ORIGINAL ARTICLE

Resilience and vulnerability of discourse-conditioned word order in heritage Spanish

Bradley Hoot[1],* and Tania Leal[2],*

[1]Department of Modern Languages, DePaul University, Chicago, IL 60614, USA and [2]Department of Spanish & Portuguese, University of Arizona, Tucson, AZ 85721, USA
*Corresponding authors. Email: tanialeal@arizona.edu; bhoot@depaul.edu

Abstract

Heritage speakers—bilinguals who acquire minority languages naturalistically in infancy but are typically majority-language-dominant in adulthood—generally acquire grammars that differ systematically from the *baseline* input received in childhood. Yet not all areas diverge equally; understanding what characterizes divergence or resilience of a given feature is crucial to understanding heritage language acquisition. In this realm, we investigate the discourse-conditioned non-canonical word orders that mark *information focus* in Spanish. Focus bears the hallmarks of structures that diverge from the baseline, yet the evidence is mixed. We use an offline forced-choice task and an online self-paced reading task to compare heritage speakers' judgments and processing to the baseline's, and we find, echoing recent work, that the heritage speakers largely resemble baseline speakers. We interpret this convergence with reference to seven factors potentially affecting heritage language acquisition and identify one hypothesis—that focus facilitates processing due to its structural and pragmatic salience—as a promising explanation.

Keywords: heritage speakers; heritage language acquisition; focus; self-paced reading; information structure; Spanish

1. Background

1.1. Divergence and resilience in heritage language acquisition

Although heritage speakers are native speakers of their heritage languages, their grammars can vary substantially from those of other native speakers. To understand the source of these differences, we must distinguish divergence that arises during acquisition from changes already present in the baseline (Polinsky & Scontras, 2020, p. 4). Because baseline speakers are often first-generation immigrants with varying proficiency in the majority language, their L1 can undergo attrition or influence from the majority language. Hence, apparent divergences in heritage grammars may represent, instead, *successful* acquisition of the input to which they were

exposed, which may diverge from homeland varieties—a situation termed "intergenerational attrition."

When a given linguistic feature is present in the input provided by baseline speakers, its absence in adult heritage grammars can derive from either *attrition*—acquiring and then losing a grammatical feature—or *divergent* (sometimes called "incomplete") *acquisition*—never fully acquiring a given feature or acquiring a different representation than the baseline's. Although distinguishing between these explanations requires longitudinal research, a methodology we do not employ here, both attrition and incomplete acquisition share a common root cause: reduced input.[1] Typically, the input directed toward child heritage speakers is reduced, limited, or interrupted—sometimes dramatically—especially after starting school.

Crucially, not all areas of the grammar are affected equally by input interruptions: some features (e.g., tense) are more resilient, while some (e.g., case morphology) are more prone to divergence (Polinsky & Scontras, 2020). The key question, then, becomes: what predicts whether a given area of the heritage grammar is conserved or not?

1.2. Word order in heritage languages

Word order is useful for evaluating factors affecting resilience because evidence suggests both resilience and vulnerability. Benmamoun et al. (2013) point out that phrase structure and basic word order are resilient, as shown by successful V2 placement in heritage Swedish (Håkansson, 1995; Larsson et al., 2015), Norwegian (Eide & Hjelde, 2015), and German (Hopp & Putnam, 2015). Moreover, English heritage speakers in contact with pro-drop languages (e.g., Japanese) do not delete subjects in English (Polinsky, 2018), Russian heritage speakers perform well with VS in existential constructions (Polinsky, 2006), and Spanish heritage speakers also display successful acquisition of the syntax of VS orders with intransitive verbs (Gellon, 2015; Montrul, 2005) and in embedded questions (Hoot & Ebert, 2021).

Yet word order is not always unproblematic. Polinsky notes that, in comparison to syntactic knowledge (e.g., A- and A' dependencies), "word order appears to be a more vulnerable domain, subject to general change and sometimes to transfer" (2018, p. 273). This is true with "basic" word orders—for example, not all cases of V2 report success (Norwegian: Johannessen, 2015; Yiddish: Kahan Newman, 2015)—and for word orders conditioned by lexical semantics—for example, unaccusative/unergative alternations in Spanish (de Prada Pérez & Pascual y Cabo, 2012)—but especially for word orders interacting with discourse context.

Key to the present study is evidence for a "reduction of word order flexibility" when expressing information structure in heritage languages (Laleko, 2021, p. 696). That is, heritage speakers generally display a reduced ability to use word order alterations to fit sentences felicitously into the discourse context. In subject focus contexts in which monolingual Russian speakers prefer subject-final word orders, Laleko (2022) found that heritage speakers prefer canonical SV or SVO orders. Comprehension is also affected. Montrul (2010) found heritage Spanish speakers accurately interpreted SVO sentences but not left-dislocated (topical) objects. More broadly, increased use of canonical orders has been found

for Russian (Isurin & Ivanova-Sullivan, 2008; Montrul, 2016, p. 71; Polinsky, 2006), Spanish (Montrul, 2016, p. 71), Korean (Song et al., 1997), and Romanian (Montrul et al., 2015). Egyptian and Palestinian Arabic heritage speakers similarly tend to overuse their variety's canonical order (SVO and VSO, respectively; Albirini et al., 2011). These results align with research showing that L1 and L2 speakers are initially drawn to canonical word orders (O'Grady et al., 2005).

Other research shows more encouraging results for non-canonical orders indexing pragmatic information. Although Montrul (2010) found difficulty in comprehension of left dislocations, two investigations found that heritage speakers did not differ from the baseline in judging their contextual felicity (Leal Méndez et al., 2015; Leal et al., 2014). Sequeros-Valle et al. (2020) found that heritage speakers could distinguish (in)felicitous contexts in a speeded production task, although they produced more infelicitous dislocations than baseline speakers. Similarly, Laleko and Dubinina (2018) documented felicitous use of non-canonical orders by Russian heritage speakers, albeit accompanied by some infelicitous uses differing from the baseline. Heritage speakers of Hungarian judged the syntactic realization and interpretation of non-canonical word orders for focus realization like baseline speakers (Hoot, 2019), and heritage Spanish speakers have also shown sensitivity to discursive restrictions on the subjects of psych verbs (Gómez Soler & Pascual y Cabo, 2016), acceptability judgments of focus realization (Gellon, 2015; Gómez Soler & Pascual y Cabo, 2018; Hoot, 2017), and focus production (Leal et al., 2018).

In summary, evidence of the acquisition of word order in heritage languages is mixed: the syntax of word order variation is largely (but not always) resilient, while word orders indexing pragmatics are sometimes found to be vulnerable and sometimes resilient. Several studies have found word orders marking focus—that is, making new or non-presupposed information prominent—to be examples of resilience. We extend this research here, investigating online processing and offline judgments of non-canonical word orders marking focus in heritage Spanish.

Considering the existing evidence, the question regarding what factors play a role in resilient acquisition remains. Scholars have proposed language-internal and language-external variables affecting acquisition (see Benmamoun et al., 2013; Polinsky, 2018; Polinsky & Scontras, 2020 for useful overviews). In what follows, we identify potential explanations for divergence and a potential explanation for why focus appears to buck the trend. We then consider possible methodological effects.

1.3. Why are some word orders harder than others?
Here we consider four potential reasons for divergence, along with a proposal for why discourse-related word order alterations could be maintained.

1.3.1. Transfer
One possible explanation for divergence is transfer/cross-linguistic influence. Many findings reported in section 1.2 involve languages with relatively flexible word order becoming more rigid when in contact with syntactically rigid English. Evidence pointing to transfer comes from cases in which word order flexibility is *maintained*

when the dominant language is also flexible, such as heritage Russian in contact with German (Brehmer & Usanova, 2015) and heritage Spanish in contact with Dutch (van Osch & Sleeman, 2018). Even basic word order can be affected by transfer, as shown by V2 effects appearing in heritage English in contact with Dutch (Bosch & Unsworth, 2021). Conversely, many take Albirini et al.'s (2011) findings—increased word order rigidity in heritage Arabic—as evidence against transfer effects because heritage Palestinian Arabic speakers overused VSO, despite having English as a majority language.

1.3.2. Frequency

Frequency could explain why common, often canonical, orders are relatively resilient, whereas less common, often non-canonical, orders are more vulnerable. In the usage-based literature, (morpho)syntactic productivity—defined as "the ability of a pattern to apply to novel items" (Poplack, 2001, p. 408)—has been proposed to relate to type frequency (Bybee & Thompson, 1997).[2] For example, in Russian, some of the six possible permutations of the base word order are quite rare (Miller & Weinert, 2009). Their low type frequency could explain why heritage speakers avoid them in favor of more frequent canonical orders.

In our case, this is relevant because in 20 hours of informal conversation in Spanish, Ocampo (2009) documented only two cases of VSO order and none of VOS (both non-canonical), while finding ample use of SVO (canonical). More recently, Davidson (2016), using the Corpus del Español (Davies, 2002), found only 12 exemplars of VSO in over 70,000 tokens. Although we do not have comparable corpus data on VOS, these two studies together suggest the type frequency of V-initial structures is very low, which could complicate their acquisition.

