imperfect indicative or subjunctive in this context, whereas L2 learners did not have a default grammatical form that they consistently used for this linguistic context. This pattern of a reduced but present use of subjunctive (compared to monolingually raised Spanish speakers residing in the US) is consistent with research on variationist approaches to heritage language acquisition (Silva-Corvalán; Viner, 2018), and also highlights the differences in starting point for these two groups of learners, whose prior learning experience may have likely influenced the underlying grammar they drew on before being exposed to grammar instruction.

In contrast with L2 learners, instructed HLs only exhibited positive changes over time with ungrammatical items. Instructed HLs' ratings for ungrammatical sentences decreased significantly over time, and they showed a significant increase in the use of imperfect subjunctive in sentences that contained a matrix clause with a non-specific reference in the EIT upon receiving instruction, compared to participants in the control group, who did not experience learning gains. The AJT results from the present study are not consistent with Potowski et al. (2009) who found that explicit instruction is not able to alter HLs' acceptability perception of sentences' grammaticality, and with Montrul and Bowles (2010), whose participants experienced no learning gains with ungrammatical sentences. With regard to EIT results, our findings contrast slightly with those of Fernández Cuenca and Bowles (2022), who found that HLs receiving explicit instruction (identical module to the one in the present study) improved with grammatical and ungrammatical sentences in the EIT post-instruction. In the present study, instructed HLs only exhibited learning gains with ungrammatical sentences in the EIT. A possible explanation for this difference could be the approach used to analyze the data. Whereas the analyses in the present study were conducted separately for grammatical and ungrammatical sentences, Fernández Cuenca and Bowles (2022) explored accuracy with both types of items together. Our AJT findings match those of Fernández Cuenca and Bowles (2022) with HLs only exhibiting a decrease in acceptability with ungrammatical items. Overall, these studies and the present study provide evidence that explicit language instruction is beneficial for adult HLs who still displayed learning gains two weeks post instruction.

Our third research question aimed to take a closer look at learning gains by outcome measure. As formerly discussed, heritage and L2 learners differ in the way they acquire the Spanish language, and this may have affected how learners approached completion of these assessment tasks. Previous research suggests that HLs are less familiar with written focus on form tasks, such as GJTs, due to acquiring Spanish in a naturalistic fashion that did not bring their attention to form until later in life when they enrolled in Spanish language courses (Bowles, 2011). On the contrary, these studies also point to Spanish L2 learners being more comfortable completing focus on form tasks due to their robust experience with formal language instruction. Our findings do not necessarily match this task modality distinction. L2 learners evidenced learning gains similarly in both assessment tasks, and HLs also did so across assessment tasks, even if positive learning outcome were only present with ungrammatical sentences. We did not find HLs' learning gains to be greater in the oral than in the written task or observed L2 learners performing better in written than oral tasks, in contrast with previous studies (Sanz, 1997; Torres, 2022). Therefore, we conclude that task modality, at least in the current study, did not modulate learning gains for heritage and L2 learners differently.

What we did find is an interesting pattern with the grammaticality variable that we did not anticipate, despite it already being somewhat present in previous studies (e.g., Montrul & Bowles, 2010). As Montrul and colleagues (2014) wisely pointed out, some studies that sought to compare heritage and L2 learners' performance with different types of tasks differing in modality, often confounded modality with the type of knowledge (explicit or implicit) these learners had to rely on to complete the language task. The two tasks employed as assessment tasks in the present study raise the notion of grammaticality despite it being presented in different modalities. EITs have been traditionally associated with measuring implicit knowledge, and untimed GJTs are associated with explicit knowledge (Bowles, 2011; Ellis, 2005). The EIT used in the present study is slightly different than the type of EIT used in the cited previous studies in that it includes a focus-on-form approach that alerts participants to pay attention to form in the sentences they listen to and asks them to correct such mistakes when participants proceed to repeat the stimulus sentences. Therefore, one could easily say that the EIT task employed in the present study does not fit the criteria of the EITs used in previous research and both tasks are characterized for adopting an explicit approach.

Nevertheless, one important point we can draw on from this research is their findings with regard to the type of stimulus sentence. Vafaee and colleagues (2017) found that ungrammatical sentences in a GJT are a good measure of explicit knowledge, at least for L2 learners. Similarly, Gutiérrez (2013) found evidence that L2 learners draw on different types of knowledge, which one could interpret as explicit vs. implicit knowledge, when rating grammatical and ungrammatical sentences in GJTs. One possible explanation could be that instruction provided both types of learners with explicit knowledge that they were able to apply in discrete focus-on-form tasks such as the ones used in our study. As we know, L2 learners possess a dynamic linguistic system that is constantly being restructured based on negative or positive feedback, similar to the type of feedback received in an instructional setting. Therefore, we could speculate that since L2 learners are