1.3.3. Optional movements and variability

Polinsky (2018) speculates that one cause of divergence could be the "optional" nature of certain movements, such as focus in Hungarian. This idea fits well for information structure because there are often multiple ways to encode a given discourse relationship. For instance, a topical/given constituent could be realized as a hanging topic, dislocated, deaccented in situ, scrambled, replaced with a pronoun, or simply deleted. Focused constituents likewise can be stressed in situ, moved to particular positions, presented in a cleft, or uttered alone. Thus, although discourse-related movements are not strictly optional because they can bring about subtle changes in interpretation (Bresnan et al., 2007), they unquestionably exist within a range of possible constructions that speakers can employ. Evidence from monolingual speakers supports the view that focus realizations are variable. In Spanish, monolingual speakers accept and produce multiple word orders to realize focus on subjects and objects (Hoot, 2016; Hoot et al., 2020; Hoot & Leal, 2020; Leal et al., 2018). Given the input's variability, heritage speakers may acquire constructions that meet their communicative needs without acquiring the full range of possibilities.

1.3.4. Processing complexity

Finally, it has been proposed that certain word orders tax the processing resources of bilinguals in ways that canonical, unmarked orders do not. Sorace and colleagues' Interface Hypothesis contends that aligning sentences to contextual or discourse information incurs a high processing cost (Sorace, 2011). Because bilinguals are hypothesized to operate under limited memory and cognitive resources (Dekydtspotter & Renaud, 2014), such high costs may overwhelm them, leading to divergent performance. When extended to heritage speakers (see Montrul & Polinsky, 2011), this view could explain the apparent difficulty observed with some discourse-conditioned word orders.

1.3.5. Salience

Although infrequent, focus is salient in terms of its structural prominence, its informational value to the interpretation of the utterance in context, and its role in real-time processing. Studies of sentence processing find that focus is processed early and incrementally in L1 comprehension, such that the discourse context modulates sentence processing, facilitating processing of marked word orders (Kaiser, 2016). Laleko (Laleko, 2021, p. 721) thus concluded that focus may be resilient because "focusing structures increase the availability of the representation in short-term memory and have facilitative effects on processing." This proposal is partially at odds with the Interface Hypothesis (Sorace, 2011), which suggests that any structure at the syntax-discourse interface should be problematic. Laleko instead suggests that focus *facilitates* processing, which means that focus stands apart from other syntax-discourse structures, such as (but not limited to) topicalizations.[3]

1.4. Methodological factors

Inconsistent results for focus may also be related to methodology.

1.4.1. Comparison to the baseline

Polinsky (2018) argues that the appropriate comparison group for heritage speakers is *baseline* speakers (i.e., first-generation immigrants), not monolinguals (who Polinsky calls *homeland* speakers).[4] Comparing heritage speakers to monolinguals risks incorrectly ascribing a difference to divergences in acquisition when said feature was already attrited in the baseline population (as noted in section 1.1). Although the existence of "intergenerational attrition" is widely recognized, it is not always carefully controlled for. Here, although we compare heritage speakers to a baseline group, we addressed the possibility that the baseline group may have undergone attrition or contact-induced language change by comparing them to a monolingual group reported in a previous study, finding no differences between the baseline and monolingual groups.

1.4.2. Variation in proficiency across heritage speaker populations

Heritage speaker populations may also differ in proficiency, relative language dominance, and literacy. Laleko (2021, p. 716) highlights the role of literacy and community factors, noting that studies reporting successful acquisition of information structure often involve communities "characterized by high degrees of minority language maintenance" such as Spanish in the U.S. Thus, some of the differences observed in research on information structure (e.g., between Russian and Spanish) could relate to differences in proficiency from broader societal factors (i.e., public presence of Spanish, similar orthography, etc.).

Additionally, within a given population, heritage speakers are heterogeneous in terms of proficiency. Variations in proficiency could affect previously reported results, with groups of different proficiency levels producing different outcomes. We therefore measure proficiency and relative language dominance to quantify these variables.

1.4.3. Task type

Task type can have contradictory effects: heritage speakers tend to better at tasks measuring interpretation or comprehension than those requiring production (Polinsky, 2018), yet they also perform better on naturalistic production than on metalinguistic tasks (Montrul et al., 2008). It is important to note that heritage speakers tend to display a "*yes*-bias," being less willing to reject ill-formed linguistic structures due to linguistic insecurity (Polinsky, 2018). In section 6.3, we interpret how task type affects our results.

2. Linguistic phenomenon

We investigate Spanish word order alterations associated with *focus*—a notion that indicates the presence of relevant alternatives for interpreting linguistic expressions (Krifka, 2008, p. 247).

2.1. Focus marking in Spanish

Speakers package information in linguistic propositions taking into consideration whether information is known to the interlocutor or not—namely, whether information is given or new. When a question such as *What did you eat?* arises, language users expect that a DP such as *an apple* could follow. Language users also expect that this element, which closes the variable opened by the *wh*-word *what*, will receive some sort of linguistic **prominence**, expressed via prosodic, syntactic, and/or morphological means.

Spanish can encode prominence syntactically, as shown in (1), where the subject DP *Kaori Sakamoto* (1a) closes the variable opened by the *wh*-operator *quién* "who," displaying non-canonical VOS word order. It has been claimed that Spanish *primarily* marks information focus by placing focal material in rightmost position via syntactic movement, since this is the position where it would receive main sentence stress (Büring & Gutiérrez-Bravo, 2001; Casielles-Suárez, 2004; Domínguez, 2004; Olarrea, 2012; Zubizarreta, 1998).

(1) Subject focus context: ¿Quién ganó el bronce en patinaje femenino en Beijing?
     "Who won the bronze (medal) in female figure skating in Beijing?"

     a. Por suerte ganó el   bronce [Kaori Sakamoto]$_F$. [VOS]
         for luck   won   the bronze KS

     b. #Por suerte ganó [Kaori Sakamoto]$_F$ el   bronce. [#VSO]
         for   luck   won   KS   the bronze

     c. Por suerte [Kaori Sakamoto]$_F$ ganó el   bronce. [SVO]
         for luck   KS   won   the bronze
     "Fortunately, Kaori Sakamoto won the bronze (medal)."

Although (1a) is a suitable answer, it is merely one way to encode this proposition. Notably, potential alternatives are constrained such that not every equivalent proposition will felicitously fit in the context. Other non-canonical orders such as VSO (1b) are not felicitous. Nevertheless, VOS is not the only felicitous answer: canonical SVO (1c) could also be used (although this point is not undisputed; see Hoot & Leal, 2020).

Thus, non-canonical word orders are restricted to discourse contexts into which they fit felicitously. While VOS can be used to answer subject focus questions, as in (1), VSO is appropriate for questions focusing on the object (e.g., *What medal did Sakamoto win?*), while VOS is not. In this investigation, we capitalize on this distinction between these two infrequent, non-canonical V-initial word orders and test whether Spanish speakers with different linguistic backgrounds can distinguish the contexts in which each order is felicitous.

Word order alterations to mark focus affect other constructions. When a question targets the object (2), if the answer contains a final prepositional phrase, one possible answer is (2a), with non-canonical VPPO order. Like the non-canonical orders in (1), this word order alteration is only available under the appropriate context, but it is not the only possible answer to (2); canonical VOPP order (2b) is also available (as is elision of post-focal material).

(2) Object focus context: ¿Fuiste al mercado ya? ¿Qué compraste?
     "You went to the market already? What did you buy?"

     a. Compré en   el   mercado [unos postres exquisitos]$_F$. [VPPO]
         I-bought in   the market   some   desserts exquisite

     b. Compré [unos postres exquisitos]$_F$ en   el   mercado. [VOPP]
         I-bought some   desserts exquisite   in   the market
     "I bought some delicious desserts at the market."

This example shows that non-canonical word orders in Spanish are but one option for marking focus as prominent. As with the V-initial subject focus orders, we examine whether heritage speakers of Spanish are sensitive to these discursive restrictions.

Because our experimental group is bilingual and dominant in English, and because transfer from the dominant language may play a role in heritage language divergence, we will briefly review focus in English.

2.2. Focus marking in English

Although Spanish can use sentence-level scrambling operations so that focused elements end up at the rightmost edge of the phrase, English has been argued to use phonological de-stressing and in situ stress. Out-of-the-blue sentences have nuclear stress, with the most prominent pitch accent on the rightmost constituent (3). Narrow focus entails shifting the stress to the focal constituent in situ (and de-stressing post-focal material), as in (4) (Selkirk, 1995).

(3) Out-of-the-blue context: What happened?
     [John ate the **pie**]_F.

(4) Subject focus context: Who ate the pie?
     [**John**]_F ate the pie.

Although pitch accent plays a role in making Spanish focus prominent (Feldhausen & del Vanrell, 2014; Zubizarreta, 1998), the word orders available to mark focus in Spanish—VOS, VSO, VPPO—are ungrammatical in English. While English shares its canonical word order with Spanish—SVO or SVOPP—the non-canonical word orders that realize Spanish focus in the cases we consider have no direct correlate in English.[5]

3. Research questions

We examine the factors affecting resilience in heritage languages by investigating whether heritage Spanish speakers are sensitive to the discourse restrictions on non-canonical word orders used to mark focus in offline judgments and online processing. Our research questions are in (5).

(5) a. Do heritage Spanish speakers pattern with baseline speakers in their
         judgments of non-canonical word orders in focus contexts?
     b. Do heritage Spanish speakers pattern with baseline speakers in their
         processing of non-canonical word orders in focus contexts?

4. Methods

Our experiment included two tasks: an offline forced-choice judgment task (FCT) and an online self-paced reading task (SPR). Participants also completed a proficiency test and background questionnaire.

4.1. Participants

We used the Bilingual Language Profile (BLP; Birdsong et al., 2012) to collect information about participants' demographics, language acquisition history, language use, and attitudes. The BLP produces a *dominance* score (−218 to 218), indicating greater or lesser relative dominance in each language, so we analyze language dominance as a continuous independent variable.

Table 1. Participant characteristics[6]

	Group means (range)	
	Heritage speakers ($n = 46$)	Baseline ($n = 31$)
Gender	36 F/10 M	16 F/15 M
Age	21.7 (18 to 45)	39.5 (18 to 80)
LexTALE_Esp Score	21.1 (−1 to 52)	43.1 (27 to 56)
BLP Dominance Score (lower = Spanish-dominant)	24.8 (−92 to 100)	−84 (−161 to −6)
Years of education in Spanish	5.2 (0 to 18)	11.4 (2 to 20+)
Years of education in English	13.8 (2 to 18)	4.5 (0 to 14)
% weekly Spanish use: family	70% (10 to 100%)	85% (0 to 100%)
% weekly English use: family	36% (0 to 100%)	17% (0 to 100%)
% weekly Spanish use: friends, self	28% (0 to 100%)	70% (10 to 100%)
% weekly English use: friends, self	74% (10 to 100%)	33% (0 to 90%)
% weekly Spanish use: work/school	27% (0 to 100%)	43% (0 to 100%)
% weekly English use: work/school	80% (30 to 100%)	57% (0 to 100%)
Self-rating, Spanish, speak & understand	5.3/6 (2 to 6)	5.9/6 (4 to 6)
Self-rating, English, speak & understand	5.7/6 (3 to 6)	4.0/6 (1 to 6)
Self-rating, Spanish, read & write	4.7/6 (1 to 6)	5.8/6 (4 to 6)
Self-rating, English, read & write	5.6/6 (3 to 6)	4.1/6 (1 to 6)

Participants belonged to two groups according to language experience. The *heritage speaker group* was composed of adults who were born in the U.S. or arrived before age 8, were raised in homes where Spanish was spoken, and began acquiring both Spanish and English at or before age 8. Their mean self-reported age of first exposure to English was 3.3 years (range 0-8; 13 people reported exposure "since birth"). They lived in or near Chicago, Illinois, at the time of testing. The *baseline group* were adults who were born in a Spanish-speaking country[7] and moved to the U.S. (where they now live), at or after age 12, were raised in homes where only Spanish was spoken, and were classified as Spanish-dominant on the BLP. They resided mostly in Hattiesburg, Mississippi; Reno, Nevada; and Chicago, Illinois. Participants who had significant contact with other languages before age 12 were excluded. Additionally, one participant's data were excluded from the SPR task results for having overall reading times more than 2.5 SDs above the group mean RT, and two participants' SPR data were excluded for answering more than 20% of the comprehension questions incorrectly.

After exclusions, 46 people in the heritage speaker group completed the FCT and other materials. Half of them ($n = 23$) completed the SPR task. Thirty-one baseline speakers completed the FCT, of whom 22 also completed the SPR task. Relevant group characteristics are provided in Table 1.

To measure proficiency, we used the LexTALE_Esp (Izura et al., 2014), a lexical decision task requiring participants to determine whether a letter string constitutes a real Spanish word. Sixty real words were presented alongside 30 plausible non-words; participants were awarded one point for each real word and penalized two points for each incorrectly accepted non-word (to adjust for guessing), producing possible scores ranging from -60 to 60. While vocabulary size does not directly index the more complex notion of language proficiency, it can serve as a rough estimate of overall language abilities because it correlates with several other measures of general proficiency (Lemhöfer & Broersma, 2012). The LexTALE_Esp correlates well with proficiency for L2 learners (Izura et al., 2014) and can distinguish between bilingual groups at high proficiency levels (Ferré & Brysbaert, 2017). Mean LexTALE_Esp scores are reported in Table 1.

4.2. Forced-choice task

We used a contextualized forced-choice task (FCT) because of the advantages such tasks provide. As Stadthagen-González et al. (2018) note, comparative judgments are easier and more reliable than stand-alone ratings, making these tasks less taxing on memory resources. FCTs can also capture relatively small differences in terms of acceptability, especially for modest or small effect sizes (Schütze & Sprouse, 2013). We expected such effects because our sentences were all grammatical and only differed in terms of their contextual felicity. Finally, FCTs can effectively address the "*yes*-bias" displayed by many heritage speakers due to linguistic insecurity (Polinsky, 2018) by shifting the focus away from rejecting or accepting sentences to expressing preferences.

4.2.1. Procedure

Participants were presented with a picture setting the scene, followed by a *wh*-question focusing on the subject (e.g., *¿Quién bebió la leche?* "Who drank the milk?"), the object (e.g., *¿Qué plantó en el jardín?* "What did (s)he plant in the garden?"), or the (adjunct) prepositional phrase (e.g., *¿Dónde perdió el zapato?* "Where did (s)he lose the shoe?"). Participants were instructed to choose the most acceptable sentence in the context of the preceding question and saw either two (subject focus) or three (object focus, PP focus) sentence options. Each condition included 16 lexicalizations, for a total of 48 test sentence pairs/triplets. These lexicalizations were distributed into two lists, such that each participant judged 24 test sentence sets (8 per condition), along with 24 fillers. Fillers always had two choices, so the only items with three choices were the 16 object/PP focus trials; that these items were different is a limitation of this design. More generally, we recognize the limitation that forced-choice items varying in word order makes word order relatively salient to participants, which is a trade-off for the increased power to detect differences that this method can offer.

Test sentences featured words from the 5000 most common Spanish words (Davies, 2006). Sentences were randomized per trial, and trials were randomized per participant. The entire FCT, including instructions and two practice items,

Sample FCT trial, subject focus

Translation

Who turned the lamp on?

It looks like turned on the singer the lamp. (VSO)

It looks like turned on the lamp the singer. (VOS)

"It looks like the singer turned on the lamp."

¿Quién prendió la lámpara?

Se ve que prendió la cantante la lámpara.

Se ve que prendió la lámpara la cantante.

Figure 1. Sample FCT Trial, Subject Focus.

was presented in Spanish, via Qualtrics but under researcher supervision, and usually lasted around 15 minutes.

4.2.2. Materials: Subject focus condition

For the subject focus condition, the relevant factor was word order. The options were two: either focus-final (VOS) or non-final focus (VSO). We intentionally avoided SVO because previous research has showed that SVO, as canonical, non-marked, and the most frequent order in Spanish, can be used under almost any information-structural context. Crucially for our purposes, these two V-initial orders differ in that VOS can be used for subject focus marking, while VSO can only mark either broad focus or narrow focus on the object (Domínguez, 2004, p. 74; Zubizarreta, 1998, p. 125). Test sentences included material preceding the verb because Gutiérrez-Bravo (2020) has noted that in Mexican Spanish, VSO/VOS can only be grammatical when following another constituent. As shown in Figure 1, which depicts a subject focus trial, our items were embedded as subordinate clauses.

Other controls included matching the number of syllables of subjects and objects because phonological weight can affect the order (Heidinger, 2015). Subjects were all [+human] DPs, while objects were [−animate]. All DPs were specific and definite. We did not control for the gender of the DPs because there was no reason to believe gender influences information structure.

4.2.3. Materials: Object/PP focus condition

We tested focus type (object focus vs. PP focus) and word order. The word order options were (canonical) VOPP, VPPO, and focus fronting (OVPP or PPVO, according to the context). To avoid undesired phonological weight effects, we controlled for number of syllables. Figure 2 shows a PP focus trial.

Sample FCT trial, PP focus

Translation
Where did he lose the shoe?
I suppose that he lost the shoe at the park. (VOPP)
I suppose that he lost at the park the shoe. (Fronting)
I suppose that he lost at the park the shoe. (VPPO)
"I suppose he lost the shoe at the park."

¿Qué perdió en el parque?

Supongo que el zapato perdió en el parque.

Supongo que perdió en el parque el zapato.

Supongo que perdió el zapato en el parque.

Figure 2. Sample FCT Trial, PP Focus.

In an object/PP focus trial, Fronting should be infelicitous since focus fronting occurs in contexts of emphasis or correction, not as the answer to a *wh*-operator.

4.3. Self-paced reading task

Self-paced reading tasks can index increased processing difficulty, as measured by higher reading times (RTs) per segment, when compared to another segment in a control condition.

4.3.1. Procedure

Participants read test sentences on a computer screen in a non-cumulative (segment-by-segment) fashion, from left to right, by pressing the space bar. At the outset, participants read instructions which included four practice items followed by yes/no comprehension questions. Participants received feedback on comprehension questions.

After practice items, participants were presented with a non-moving discourse context that spanned 2-3 lines and ended with a question that focused on the subject (e.g., *¿Quién lo distrajo?* "Who distracted him?"), the object (e.g., *¿A quién distrajo?* "Whom did he distract?" or *¿Qué compró en el mercado?* "What did she buy in the market?"), or a prepositional phrase (e.g., *¿Dónde compró los caramelos?* "Where did she buy the candy?"). Non-moving contexts were followed by a sentence with non-space characters masked by dashes (-). Test sentences had seven regions, three of which constituted the critical region (regions 3–5). All sentences fit on one line.

| Context | Me dijeron que alguien importante había sorprendido al locutor de radio. ¿Sabes quién lo sorprendió? |
| | "I was told that somebody important had surprised the radio announcer. Do you know who surprised him?" |

	Region 1	Region 2	Regions 3-5 (critical)	Region 6	Region 7
VOS felicitous	Mira yo	supe que	sorprendió / al locutor / el obispo,	aunque	me sorprende.
	Look I	*knew that*	*he-surprised to-the announcer the bishop*	*although*	*me surprises*
	"Look, I found out that the bishop surprised the announcer, although it surprises me."