familiar with explicit instruction and focus-on-form tasks that tap into explicit knowledge, the notion of grammaticality did not present a problem for them and they were successful at identifying and correcting grammatical and ungrammatical sentences, upon receiving and practicing the rule. On the other hand, this may have posited a problem for HLs who most likely had implicitly acquired the imperfect indicative as the most suitable form (when the imperfect subjunctive was not chosen) for this linguistic context. In other words, HLs had a strong grammar in place that was acquired largely implicitly, given that the average number of Spanish formal education for this group was 4.5 years, and replacing this established pattern would likely require more than one instructional session. Unfortunately, this argument falls short if we consider that HLs in the control group displayed a significant decrease in accuracy with grammatical items from pre- to posttests in the EIT, suggesting that unguided exposure to sentences that incorporated past subjunctive in this linguistic context was enough to destabilize the form-meaning connections that HLs had in place before engaging with formal language instruction. What is interesting here is that this only occurred in the EIT where the stimulus sentences were presented aurally and not in writing, which brings up the question of the modality in which the stimuli are presented as it being an important factor at play here. In fact, this observation was already noted by Torres (2022) and future studies should take the necessary steps to separate the modality of the stimulus presentation and the modality in which participants are required to respond.

Pedagogical Implications

An important takeaway from this study should be that whereas L2 learners responded homogeneously to instruction, HLs did not. Both groups of learners showed positive learning gains, but the factor of grammaticality in the stimulus sentences modulated the learning outcomes of the HLs. The notion of grammaticality is highly associated with explicit knowledge, which is a type of knowledge heritage learners do not usually resort to, given that they acquired their heritage language implicitly as children. As heritage learners become more familiar with formal language instruction, it is very possible that concepts such as grammaticality or the approach to focus-on-form will become more natural to them, but patience is needed, and them and their performance should not be directly compared to that of L2 learners. Studies such as the ones described and presented here only incorporate one instructional session and perhaps multiple sessions will be necessary for HLs.

Another point raised in the discussion is the modality of the task, or more precisely, the modality in which the stimuli is presented in the assessment tasks used to measure learning gains could have impacted HLs performance learning in this study. The fact that unguided exposure to aural grammatical and ungrammatical sentences appeared to have a negative effect on HLs' competence of mood in this linguistic context, raises the question of how HLs' already existing grammar interacts with language instruction in adulthood, especially when the language construction being targeted is a vulnerable one that shows highly variable acquisitional attainment among second generation heritage speakers (Viner, 2018). Recent studies are starting to decipher how psychosocial and biographical factors

such as willingness to communicate in the heritage language, acquisitional profile, or motivation to become a better Spanish speaker, can affect heritage learners' performance completing language tasks (Torres et al., 2019). The HL profile does not fit the uniform profile of the learner that has commonly populated our language classes. The same way language changes over time, so does the student profile and we, as language educators, should become aware and make the necessary changes to update our language teaching practices.

As previously mentioned, the field of instructed heritage language acquisition is just starting to grow, and a lot of questions remain unanswered. For the time being, Spanish language instructors should bear in mind that HLs are different in the way they acquired Spanish, and they, unlike L2 learners, come to classroom with an established grammar that was acquired largely implicitly and orally before the puberty years. This is bound to interact with the way they approach language learning in our classes and most assessment tasks commonly used to measure language proficiency and longitudinal language learning do not take this into consideration.

A final point that deserves to be mentioned here and that applies to heritage language grammar instruction in general, is that no HL should be shamed for not performing well in these tasks. The language experience HLs bring to the classroom is as valid as that of L2 learners, and it is our job as educators to ensure that HLs leave our classroom feeling empowered for having learned how to label concepts they already know, and for learning how to best use different language forms depending on the context (formal or informal) in which they plan to communicate a message.

References

Adrada-Rafael, S. (2017). Processing the Spanish imperfect subjunctive: Depth of processing under different instructional conditions. *Applied Psycholinguistics, 38*(2), 477-508.

Alarcón, I. V. (2011). Spanish gender agreement under complete and incomplete acquisition: Early and late bilinguals' linguistic behavior within the noun phrase. *Bilingualism: Language and Cognition, 14*(3), 332-350.

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in psychology, 4*, 328.

Bayram, F., Prada, J., Pascual y Cabo, D., & Rothman, J. (2016). Why should formal linguistic approaches to heritage language acquisition be linked to heritage language pedagogies. In P. P. Trifonas & T, Aravossitas (Eds.), *The Handbook of research and practice in heritage language education* (pp.187-206). *Springer*.

Bayram, F., Rothman, J., Di Pisa, G., & Slabakova, R. (2020). Current trends and emerging methodologies in charting heritage language bilingual grammars. In S. Montrul & M. Polinsky (Eds.), *The Cambridge Handbook of Heritage Languages and Linguistics* (pp. 545-578). Cambridge University Press.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Beaudrie, S. (2012). Research on university-based Spanish heritage language programs in the United States: The current state of affairs. In S. Beaudrie, S. & M. Fairclough (Eds.), *Spanish as a heritage language in the United States: The state of the field* (pp. 203-221). Georgetown University Press.

Beaudrie, S. M., Ducar, C., & Potowski, K. (2014). *Heritage language teaching: Research and practice* (pp. 32-53). New York: McGraw-Hill Education Create.

Beaudrie, S. M. & Holmes, B. C. (2022). The Differential Effects of Three Types of Form-Focused, Computer-Based Grammar Instruction: The Case of Receptive Heritage Learners. In M. A. Bowles (Ed.), *Outcomes of University Spanish Heritage Language Instruction*. Georgetown University Press.

Benati, A. (2001). A comparative study of the effects of processing instruction and output-based instruction on the acquisition of the Italian future tense. *Language Teaching Research, 5*(2), 95-127.

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.

Blas Arroyo, J., & Porcar Miralles, M. (1997). Aproximación sociolingüística al fenómeno de la neutralización modal en las comunidades de habla castellonenses. *Sintagma: Revista de lingüística, 9*, 27–45.

Borgonovo, C., Bruhn de Garavito, J. & Prévost, P. (2005). Acquisition of mood distinctions in L2 Spanish. In A. Brugos, M. R. Clark-Cotton, & S. Ha (Eds.), *Proceedings of the 29th annual Boston University Conference on Language Development* (pp. 97-108). Cascadilla Press.

Bowles, M. A. (2011a). Measuring implicit and explicit linguistic knowledge: What can heritage learners contribute? *Studies in Second Language Acquisition 33*(2), 247-271.

Bowles, M. A. (2011b). Exploring the role of modality: L2-heritage learner interactions in the Spanish language classroom. *Heritage Language Journal, 8*(1), 30-65.

Bowles, M. A., Adams, R. J., & Toth, P. D. (2014). A comparison of L2–L2 and L2–heritage learner interactions in Spanish language classrooms. *The Modern Language Journal, 98*(2), 497-517.

Bowles, M. & Henshaw, F. (2015, October). What type(s) of knowledge do PI and SI result in? [Paper presentation]. Second Language Research Forum, Atlanta, GA, United States.

Bowles, M.A. (2018). Outcomes of Classroom Spanish Heritage Language Instruction. In K. Potowski (Ed.), *The Routledge Handbook of Spanish as a Heritage Language*. Routledge.

Bowles, M. A., & Torres, J. (2021). Instructed heritage language acquisition. In M. Polinsky & S. Montrul (Eds.), T*he Cambridge Handbook of Heritage Languages and Linguistics* (pp. 826-850). Cambridge University Press.

Bowles, M. A. (2022). *Outcomes of University Spanish Heritage Language Instruction in the United States*. Georgetown University Press.

Carreira, M., & Kagan, O. (2011). The results of the National Heritage Language Survey: Implications for teaching, curriculum design, and professional development. *Foreign language annals, 44*(1), 40-64.

Carreira, M. & Chik, C.H. (2018). Differentiated teaching: a primer for heritage and mixed classes. In K. Potowski (Ed.), *The Routledge Handbook of Spanish as a Heritage Language*. Routledge.

Chomón Zamora, C. (2022). The Secret Is in The Processing: Categorizing how Heritage Learners of Spanish Process. In M. A. Bowles (Ed.), *Outcomes of University Spanish Heritage Language Instruction*. Georgetown University Press.

Collentine, J. (1995). The development of complex syntax and mood-selection abilities by intermediate-level learners of Spanish. *Hispania, 78*(1), 123–136.

Collentine, J. G. (1997). The effects of irregular stems on the detection of verbs in Spanish. *Spanish Applied Linguistics, 1*(1), 3–23.

Collentine, J. (1998). Processing instruction and the subjunctive. *Hispania, 81*(3), 576-587.

Correa, M. (2011). Heritage language learners of Spanish: What role does metalinguistic knowledge play in their acquisition of the subjunctive? In Ortiz-López, L. (ed), *Selected Proceedings of the 13th Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla Proceedings Project, pp. 128-138.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in second language acquisition, 27*(2), 141-172.

Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics, 27*(3), 464–91.

Farley, A. (2001). Authentic processing instruction and the Spanish subjunctive. *Hispania, 84*(2), 289-99.

Farley, A. P. (2004a). The Relative Effects of Processing Instruction and Meaning-Based. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 183-205). Lawrence Erlbaum Associates.

Farley, A. (2004b). Processing instruction and the Spanish subjunctive. Is explicit information needed? In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 183-205). Lawrence Erlbaum Associates.