Figure 3. Sample Token, Subject Focus Condition (Felicitous VOS Order).

When participants finished reading the sentence, a new screen appeared presenting a centered yes/no comprehension question. Comprehension questions were counterbalanced: half false, half true. Additionally, half the questions focused on the context and half on the test sentence, although none involved the critical region. The software Linger (Rohde, 2003) recorded participants' reading times (per segment) as well as their accuracy on the comprehension question. Participants read 96 sentences, randomized by participant: 32 in the subject/object focus condition, 32 in the object/PP focus condition, and 32 fillers. The experiment lasted around 45 minutes.

4.3.2. Materials: Subject/Object focus condition
The 2 × 2 factorial design crossed word order (VSO/VOS) and focus type (subject/object). We employed a series of controls to avoid uninformative word order effects. Subjects and objects had the same number of syllables, since heavier arguments tend to appear sentence-finally (Heidinger, 2015). Following Gutiérrez-Bravo (2020), we avoided V-initial sentences by embedding them inside a carrier phrase. To ensure that the theta roles were not predictable, subjects and objects were both [+human] DPs, and we first designed a norming task that tested plausibility and reversibility of 54 lexicalizations (previously reported in Hoot & Leal, 2020, 2022; Leal & Hoot, 2022). Of these, 32 were retained (8 per cell) and then distributed across four lists. As with the FCT, we intentionally avoided canonical SVO orders because there are biases toward default/canonical forms, and SVO appears to fit multiple information-structural configurations. Figure 3 shows a sample subject focus trial.

4.3.3. Materials: Object/PP focus condition
This 2 × 2 factorial design also crossed word order (VOPP/VPPO) and focus type (object/PP). Objects and PPs had the same number of syllables and used only definite DPs. In this case, we did include a canonical order (VOPP) because there are no available alternatives in Spanish. We did not control for frequency given the other controls. For this condition, we had 32 lexicalizations, presented across four lists, as noted above. Figure 4 shows a sample trial.

| Context | El detective descubrió una pista cuando pudo localizar algo de la escena del crimen en el jardín. ¿Sabes qué descubrió? |
| | "The detective discovered a clue when we found something from the crime scene in the garden. Do you know what he discovered?" |

	Region 1	Region 2	Regions 3-5 (critical)	Region 6	Region 7
VOPP infelicitous	Mira yo	creo que	descubrió un pañuelo en el jardín,	aunque	falta confirmación.
	Look I	*believe that*	*he-discovered a tissue in the garden*	*although*	*lacks confirmation*
	"Look, I believe he discovered a tissue in the garden, although confirmation is still needed."				

Figure 4. Sample Token, Object Focus Condition (Infelicitous VOPP Order).

5. Results

5.1. Forced-choice task results

5.1.1. Data processing and analysis

The dependent variable was word order. Because the design for each focus type was slightly different, the analysis also differed. The subject focus condition had a binary outcome (VOS/VSO), so we analyzed it using a binomial logistic regression via a generalized linear mixed-effects model (GLMM) using the GENLINMIXED procedure in SPSS, which calculates the likelihood of each of the two options being chosen as an effect of the predictors (fixed factors). In this case, the only fixed factor was group (heritage speaker vs. baseline). To account for repeated measures, we included by-participant and by-item random intercepts. Because the only fixed factor was a between-subjects variable, no by-participant random slopes could be included.

Because the outcome for the object/PP focus conditions had three levels, we used a multinomial logistic regression (same GLMM procedure). Since our aim was to compare whether the choice of outcome varied across the two focus contexts, we analyzed them together. The multinomial logistic regression tells us whether the likelihoods of the three outcomes differ according to the fixed factors. To understand how they differ, we conducted a series of follow-up binomial models, comparing each outcome against the other two.

We concur with Meteyard and Davies (2020), who argue that the results of mixed-effects models should be displayed in tables including both fixed and random effects specifications, along with relevant output, for maximum transparency. However, presenting a table for each statistical test consumes too much space and makes for difficult reading. We therefore present the full output tables in Supplementary File 1 and only key numbers in the text.

5.1.2. Testing attrition in the baseline speakers

To test for "intergenerational attrition," we compared the results of our baseline speakers to those of monolingual Mexican Spanish speakers reported previously (Hoot & Leal, 2020, 2022). We find no evidence suggesting that baseline speakers differed from monolingual Spanish speakers for any of the phenomena we report here. Full results are available in Supplementary File 2.
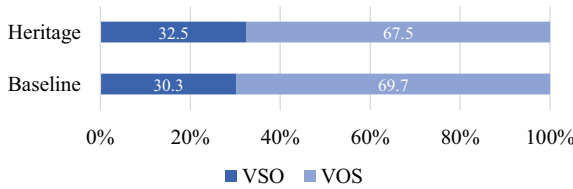
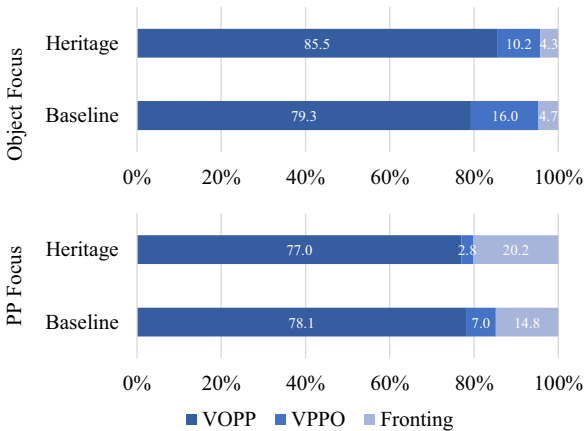**Figure 5.** Word Order Chosen by Group, Subject Focus.

Figure 6. Word Order Chosen by Focus Context and Group, Object/PP Focus.

5.1.3. Subject focus

Figure 5 shows subject focus results, indicating that both groups preferred VOS. The binomial logistic regression found no difference by group ($F = 0.08, p = .777$, odds ratio $= 0.91$; see Supplementary Table S1).

5.1.4. Object/PP Focus

The multinomial logistic regression reveals a significant effect by focus context ($F = 23.38, p < .001$) but no effects by group or its interaction with context: while the distribution of the three answers varies according to the context, the groups do not differ either in their overall distribution or how they respond to contextual differences (see Supplementary Table S2). Figure 6 displays the results.

To further explore these results, we conducted a series of binomial logistic regressions, comparing a single outcome against the other two.

First, we tested whether the likelihood of choosing VOPP, compared to the other two orders (VPPO/fronting), changed across groups or contexts. We found a marginal effect ($F = 3.27, p = .07$, odds ratio $= 1.47$) by focus type, but no evidence of group differences (Supplementary Table S3). Second, when comparing non-canonical VPPO against canonical VOPP and fronting (Supplementary Table S4),

we observed a clear effect by type ($F = 14.99$, $p < .001$, odds ratio $= 3.09$) and a marginal effect by group ($F = 3.67$, $p = .059$, odds ratio $= 0.51$). According to the odds ratios, the odds of choosing object-final VPPO were 3 times higher under object focus than under PP focus. We do not find an interaction between group and type, suggesting that the groups do not differ across contexts. Finally, we examined fronting versus VOPP/VPPO (Supplementary Table S5). We observe an effect by type ($F = 25.23$, $p < .001$, odds ratio $= 0.21$), but no effects by group. The odds ratio suggests that the odds of choosing fronting are nearly five times higher ($1/0.21 = 4.8$) under PP focus for both groups. Because we found a significant difference for fronting, which is not our result of interest, we conducted one final follow-up binomial regression comparing VOPP/VPPO, with Fronting removed from the dataset (Supplementary Table S6). When fronting is removed, we observe a difference by type ($F = 10.50$, $p = .002$, odds ratio $= 2.63$) and a marginal difference by group ($F = 2.99$, $p = .087$, odds ratio $= 0.52$), with no interaction between them, just as we saw for the test of VPPO.

In summary, the results of the object/PP focus AJT show the following: (a) an overwhelming preference for canonical VOPP, irrespective of context, in both groups; (b) increased preference for VPPO under object focus compared to PP focus, with no significant differences by group; and (c) increased preference for fronting (PPVO) under PP focus, with no difference by group.

5.1.5. Proficiency and dominance

Separately for each group, we examined the role of proficiency—measured by the lexical decision task—and language dominance—operationalized as the BLP score. When testing the effect of proficiency on the realization of VOS/VSO under subject focus, we observe no effect for baseline speakers (Supplementary Table S7). For heritage speakers (Supplementary Table S8), we observe only a marginal effect ($F = 3.46$, $p = .070$, odds ratio $= 1.03$). Similarly, we find no evidence suggesting that responses differ by dominance for either group (baseline in Supplementary Table S9, heritage in Supplementary Table S10).