Farley, A., & McCollam, K. (2004). Learner readiness and L2 production in Spanish: Processability theory on trial. *Estudios de Lingüística Aplicada, 40*, 47-69.

Fernández, C. (2008). Reexamining the role of explicit information in processing instruction. *Studies in second language acquisition, 30*(3), 277-305.

Fernández Cuenca, S. (2019). *The Effects of Language Instruction on L2 Language Learners' Processing and Learning Outcomes* [Unpublished doctoral dissertation]. The University of Illinois at Champaign-Urbana.

Fernández Cuenca, S. & Bowles, M. (2022). What type of knowledge do implicit and explicit heritage language instruction result in? In M. A. Bowles (Ed.), *Outcomes of University Spanish Heritage Language Instruction*. Georgetown University Press.

Geeslin, K. L., & Gudmestad, A. (2008). Comparing interview and written elicitation tasks in native and nonnative data: Do speakers do what we think they do. In J. Bruhn de Garavito & E. Valenzuela (Eds.), *Selected proceedings of the 10th Hispanic linguistics symposium* (pp. 64-77). Cascadilla Proceedings Project.

Giancaspro, D. (2019a). Over, under and around: Spanish heritage speakers' production (and avoidance) of subjunctive mood. *Heritage Language Journal, 16*(1), 44-70.

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*(3), 69.

Gudmestad, A. (2006). L2 variation and the Spanish subjunctive: Linguistic features predicting mood selection. In *Selected proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 170-184). Cascadilla Proceedings Project.

Gudmestad, A. (2012). Acquiring a variable structure: An interlanguage analysis of second language mood use in Spanish. *Language Learning, 62*(2), 373-402.

Gutiérrez, X. (2013). The construct validity of grammaticality judgment tests as measures of implicit and explicit knowledge. *Studies in second language acquisition, 35*(3), 423-449.

Henshaw, F., & Hetrovicz, L. (2021). The Role of Linguistic Background in Interaction. *Hispania, 104*(1), 43-59.

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, 1*, 135–163.

Jegerski, J. (2018). Psycholinguistic Perspective on Heritage Spanish. In K. Potowski (Ed.), *The Routledge Handbook of Spanish as a Heritage Language*. Routledge.

Kang, E. Y., Sok, S., & Han, Z. (2019). Thirty-five years of ISLA on form-focused instruction: A meta-analysis. *Language Teaching Research, 23*(4), 428-453.

Kanwit, M., & Geeslin, K. L. (2018). Exploring lexical effects in second language interpretation: The case of mood in Spanish adverbial clauses. *Studies in Second Language Acquisition, 40*(3), 579-603.

Kirk, R. W. (2013). The effects of processing instruction with and without output: Acquisition of the Spanish subjunctive in three conjunctional phrases. *Hispania, 96*(1), 153-169.

Lenth, R., Singmann, H., Love J., Buerkner, P.,&Herve, M. (2018). Emmeans: Estimated Marginal Means, aka Least-Squares Means. https://cran.r-project.org/web/packages/emmeans/index.html.

Lynch, A. E. (1999). *The subjunctive in Miami Cuban Spanish: Bilingualism, contact, and language variability* [Unpublished doctoral dissertation]. University of Minnesota.

Massery, L. A. (2009). *Syntactic development of the Spanish subjunctive in second language acquisition: Complement selection in nominal clauses* [Unpublished doctoral dissertation]. University of Florida, Gainesville.

McNulty- Díaz, E. (2017). The order of explicit information in processing instruction. *AppliedLanguage Learning, 27*(1-2), 41-71.

Montrul, S., & Slabakova, R. (2003). Competence similarities between native and near-native speakers: An investigation of the preterite/imperfect contrast in Spanish. *Studies in Second Language Acquisition, 25*(3), 351-398.

Montrul, S. (2007). Interpreting mood distinctions in Spanish as a heritage language. In K. Potowski & R. Cameron (Eds.), *Spanish in Contact: Policy, Social and Linguistic Inquiries* (pp. 23–40). John Benjamins.

Montrul, S., & Bowles, M. (2010). Is grammar instruction beneficial for heritage language learners? Dative case marking in Spanish. *Heritage Language Journal, 7*(1), 47-73.

Montrul, S. (2011). Morphological Errors in Spanish Second Language Learners and Heritage Speakers. *Studies in Second Language Acquisition, 33*(2), 163–192.

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.

Montrul, S., & Polinsky, M. (2011). Why not heritage speakers? *Linguistic Approaches to Bilingualism, 1*(1), 58-62

Montrul, S., Davidson, J., De La Fuente, I., & Foote, R. (2014). Early language experience facilitates the processing of gender agreement in Spanish heritage speakers. *Bilingualism: Language and Cognition, 17*(1), 118-138.

Montrul, S., & Bowles, M. (2017). Instructed heritage language acquisition. In S. Loewen (Ed.), *The Routledge handbook of instructed second language acquisition* (pp. 488-502). Routledge.

Morgan-Short, K., & Bowden, H. W. (2006). Processing instruction and meaningful output-based instruction: effects on second language development. *Studies in Second Language Acquisition, 28*(1), 31-65.

Murillo-Medrano, J. (1999). Subjuntivo e indicativo en las oraciones circunstanciales. Káñina, Revista de Artes y Letras de la Universidad de Costa Rica, 23, 143-156.

Oh, J. S., & Au, T. K. F. (2005). Learning Spanish as a heritage language: The role of sociocultural background variables. *Language, Culture and Curriculum, 18*(3), 229-241.

Pascual y Cabo, D., & Rothman, J. (2012). The (Il)logical problem of heritage speaker bilingualism and incomplete acquisition. *Applied Linguistics, 33*(4), 450-455.

Pereira, I. (1996). *Markedness and instructed SLA: An experiment in teaching the Spanish subjunctive* [Unpublished doctoral dissertation]. The University of Illinois at Urbana-Champaign, Champaign-Urbana

Perez-Cortes, S. (2022). On complexity and divergence in heritage language grammars: The case of double mood selection in reported speech contexts. *Studies in Second Language Acquisition, 44* (3), 818-842.

Perez-Cortes, S. (2021). Re-examining the role of type of mood selection in Spanish HSs' subjunctive production. *Linguistic Approaches to Bilingualism, 13*(2), 238-266.

Plonsky, L., & Derrick, D. J. (2016). A meta-analysis of reliability coefficients in second language research. *The Modern Language Journal, 100*(2), 538-553.

Polinsky, M. (2011). Reanalysis in adult heritage language: New evidence in support of attrition. *Studies in second language acquisition, 33*(2), 305-328.

Potowski, K. (2002). Experiences of Spanish heritage speakers in university foreign language courses and implications for teacher training. *Adfl Bulletin, 33*(3), 35-42.

Potowski, K., Jegerski, J., & Morgan-Short, K. (2009). The effects of instruction on linguistic development in Spanish heritage language speakers. *Language Learning, 59*(3), 537-579.

Potowski, K. (2010). *Conversaciones escritas: Spanish composition for heritage and L2 students*. John Wiley & Sons.

Potowski, K., Parada, M., & Morgan-Short, K. (2012). Developing an online placement exam for Spanish heritage speakers and L2 students. *Heritage Language Journal, 9*(1), 51-76.

R Core Team. (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing: https://www.R-project.org/.

Russell, V. (2009). *The effects of processing instruction, structured input, and visual input enhancement on the acquisition of the subjunctive in adjectival clauses by intermediate-level distance learners of Spanish* [Unpublished doctoral dissertation]. University of South Florida.

Sanz, C. (1997). Experimental Tasks in SLA Research: Amount of Production, Modality, Memory, and Production Processes. In A. Perez Leroux & W. R. Glass (Eds.), *Contemporary Perspectives on the Acquisition of Spanish* (pp. 41–56). Cascadilla Press.

Silva-Corvalán, C. (1994). The gradual loss of mood distinctions in Los Angeles Spanish. *Language Variation and Change, 6*(3), 255–272.

Shintani, N. (2015). The effectiveness of processing instruction and production-based instruction on L2 grammar acquisition: A meta-analysis. *Applied Linguistics, 36*(3), 306-325.

Terrell, T., Baycroft, B., & Perrone, C. (1987). The subjunctive in Spanish interlanguage: Accuracy and comprehensibility. In B. VanPatten, T. Dvorak, & J. Lee (Eds.), *Foreign language learning: A research perspective* (pp. 23–48). Cambridge University Press.

Torres, J. (2018). The effects of task complexity on heritage and L2 Spanish development. *Canadian Modern Language Review, 74*(1), 128-152.

Torres, J., Estremera, R., & Mohamed, S. (2019). The Contribution of Psychological and Biographical Variables to Heritage Language Learners' Knowledge of Spanish. *Studies in Second Language Acquisition, 41*(4), 695-719.

Torres, J. (2022). Modality Matters! A look at Task-Based Outcomes. In M. A. Bowles (Ed.), *Outcomes of University Spanish Heritage Language Instruction* (pp. 19-37). Georgetown University Press.

Vafaee, P., Suzuki, Y., & Kachinske, I. (2017). Validating grammaticality judgment tests: Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition, 39*(1), 59–95.

Van Osch, B., & Sleeman, P. (2018). Spanish heritage speakers in the Netherlands: Linguistic patterns in the judgment and production of mood. *International Journal of Bilingualism, 22*(5), 513-529.

VanPatten, B., & Cadierno, T. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition, 15*(2), 225-243.

VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research*. Ablex.