Turning to object focus, we first tested the effect of proficiency on the distribution of the three possible outcomes with a multinomial logistic regression. For baseline speakers (Supplementary Table S11), we find an interaction between proficiency and focus type ($F = 5.96$, $p = .003$), suggesting that the answer distribution changes in different ways per focus type as proficiency changes. To explore this interaction, we plotted the structures chosen by the baseline speakers in each of the two contexts by proficiency (Figure 7). In the object focus context, the trendlines are flat for VOPP and VPPO—the proportion of each answer does not change much as proficiency increases. In the PP focus context, the amount of VOPP increases slightly and the amount of VPPO decreases slightly as proficiency increases (fronting is flat). It appears that for the baseline speakers, VOPP is more likely (and VPPO less likely) as proficiency increases, but only for PP focus. This result is surprising; we did not expect effects by proficiency for the baseline speakers. We discuss this finding in section 6.4. For heritage speakers, we observe no effect by proficiency (Supplementary Table S12).
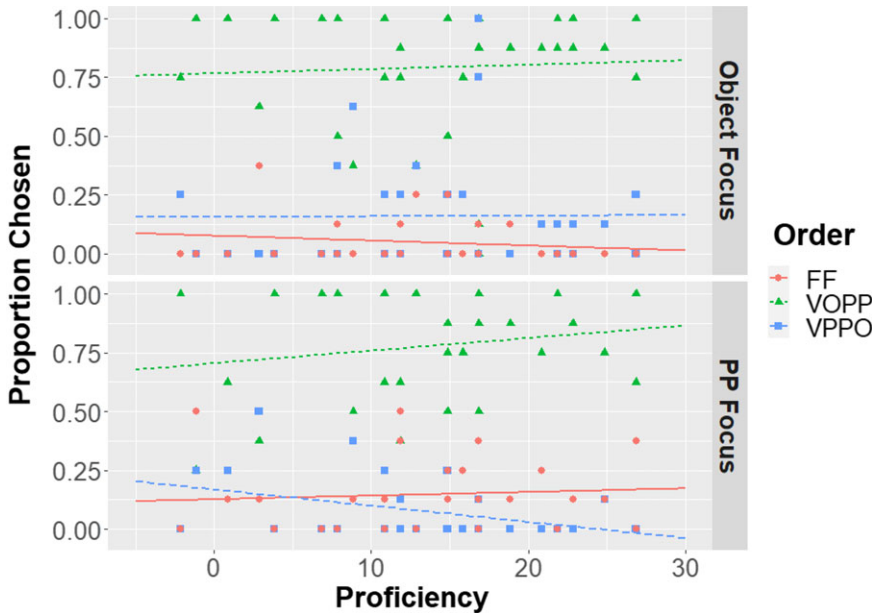
Figure 7. Word Order Chosen by Proficiency and Context, Baseline Speakers.

We also tested the effect of dominance. We do not find evidence to suggest that responses differ by dominance for baseline (Supplementary Table S13) or heritage speakers (Supplementary Table S14).

Finally, at the suggestion of an anonymous reviewer, we tested the effect of age of onset of bilingualism, operationalized as self-reported age of exposure to English, for both types of focus, for the heritage speaker group. We do not find evidence to suggest that responses differ by age of exposure for subject focus (Supplementary Table S15) or object focus (Supplementary Table S16) for the heritage speakers.

5.2. Self-paced reading task results

5.2.1. Data processing and analysis

We trimmed reading times (RTs) at 100 ms and 10,000 ms. Because RTs typically have positive skew, we log-transformed them, producing logRTs. Finally, logRTs were length-adjusted (Fine et al., 2013), which converts them to residuals from a regression of predicted RTs by word length. (Negative RTs are read faster than expected for the word length, positive RTs slower.) Length-adjusted logRTs were analyzed using a linear mixed-effects model (LMM) implemented in SPSS's MIXED command. Each model had three fixed factors: group (heritage/baseline), focus context (subject/object or object/PP), and word order (VOS/VSO or VOPP/VPPO), plus their interactions. To account for repeated measures, each model included a random effects structure (RES) determined top-down, following Barr et al. (2013). We investigated significant effects via post hoc pairwise comparisons with the Bonferroni correction for multiple comparisons.
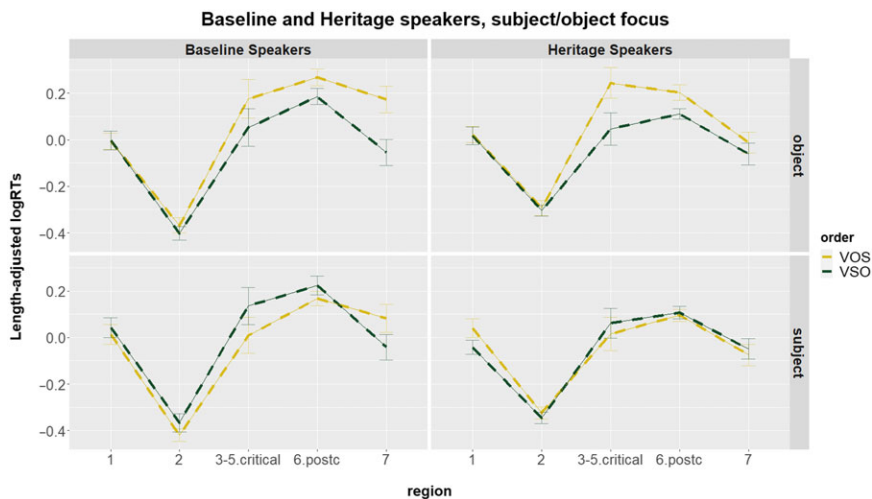
Figure 8. Length-Adjusted logRTs by Group, Order, and Focus for All Regions.

As before, we follow Meteyard and Davies (2020) by reporting the full results of the statistical tests in tables in Supplementary File 1. In the text, we present only essential numerical output. For each of the reported tests, we checked that the model met the assumptions of normality and homoscedasticity of residuals by visually examining histograms, Q-Q plots, P-P plots, and scatterplots for the residuals following procedures outlined by Eddington (2015) and West et al. (2015). We also followed West et al.'s procedure of influence testing to check for outliers and chose not to eliminate any participant or item based on these tests.

5.2.2. Testing attrition in baseline speakers
First, we compared baseline speakers' results to those of monolingual Mexican Spanish speakers previously reported (Hoot & Leal, 2020, 2022). We find no evidence suggesting the baseline speakers differ from monolingual Spanish speakers. Full results are available in Supplementary File 2.

5.2.3. Subject/Object focus
Figure 8 presents the length-adjusted logRTs for all sentence regions for the baseline and heritage speaker groups. It compares object focus (top panels) to subject focus (bottom panels) and compares VOS (yellow line) to VSO (green line) within each panel. We observe that both groups have similar patterns and appear to read VOS more slowly under object focus.

We conducted two LMMs. First, we examined the critical region (regions 3-5), comparing group (baseline/heritage), focus (subject/object), and order (VOS/VSO). We observe a significant focus by order interaction ($F = 5.79$, $p = .027$), suggesting that the RTs for each word order vary across focus contexts (Supplementary Table S17). Post hoc pairwise comparisons indicate a difference between contexts within VOS ($p = .040$): VOS is read faster under subject focus than under object
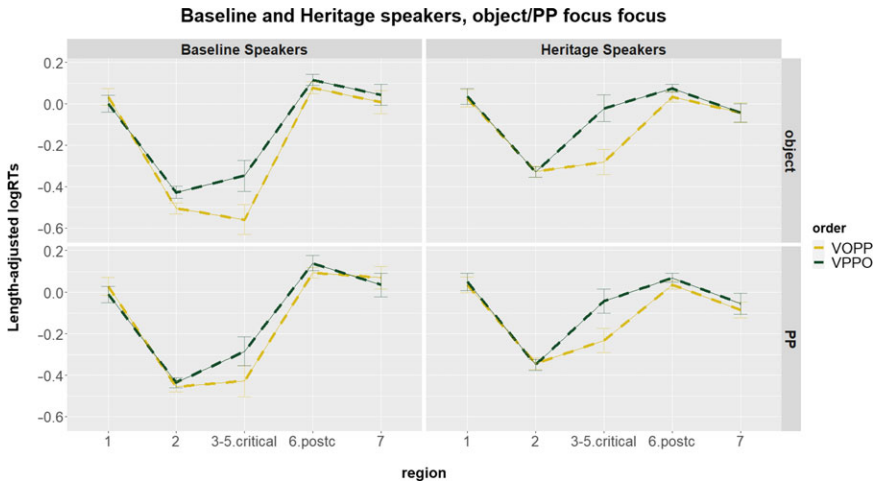
Figure 9. Length-Adjusted logRTs by Group, Order, and Focus for All Regions.

focus. Additionally, we see a difference between word orders within subject focus ($p = .034$): in this context, VOS is read faster than VSO. The other post hoc comparisons show no apparent differences between contexts for VSO nor between orders under object focus. Finally, we observe no effect by group. These results indicate that both groups attend to focus context and its relationship to word order in real-time processing. Specifically, both groups associate VOS with subject focus, processing it faster under subject focus than under object focus, and processing VOS faster than VSO within subject focus.

Because processing difficulty can appear beyond the critical region, we examined the spillover region with the same fixed factors. As in the critical region, we observe an interaction between focus context and word order ($F = 5.67$, $p = .026$), suggesting that RTs for a given order varied by context (Supplementary Table S18). Post hoc pairwise comparisons indicate a difference between contexts within VOS ($p = .005$): VOS is read faster under subject focus than under object focus. Additionally, we found a difference between word orders within object focus ($p = .012$): in this context, VSO is read faster than VOS. We observe no differences between the orders for subject focus, nor between the contexts for VSO.

In terms of group differences, heritage speakers read the post-critical region faster than baseline speakers overall ($F = 4.62$, $p = .037$), but this difference is not the outcome of interest. Crucially, we do not find evidence of interaction between group and other factors, suggesting that the groups do not differ in their reactions to the different orders according to focus context. These results echo the findings for the critical region: The groups do not apparently differ in their processing patterns; rather, both read focus in sentence-final position faster.