VanPatten, B. (2015). Foundations of processing instruction. *International Review of Applied Linguistics in Language Teaching, 53*(2), 91-109.

Viner, K. M. (2018). The optional Spanish subjunctive mood grammar of New York City heritage bilinguals. *Lingua, 210-211*, 79-94.

Waltermire, M. (2017). Mood variation after expressions of uncertainty in Spanish. *Studies in Hispanic and Lusophone Linguistics, 10*(2), 349-372.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. A., François, R., Golemund, G., Hayes, A., Lionel, H, Hester, J., Kuhn, M., Lin Pedersen, T., Miller, E., Milton Bache, S., Muller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., … & Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source software, 4*(43), 1686.

Yan, X., Maeda, Y., Lv, J., & Ginther, A. (2016). Elicited imitation as a measure of second language proficiency: A narrative review and meta-analysis. *Language Testing, 33*(4), 497–528.

Appendix A

Sample elicited imitation task

Listen to the following sentences carefully and indicate whether these sentences are consistent with your personal experience by saying "Sí" "No" o "No se aplica" then repeat them out loud in correct Spanish (Sometimes this will mean repeating the sentence exactly as you heard it, and sometimes this will mean changing some part of the sentence).

One more thing…if the statement you hear is negative, but your answer is yes, make sure to repeat the sentence the same way you heard it in the second screen.

Here is an example to clarify this:

Screen 1: you hear: No vivía en un apartamento
Sí" "No" "No se aplica"
you say: Sí
Screen 2: Repite la frase
you say: No vivía en un apartamento

Context for these sentences: Cuando era más joven….. (context)

Practice item:

- No me gustaba pasar tiempo con mis padres

Experimental (ungrammatical)- 4 total	Experimental (grammatical)- 4 total	Fillers (grammatical and ungrammatical) – 8 total
No tenía hermanos que escuchaban mis problemas No leía libros que eran de ficción No tenía una clase que me gustaba No jugaba con niños que eran muy mayores	No usaba excusas que mis padres conocieran No tenía hermanos que interrumpieran mis sesiones de videojuegos No tenía amigos que hablaran chino mandarín No escuchaba música que enojara a mis padres	Tenía muchas fiestas de cumpleaños Mis hermanos jugaba con juguetes peligrosos Pasaba mucho tiempo fuera de casa Mis amigas eran tímidos e inocentes Me peleaba mucho con mis hermanos Mi madre no trabajan muchas horas Decía mentiras que eran inocentes La escuela no era divertido Practicaba deporte con mis amigos después de clase

Appendix B

Sample acceptability judgment task

Read or listen to the following sentences and using the scale below indicate what you think their acceptability is. If you think that a sentence is not totally acceptable, please make the necessary corrections to make it sound better.

Context for these sentences: when I was young…

Example:

1. Las tiendas tenían menos variedad de juguetes

1- totally unacceptable	2 - unacceptable	3 - neutral	4 - acceptable	5 - perfectly acceptable

If you did not click on "perfectly acceptable" please make the necessary changes to make it sound better.

Experimental stimuli

1. En mi escuela no había profesores que hablaran bien español (grammatical)
2. No había gente que quisiera otra línea de tren nueva (grammatical)
3. En la interestatal 57 no había luces que alumbraran la carretera por la noche (grammatical)
4. El profesor no encontró estudiantes que supieran la respuesta a la pregunta (grammatical)
5. Susana no compré un televisor que costaba $ 2.000 dólares en Amazon (ungrammatical)
6. En el centro comercial no había tiendas que tenían productos orgánicos (ungrammatical)
7. De pequeño no tenía amigos que hablaban muchas lenguas diferentes (ungrammatical)
8. Los trabajadores no aceptaron un contrato que cumplía con sus demandas (ungrammatical)

Fillers

1. Las clases de gimnasia en la escuela eran obligatorias (grammatical)
2. En los años 50, la mujer no trabajaban fuera de la casa (ungrammatical)
3. Las fiestas de cumpleaños eran muy importantes cuando era joven (grammatical)
4. Los carreteras de Chicago estaban en construcción el invierno pasado (ungrammatical)
5. Los parques de Chicago estaban más limpios hace 10 años (grammatical)
6. Mi amigo Juan tenían una bicicleta de segunda mano (grammatical)
7. Cuando mis abuelos eran joven la gente se casaba muy pronto (ungrammatical)
8. Hace 20 años las películas no eran tan largos y aburridas (ungrammatical)

Appendix C

Statistical outputs from L2 learners' analyses

Output from mixed effects linear regression model with L2 learners' AJT ratings for grammatical and ungrammatical items

	Estimate	*SE*	*t*	*p*
Grammatical items				
Intercept	3.24	0.20	15.56	0.00
Instruction	0.26	0.28	0.90	0.36
Time (delayed posttest)	0.38	0.22	1.67	0.09
Time (posttest)	0.04	0.22	0.19	0.84
Instruction x Time (delayed posttest)	0.68	0.31	2.15	0.03
Instruction x Time (posttest)	0.59	0.31	1.90	0.05
Ungrammatical items				
Intercept	4.13	0.20	19.97	0.00
Instruction	0.08	0.27	0.31	0.75
Time (delayed posttest)	0.09	0.21	0.42	0.66
Time (post)	-0.04	0.20	-0.22	0.81
Instruction x Time (delayed posttest)	-0.62	0.28	-2.17	0.03
Instruction x Time (posttest)	-1.25	0.29	-4.30	0.00

Pairwise comparisons:

Grammatical items		
	Control group	Instructed group
Pre to posttest	estimate = -0.04, *SE* = 0.22, *t* = -0.19, *p* = 0.98, *d* = 0.01	estimate = -0.64, *SE* = 0.22, *t* = -2.84, *p* = 0.01, *d* = 0.34
Post to delayed posttest	estimate = 0.34, *SE* = 0.23, *t* = 1.46, *p* = 0.31, *d* = 0.16	estimate = 0.42, *SE* = 0.24, *t* = 1.72, *p* = 0.19, *d* = 0.27
Pre to delayed posttest	estimate = -0.38, *SE* = 0.22, *t* = -1.67, *p* = 0.21, *d* = 0.10	estimate = -1.06, *SE* = 0.23, *t* = -4.46, *p* = 0.00, *d* = 0.70
Ungrammatical items		
	Control group	Instructed group
Pre to posttest	estimate = 0.04, *SE* = 0.20, *t* = 0.22, *p* = 0.97, *d* =0.01	estimate = 1.29, *SE* = 0.20, *t* = 6.25, *p* = 0.00, *d* = 0.68
Post to delayed posttest	estimate = 0.13, *SE* = 0.21, *t* = 0.63, *p* = 0.79, *d* = 0.02	estimate = 0.76, *SE* = 0.23, *t* = 3.27, *p* = 0.00, *d* = 0.36
Pre to delayed posttest	estimate = -0.09, *SE* = 0.21, *t* = -0.42, *p* = 0.90, *d* = 0.04	estimate = 0.53, *SE* = 0.22, *t* = 2.38, *p* = 0.04, *d* = 0.31

Output from mixed effects binomial logistic regression model with L2 learners' EIT accuracy responses for grammatical and ungrammatical items

	Estimate	*SE*	*t*	*p*
Grammatical items				
Intercept	-0.60	0.25	-2.31	0.02
Instruction	-0.47	0.38	-1.22	0.22
Time (delayed posttest)	-0.30	0.35	-0.85	0.39
Time (posttest)	-2.26	0.36	-0.72	0.47
Instruction x Time (delayed posttest)	1.70	0.50	3.36	0.00
Instruction x Time (posttest)	1.27	0.53	2.37	0.01
Ungrammatical items				
Intercept	-2.23	0.37	-6.01	0.00
Instruction	-0.58	0.63	-0.91	0.36
Time (delayed posttest)	0.29	0.54	0.53	0.59
Time (post)	-0.26	0.56	-0.46	0.64
Instruction x Time (delayed posttest)	1.23	0.81	1.52	0.12
Instruction x Time (posttest)	1.91	0.80	2.36	0.01

Pairwise comparisons:

Grammatical items		
	Control group	Instructed group
Pre to posttest	estimate = 0.26, *SE* = 0.36, z = 0.72, *p* = 0.75, *d* = 0.08	estimate = -1.00, *SE* = 0.39, *z* = -2.58, *p* = 0.02, *d* = 0.39
Post to delayed posttest	estimate = -0.04, *SE* = 0.35, *z* = -0.11, *p* = 0.99, *d* = 0.01	estimate = -1.39, *SE* = 0.35, *z* = -3.91, *p* = 0.00, *d* = 0.56
Pre to delayed posttest	estimate = 0.30, *SE* = 0.35, *z* = 0.85, *p* = 0.67, *d* = 0.10	estimate = 0.38, *SE* = 0.33, *z* = 1.15, *p* = 0.47 *d* = 0.16.
Ungrammatical items		
	Control group	Instructed group
Pre to posttest	estimate = 0.26, *SE* = 0.36, z = 0.72, *p* = 0.75, *d* = 0.05	estimate = -1.65, *SE* = 0.57, *z* = -2.85, *p* = 0.01, *d* = 0.43
Post to delayed posttest	estimate = -0.04, *SE* = 0.35, *z* = -0.11, *p* = 0.99, *d* = 0.10	estimate = -1.52, *SE* = 0.59, *z* = -2.55, *p* = 0.02, *d* = 0.40
Pre to delayed posttest	estimate = 0.30, *SE* = 0.35, *z* = 0.85, *p* = 0.67, *d* = 0.05	estimate = -0.12, *SE* = 0.40, *z* = -0.31, *p* = 0.94, *d* = 0.03