5.2.4. Object/PP focus
Figure 9 shows length-adjusted logRTs for all sentence regions. It compares object focus (top panels) to PP focus (bottom panels), and VOPP (yellow line) to VPPO

(green line). Note that both groups appear to read the VPPO order (green) slower in both contexts.

We conducted two LMMs examining the critical region (3-5) and comparing group (heritage/baseline), focus (object /PP), and order (VOPP /VPPO). We observe significant main effects for order ($F = 19.39$, $p < .001$) and group ($F = 5.89$, $p = .020$), indicating that VOPP is read faster overall, irrespective of context, and that baseline speakers read everything faster overall. We find no interactions, suggesting that reading times do not vary by context and groups do not vary in their patterns (Supplementary Table S19.) When we examined the spillover region (Supplementary Table S20), we found no effects.

In summary, we observe a significant processing advantage for canonical VOPP order and an unsurprising faster overall RT for baseline speakers in the object/PP focus experiment, but no differences in processing patterns by group.

5.2.5. Proficiency and dominance

We examined the role of proficiency and dominance, analyzing each group separately. For the sake of space, we only report on the critical region here. We tested the post-critical region as well but found no effects.

We fit an LMM for each group and each experiment. In all cases, the dependent variable was logRT, and the fixed factors were focus (subject/object or object/PP), order (VSO/VOS or VOPP/VPPO), and proficiency (mean-centered LexTALE_Esp score). For the subject/object focus condition (Supplementary Table S21), we observe an effect of proficiency by the baseline speakers: in addition to an overall effect for order, the outcome of interest is its interaction with proficiency ($F = 8.01$, $p = .005$), which indicates that the effect of order changes as proficiency increases for the baseline group.

To visualize this effect, we plotted the relationship between proficiency and RTs for each word order separately (Figure 10). We observe no obvious relationship between proficiency and RT for VSO order, while for VOS order, increased proficiency results in slower processing. As for the isolated result by proficiency for the FCT, this result is surprising; we did not expect effects by proficiency for the baseline speakers. We return to this finding in section 6.4. For heritage speakers, we observe no effects by proficiency (Supplementary Table S22).

Turning to the object/PP focus experiment, we see no effect of proficiency for baseline speakers (Supplementary Table S23), unlike for the subject/object focus experiment. The same is true for the heritage speakers (Supplementary Table S24).

To analyze dominance, we again fit an LMM for each group and experiment. In all cases, the dependent variable was logRT, and the fixed factors were focus (subject/object or object/PP), order (VSO/VOS or VOPP/VPPO), and dominance (BLP score). For the subject/object experiment, neither the baseline (Supplementary Table S25) nor heritage speakers (Supplementary Table S26) vary by dominance. For the object/PP focus experiment, we again observe no apparent effects by dominance for the baseline (Supplementary Table S27) or heritage speakers (Supplementary Table S28).

Finally, at the suggestion of an anonymous reviewer, we tested the effect of age of onset of bilingualism, operationalized as self-reported age of exposure to English,
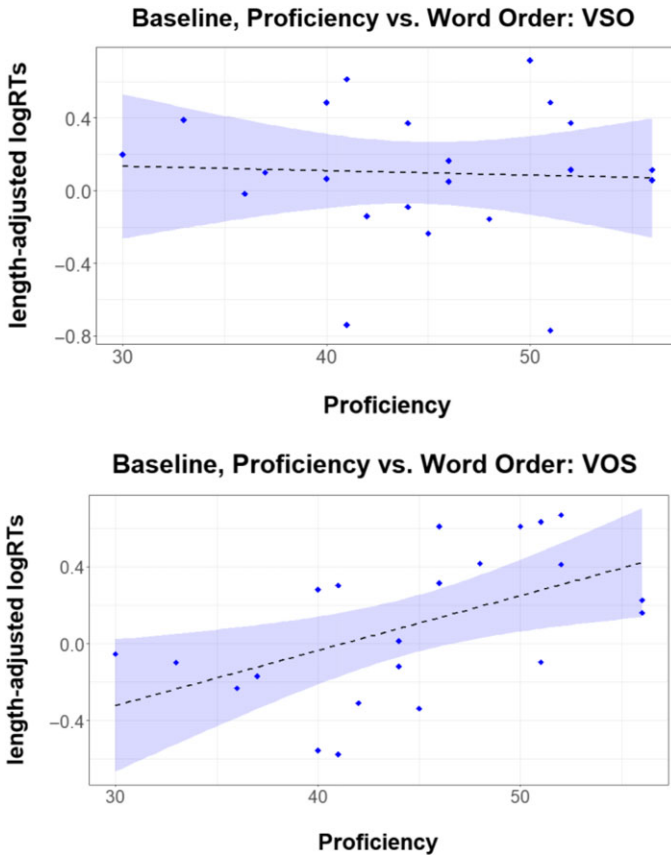
Figure 10. RTs by Proficiency and Word Order, Baseline Speakers.

for both types of focus for the heritage speaker group. We do not find evidence to suggest that responses differ by age of exposure for subject focus (Supplementary Table S29) or object focus (Supplementary Table S30) for the heritage speakers.

Although we find only one significant effect, for the sake of offering a full panorama of the data, we provide plots by proficiency and dominance in Supplementary File 1 as well.

6. Discussion

6.1. Summary of results

Overall, the results of both tasks show that the judgments and processing signature of heritage speakers very closely resemble those of the baseline speakers of Spanish.

The forced-choice task shows that in the subject focus conditions, both groups displayed the same preference, choosing VOS at a higher rate than VSO, as predicted by the syntactic literature. This preference is not absolute, since both groups hovered around two-thirds, evincing no group differences. The results from

the object and PP focus conditions, analogously, show that the groups did not differ, since the tests revealed a significant effect of focus context but no group effect nor interactions. Importantly, we see the effects of canonical order, because both groups chose canonical VOPP at higher rates, both under PP focus (predicted to be felicitous in the syntactic literature) and under object focus (predicted to be infelicitous). When examining the odds ratios, however, we found that the probability of choosing VPPO under object focus, as compared to PP focus, was three times higher for the heritage speakers and two times higher for the baseline speakers, suggesting both groups associate VPPO with object focus.

Results from the self-paced reading task show that bilingual speakers attend to the discourse context and integrate it incrementally, as shown by interactions of focus type by word order in the subject focus condition (VOS was read faster than VSO), with no differences between the two groups' processing signatures. In the object/PP focus condition, as with the judgment results, we see that canonical VOPP is *always* read faster, irrespective of context. Again, we find no group effects or interactions, showing that this processing advantage for canonical VOPP is present in both groups. For both tasks, we observe no effects by proficiency or dominance for the heritage speakers, while we find limited proficiency effects for the baseline speakers (discussed in section 6.4).

Overall, we find evidence that heritage speakers resemble baseline speakers in their judgments and processing signatures, with no differences between the groups on either task. We also observed strong advantages for canonical order where available.

6.2. Implications for factors affecting resilience

Let us consider the implications of our results by returning to the factors we identified in section 1, which purport to explain why some word orders are more vulnerable and others more resilient.

6.2.1. Transfer

We view cross-linguistic influence as an unlikely explanation for our results. In the subject/object focus conditions, we do not evince influence because English disallows V-initial orders such that no differences between the orders we studied (VSO/VOS) would have been expected, contrary to what we found. In terms of the object/PP focus conditions, it would be reasonable to posit influence from English's canonical order (SVOPP) increasing preference for the same order ([S]VOPP) in both judgments and processing. However, in the judgment experiment, object-final VPPO was more likely to be chosen under object focus than under PP focus, an effect consistent with an association between focus and final position, and which cannot be explained by transfer from English, in which VPPO is largely ungrammatical (excepting heavy NP-shift, see note 1). Furthermore, the advantage in processing canonical orders is not unique to the heritage speakers: both the baseline speakers and a monolingual group we previously examined (Hoot & Leal, 2020, 2022) show the same strong preference for canonical

orders, which cannot be the result of transfer from English (at least for the monolinguals).

6.2.2. Frequency
Given the reduced input characterizing heritage language acquisition, it is reasonable to hypothesize that infrequent constructions would be less likely to be acquired. The available evidence suggests that these orders are quite infrequent, especially VOS/VSO (Davidson, 2016; Ocampo, 2009). Given the limited evidence available, we must be careful not to over-interpret our results, but it is nonetheless noteworthy that our heritage speakers appear to pattern just like the baseline speakers in their judgments and processing of these very infrequent word orders, suggesting frequency may not be a good explanation.

6.2.3. Optional movements and variability
Polinsky (2018) has speculated that one reason for divergence could be the optional nature of some information-structural movements. Previous studies have shown that focus-final marking is optional because canonical orders fit many focus-marking strategies, including narrow focus on subjects and objects (Hoot & Leal, 2020; Leal et al., 2018). Yet word orders to mark focus appear to be resilient, suggesting the optional nature of movement is an unlikely reason for divergent acquisition.

6.2.4. Processing complexity
Importantly, our study confirms that bilingual speakers—both heritage and baseline—integrate the discourse content in online processing as soon as the information becomes available, corroborating the results of L1 processing studies (Kaiser & Trueswell, 2004; Slioussar, 2011; Weskott et al., 2011). Additionally, unlike what the Interface Hypothesis predicts (Sorace, 2011), we find no evidence that processing focus in context presents special difficulty for either bilingual group: the baseline speakers did not differ from the monolinguals, and the heritage speakers did not differ from the baseline speakers. Our results thus add evidence not only that offline knowledge of focus marking is resilient, but that its real-time processing is as well.