Appendix D

Statistical outputs from Heritage learners' analyses

Output from mixed effects linear regression model with HLs' AJT ratings for grammatical and ungrammatical items

	Estimate	*SE*	*t*	*p*
Grammatical items				
Intercept	2.77	0.51	5.37	0.00
Instruction	0.11	0.60	0.18	0.85
Time (delayed posttest)	1.18	0.53	2.22	0.02
Time (posttest)	0.67	0.56	1.18	0.24
Instruction x Time (delayed posttest)	-0.29	0.67	-0.43	0.66
Instruction x Time (posttest)	-0.10	0.83	-0.12	0.90
Ungrammatical items				
Intercept	4.53	0.21	21.50	0.00
Instruction	-0.19	0.24	-0.78	0.43
Time (delayed posttest)	-0.07	0.27	-0.28	0.77
Time (post)	-0.55	0.27	-2.03	0.04
Instruction x Time (delayed posttest)	-0.59	0.35	-1.65	0.10
Instruction x Time (posttest)	-0.01	0.37	-0.05	0.95

Pairwise comparisons:

Ungrammatical items		
	Control group	Instructed group
Pre to posttest	estimate = -0.44, *SE* = 0.24, *t* = -1.80, *p* = 0.17, *d* = 0.22	estimate = -1.09, *SE* = 0.21, *t* = -4.99, *p* = 0.00, *d* = 0.37
Post to delayed posttest	estimate = 0.17, *SE* = 0.25, *t* = 0.67, *p* = 0.78, *d* = 0.32	estimate = 0.00, *SE* = 0.22, *t* = 0.08, *p* = 0.99, *d* = 0.02
Pre to delayed posttest	estimate = -0.27, *SE* = 0.26, *t* = -1.06, *p* = 0.53, *d* = 0.08	estimate = -1.07, *SE* = 0.21, *t* = -4.98, *p* = 0.00, *d* = 0.33

Output from mixed effects binomial logistic regression model with HLs' EIT accuracy responses for grammatical and ungrammatical items

	Estimate	*SE*	*t*	*p*
Grammatical items				
Intercept	0.29	0.27	1.08	0.27
Instruction	-0.44	0.35	-1.26	0.20
Time (delayed posttest)	-1.01	0.36	-2.79	0.00
Time (posttest)	-0.68	0.35	-1.90	0.05
Instruction x Time (delayed posttest)	1.29	0.47	2.72	0.00
Instruction x Time (posttest)	1.21	0.48	2.48	0.01
Ungrammatical items				
Intercept	-3.32	0.58	-5.65	0.00
Instruction	0.48	0.78	0.62	0.53
Time (delayed posttest)	-0.50	1.16	-0.43	0.66
Time (post)	0.37	0.83	0.45	0.65
Instruction x Time (delayed posttest)	2.69	1.30	2.06	0.03
Instruction x Time (posttest)	1.37	1.01	1.35	0.17

Pairwise comparisons:

Grammatical items		
	Control group	Instructed group
Pre to posttest	estimate = 1.01, *SE* = 1.01, *z* = 2.79, *p* = 0.01, *d* = 0.35	estimate = -0.53, *SE* = 0.33, *z* = -1.60, *p* = 0.24, *d* = 0.42
Post to delayed posttest	estimate = -0.32, *SE* = 0.32, *z* = -0.99, *p* = 0.57, *d* = 0.13	, estimate = -0.24, *SE* = 0.32, *z* = -0.75, *p* = 0.73, *d* = 0.09
Pre to delayed posttest	estimate = 0.68, *SE* = 1.90, *z* = 0.13, *p* = 0.13, *d* = 0.39	estimate = -0.28, *SE* = 0.31, *z* = -0.92, *p* = 0.62, *d* = 0.22
Ungrammatical items		
	Control group	Instructed group
Pre to posttest	estimate = -0.37, *SE* = 0.83, z = -0.45, *p* = 0.89, *d* = 0.08	estimate = -1.75, *SE* = 0.57, *z* = -3.07, *p* = 0.00, *d* = 0.45
Post to delayed posttest	estimate = -0.88, *SE* = 1.17, *z* = -0.75, *p* = 0.73, *d* = 0.12	estimate = 0.44, *SE* = 0.37, *z* = 1.19, *p* = 0.45, *d* = 0.15
Pre to delayed posttest	estimate = 0.50, *SE* = 1.16, *z* = 0.43, *p* = 0.90, *d* = 0.04	estimate = -2.19, *SE* = 0.58, *z* = -3.75, *p* = 0.62, *d* = 0.63