6.2.5. Salience
Laleko (2021) suggests that the resilience of focus structures may be due to *salience,* which she argues strengthens their representation in short-term memory, facilitating processing. In our experiment, salience is represented by the association of final sentence position (via syntactic movement) and focus. Our results show that the discourse context clearly modulates the processing of non-canonical word orders for both groups, at least when canonical orders are removed. If such a facilitatory effect is confirmed, it could lend credence to the idea that focus is resilient because it is informationally salient.

If focus is retained due to its salience, the question arises whether this explanation can be extended to other word orders indexing pragmatic information. For instance, although we reviewed several studies showing successful acquisition of focus-related word orders in section 1.2, other non-canonical orders may not show the same degree of success. For example, as mentioned earlier, Montrul (2010) found that dislocated objects instantiating a topic rather than a focus structure were difficult to interpret for heritage speakers, and topic-related structures have also been shown to present difficulties in L2 acquisition (Sorace et al., 2009) and L1 acquisition (Shin & Cairns, 2009). Perhaps (thematic) topics, which often express given or background information, are less resilient than focus because they are less informationally salient. Yet Leal Méndez, Rothman, and Slabakova (2015) provide evidence of successful acquisition of topic constructions (viz., clitic left dislocation) by heritage Spanish speakers, so the effect of informational salience on acquiring these discourse-related word orders remains an open question for future research.

6.3. Task effects
We noted in section 1.6 that the methods used could affect our interpretation of the data. Of the issues we pointed out, two merit further comment.

6.3.1. Task type
As noted in section 1.6, heritage speakers' performance may vary by task type. Our written FCT requires comparative judgments, which can increase the sensitivity and reliability of the instrument (Schütze & Sprouse, 2013; Stadthagen-González et al., 2018) while attenuating a possible "*yes*-bias" (Polinsky, 2018). Yet any judgment task is somewhat metalinguistic. The SPR task, conversely, has the advantage of being less metalinguistic and more implicit in nature. In this context, it is worth noting the *lack* of apparent task effects. That is, we observe convergence on the more metalinguistic FCT and the less metalinguistic SPR task, and we have previously observed similar convergence on other judgment tasks (Hoot, 2017) and with production tasks (Hoot et al., 2020; Leal et al., 2018). Although we recognize that task effects may well modulate our interpretation of the extant data, in the case of focus in Spanish there is remarkable consistency across task type that supports the conclusion that this phenomenon is largely resilient.

6.3.2. Proficiency and dominance
We found no effects of language dominance (or age of exposure) in either task, nor did we find effects of proficiency for the heritage speakers in either task. This latter result is somewhat surprising, given the heterogeneity among heritage speakers, which can lead to wide ranges in proficiency.

For baseline speakers, we found two effects by proficiency, which we considered surprising, given that their proficiency range was smaller than that of the heritage speakers and we had no expectation of proficiency effects for baseline speakers. On the FCT, higher proficiency correlated with greater likelihood of choosing focus-final VOPP than non-final VPPO under PP focus, but no other effects (e.g., no effect on the relative distribution of VOPP/VPPO under object focus).

On the SPR task, higher proficiency correlated with slower reading of the VOS word order only (not the other orders and not variable according to context). The proficiency effects for the baseline speakers in general were unexpected, and it is difficult to assign a unified interpretation to them.

One possible explanation is the nature of the proficiency task. Although the aim of the LexTALE_Esp is to discriminate among speakers by global proficiency (Izura et al., 2014), it is ultimately a measure of vocabulary size, whereby participants indicate knowledge of specific words (presented alongside non-words). Since vocabulary size has shown a close association with reading skills in L2 learners (Qian & Lin, 2020), we speculate that baseline speakers with a larger vocabulary might have more experience encountering non-canonical orders such as VPPO or VOS/VSO, which might be found more frequently in print. This conjecture does not, however, explain why the effects are not more consistent, instead appearing only in one context for each task, nor does it explain why VOS would be read slower as proficiency increases.

Another possible explanation is that these findings are Type I errors, that is, false positives. In null hypothesis significance testing, setting alpha at .05 equates to a 5% chance in any given test of rejecting the null hypothesis incorrectly. We performed more than twenty tests. Following standard practice in the field, we adjusted alpha to control the familywise error rate for post hoc pairwise comparisons within any given test, but we did not adjust alpha across models involving different factors.[8] Because these proficiency effects are inconsistent and not theoretically expected, it is reasonable to conclude that they may be spurious.

7. Conclusion

The key empirical takeaways from our study are presented in (6).

(6)   Main empirical findings
   a.   Heritage speakers did not differ from the baseline speakers in either their judgments or processing of non-canonical word orders to mark focus.
   b.   Canonical word orders have a processing advantage that can obscure contextual effects, but when canonical orders are removed, both groups are sensitive to discourse context in real-time processing.
   c.   In judgments, both groups associate non-canonical orders with the expected discourse contexts.

We interpreted these findings against the backdrop of factors previously proposed to explain heritage language acquisition of word order variation and concluded that our results did not support the proposals that heritage language divergence is the result of majority-language transfer/CLI, that divergence occurs due to non-canonical orders' low frequency, that divergence occurs when movements are optional, or that divergence occurs due to the inherent processing complexity of such constructions. We also found a degree of support for the proposal that focus was more likely than other non-canonical word orders to be resilient in heritage grammars due to its informational salience. This idea fits well with the evidence from L1 processing showing that listeners use the information structure of a

sentence to interpret it from the earliest stages of processing, facilitating sentence interpretation. Such an explanation has the promise of unifying two disparate facts—the apparent resilience of information-structural word order variation in the face of divergence for other types of word order variation and the L1 processing evidence—which we view as a fruitful avenue for future research.

Notes

1 We acknowledge that not only reduced input but the concomitant reduced opportunities to process language and produce output can contribute to divergences. For instance, Putnam and Sánchez (2013) note "[w]hat is crucial is the frequency of processing for comprehension and production purposes" (p. 480). We encapsulate all reduction in language use under the label of reduced input for ease of reference.

2 "Type frequency," which denotes the "number of distinct lexical items that can be substituted in a given [. . .] syntactic construction specifying the relation among words" (Ellis & Collins, 2009, p. 330), as opposed to "token frequency," which represents the simple count of the occurrence of a particular word, is the relevant construct for our study. Yet establishing type frequency for discourse-conditioned word orders can be a thorny endeavor, as one must consider not only the word orders but the discourse context, so not all cases of a given word order may serve as exemplars of the relevant type.

3 We should note that there is a debate regarding whether focus structures are at external interface or not. In this regard, we side with Slabakova (2011) and many others (e.g., Belletti, 2004; Rizzi, 1997) in suggesting that Topic and Focus both constitute external interface phenomena. However, Tsimpli and Sorace (2006) place it in core syntax, noting that focus is a relational feature that identifies new information with respect to the topic.

4 A reviewer rightly points out that the term "homeland" speaker can imply that heritage languages are the result of (recent) immigration, which is not the case in all heritage speaker communities. Although we agree that this interpretation is available, we retain the term because Polinsky herself acknowledges that such labels involve many degrees of idealization that do not precisely describe the reality on the ground because these terms tend to assume "static representations" (Polinsky, 2018, p. 9).

5 English allows VPPO in cases of "Heavy-NP Shift," as in *I finally bought at the market those exquisite pastries that I've had my eye on for the last two weeks*, where the object is (much) heavier prosodically. However, in simple sentences of the type we examine here, English does not permit VPPO: **I bought at the market the pastries*.

6 The BLP does not require that weekly-use averages add up to 100% across languages, so participants sometimes report totals exceeding 100%. Participants also varied in their interpretation of "How many years of classes (grammar, history, math, etc.) have you had in ENGLISH (primary school through university)?". We suspect those reporting very low numbers understood the question to mean language classes, specifically. Our heritage speakers were all college students who grew up in the U.S. and presumably graduated high school there; it seems unlikely anyone had only two years of school in an English-speaking environment. Nevertheless, we report the data as it was provided.

7 Concretely: Colombia, 6; Costa Rica, 1; Cuba, 1; Ecuador, 1; Honduras, 5; Mexico, 2; Venezuela, 1. For nine people, the data were lost due to an error with the BLP; they were all from the Chicago group and were most likely from Mexico. One person reported birth in Southern California but no exposure to English before age 20; knowing this person, we are certain they grew up outside the US, moving out of California as an infant, so we decided to include them nevertheless.

8 For further discussion, see Maxwell and Delaney (2004, Chapter 5), who advocate controlling the experimentwise error rate (i.e., adjusting alpha over every test conducted within a given study) rather than the familywise error rate, while also noting that ultimately the decision "involves a trade-off between Type I and Type II errors" (p. 196).

References

Albirini, A., Benmamoun, E., & Saadah, E. (2011). Grammatical features of Egyptian and Palestinian Arabic heritage speakers' oral production. *Studies in Second Language Acquisition*, **33**(2), 273–303. https://doi.org/10.1017/S0272263110000768

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, **68**(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Belletti, A. (2004). Aspects of the low IP area. In L. Rizzi (Ed.), *The Structure of CP and IP: The Cartography of Syntactic Structures, Volume 2* (pp. 16–52). Oxford University Press.

Benmamoun, E., Montrul, S., & Polinsky, M. (2013). Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics*, **39**(3/4), 129–181. https://doi.org/10.1515/tl-2013-0009

Birdsong, D., Gertken, L. M., & Amengual, M. (2012). *Bilingual Language Profile: An easy-to-use instrument to assess bilingualism.* COERLL, University of Texas at Austin. https://sites.la.utexas.edu/bilingual/

Bosch, J. E., & Unsworth, S. (2021). Cross-linguistic influence in word order: Effects of age, dominance and surface overlap. *Linguistic Approaches to Bilingualism*, **11**(6), 783–816. https://doi.org/10.1075/lab.18103.bos

Brehmer, B., & Usanova, I. (2015). Let's fix it? Cross-linguistic influence in word order patterns of Russian heritage speakers in Germany. In H. Peukert (Ed.), *Transfer effects in multilingual language development* (pp. 161–188). John Benjamins. https://doi.org/10.1075/hsld.4.08bre

Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. In G. Bouma, I. Krämer, & J. Zwarts (Eds.), *Cognitive foundations of interpretation* (pp. 69–94). Royal Netherlands Academy of Arts and Sciences.

Büring, D., & Gutiérrez-Bravo, R. (2001). Focus-related constituent order variation without the NSR: A prosody-based crosslinguistic analysis. *Syntax at Santa Cruz*, **3**, 41–58.

Bybee, J., & Thompson, S. (1997). Three frequency effects in syntax. *Annual Proceedings of the Berkeley Linguistics Society*, **23**(1), 378–388. https://doi.org/10.3765/bls.v23i1.1293

Casielles-Suárez, E. (2004). *The syntax-information structure interface: Evidence from Spanish and English.* Routledge.

Davidson, C. L. (2016). *Deriving VSO sentences in Spanish* [Dissertation, Indiana University]. http://www.proquest.com/docview/1868414790/abstract/E2F06A26272140F4PQ/1

Davies, M. (2002). *Corpus del Español: Genre/Historical.* http://www.corpusdelespanol.org/hist-gen/

Davies, M. (2006). *A frequency dictionary of Spanish: Core vocabulary for learners.* Routledge.

de Prada Pérez, A., & Pascual y Cabo, D. (2012). Interface heritage speech across proficiencies: Unaccusativity, focus, and subject position in Spanish. In K. Geeslin & M. Díaz-Campos (Eds.),

Selected proceedings of the 14th Hispanic linguistics symposium (pp. 308–318). Cascadilla Proceedings Project.

Dekydtspotter, L., & Renaud, C. (2014). On second language processing and grammatical development: The parser in second language acquisition. *Linguistic Approaches to Bilingualism*, **4**(2), 131–165. https://doi.org/10.1075/lab.4.2.01dek

Domínguez, L. (2004). *The syntax and prosody of focus in Spanish* [Dissertation, Boston University].

Eddington, D. (2015). *Statistics for linguists: A step-by-step guide for novices*. Cambridge Scholars Publishing.

Eide, K. M., & Hjelde, A. (2015). Borrowing modal elements into American Norwegian: The case of suppose(d). In J. B. Johannessen & J. C. Salmons (Eds.), *Germanic heritage languages in North America: Acquisition, attrition and change* (pp. 256–280). John Benjamins. https://doi.org/10.1075/silv.18.12eid

Ellis, N., & Collins, L. (2009). Input and second language acquisition: The roles of frequency, form, and function: Introduction to the Special Issue. *The Modern Language Journal*, **93**(3), 329–335. https://doi.org/10.1111/j.1540-4781.2009.00893.x

Feldhausen, I., & del Vanrell, M. M. (2014). Prosody, focus and word order in Catalan and Spanish: An Optimality Theoretic approach. In S. Fuchs, M. Grice, A. Hermes, L. Lancia, & D. Mücke (Eds.), *Proceedings of the 10th International seminar on speech production (ISSP)* (pp. 122–125). Universität zu Köln.

Ferré, P., & Brysbaert, M. (2017). Can Lextale-Esp discriminate between groups of highly proficient Catalan–Spanish bilinguals with different language dominances? *Behavior Research Methods*, **49**(2), 717–723. https://doi.org/10.3758/s13428-016-0728-y

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLOS ONE*, **8**(10), e77661. https://doi.org/10.1371/journal.pone.0077661

Gellon, S. (2015). *La adquisición del orden verbo sujeto en español por estudiantes de herencia de español* [Dissertation, University of Houston]. https://uh-ir.tdl.org/handle/10657/1268

Gómez Soler, I., & Pascual y Cabo, D. (2016). Syntactic reflexes of information structure in heritage Spanish: Evidence from psych-predicate constructions. *International Review of Applied Linguistics in Language Teaching*, **54**(4), 291–317. https://doi.org/10.1515/iral-2015-0036

Gómez Soler, I., & Pascual y Cabo, D. (2018). On focus and weight in Spanish as a heritage language. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics*, **31**(2), 437–466. https://doi.org/10.1075/resla.16021.gom

Gutiérrez-Bravo, R. (2020). La sintaxis del español de México: Un esbozo. *Cuadernos de la ALFAL*, **12**(2), 44–70.

Håkansson, G. (1995). Syntax and morphology in language attrition: A study of five bilingual expatriate Swedes. *International Journal of Applied Linguistics*, **5**(2), 153–169. https://doi.org/10.1111/j.1473-4192.1995.tb00078.x

Heidinger, S. (2015). Optionality and preferences in Spanish postverbal constituent order: An OT account without basic constituent order. *Lingua*, **162**, 102–127. https://doi.org/10.1016/j.lingua.2015.05.003

Hoot, B. (2016). Narrow presentational focus in Mexican Spanish: Experimental evidence. *Probus*, **28**(2), 335–365. https://doi.org/10.1515/probus-2014-0004

Hoot, B. (2017). Narrow presentational focus in heritage Spanish and the syntax-discourse interface. *Linguistic Approaches to Bilingualism*, **7**(1), 63–95. https://doi.org/10.1075/lab.14021.hoo

Hoot, B. (2019). Focus in heritage Hungarian. *Language Acquisition*, **26**(1), 46–72. https://doi.org/10.1080/10489223.2017.1393078

Hoot, B., & Ebert, S. (2021). On the position of subjects in Spanish: Evidence from code-switching. *Glossa: A Journal of General Linguistics*, **6**(1), 73. https://doi.org/10.5334/gjgl.1449

Hoot, B., & Leal, T. (2020). Processing subject focus across two Spanish varieties. *Probus*, **32**(1), 93–127. https://doi.org/10.1515/probus-2019-0004

Hoot, B., & Leal, T. (2022). Crosslinguistic influence from Catalan and Yucatec Maya on judgments and processing of Spanish focus. *Linguistic Approaches to Bilingualism*, *Online Ahead of Print*. https://doi.org/10.1075/lab.21020.hoo

Hoot, B., Leal, T., & Destruel, E. (2020). Object focus marking in Spanish: An investigation using three tasks. *Glossa: A Journal of General Linguistics*, **5**(1), 70. https://doi.org/10.5334/gjgl.1160

Hopp, H., & Putnam, M. T. (2015). Syntactic restructuring in heritage grammars: Word order variation in Moundridge Schweitzer German. *Linguistic Approaches to Bilingualism*, **5**(2), 180–214. https://doi.org/10.1075/lab.5.2.02hop

Isurin, L., & Ivanova-Sullivan, T. (2008). Lost in between: The case of Russian heritage speakers. *Heritage Language Journal*, **6**(1), 72–103. https://doi.org/10.46538/hlj.6.1.4

Izura, C., Cuetos, F., & Brysbaert, M. (2014). Lextale-Esp: A test to rapidly and efficiently assess the Spanish vocabulary size. *Psicologica: International Journal of Methodology and Experimental Psychology*, **35**(1), 49–66.

Johannessen, J. B. (2015). Attrition in an American Norwegian heritage language speaker. In J. B. Johannessen & J. C. Salmons (Eds.), *Germanic heritage languages in North America: Acquisition, attrition and change* (pp. 46–71). John Benjamins. https://doi.org/10.1075/silv.18.02joh

Kahan Newman, Z. (2015). Discourse markers in the narratives of New York Hasidim: More V2 attrition. In J. B. Johannessen & J. C. Salmons (Eds.), *Germanic heritage languages in North America: Acquisition, attrition and change* (pp. 178–198). John Benjamins. https://doi.org/10.1075/silv.18.08kah

Kaiser, E. (2016). Information structure and language comprehension. In C. Féry & S. Ishihara (Eds.), *The Oxford handbook of information structure* (pp. 523–540). Oxford University Press.

Kaiser, E., & Trueswell, J. C. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition*, **94**(2), 113–147. https://doi.org/10.1016/j.cognition.2004.01.002

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, **55**(3–4), 243–276. https://doi.org/10.1556/aling.55.2008.3-4.2

Laleko, O. (2021). Discourse and information structure in heritage languages. In M. Polinsky & S. Montrul (Eds.), *The Cambridge Handbook of Heritage Languages and Linguistics* (pp. 691–727). Cambridge University Press. https://doi.org/10.1017/9781108766340.031

Laleko, O. (2022). Word order and information structure in heritage and L2 Russian: Focus and unaccusativity effects in subject inversion. *International Journal of Bilingualism*, *Online Ahead of Print*. https://doi.org/10.1177/13670069211063674

Laleko, O., & Dubinina, I. (2018). Word order production in heritage Russian: Perspectives from linguistics and pedagogy. In S. Bauckus & S. Kresin (Eds.), *Connecting across languages and cultures: A heritage language festschrift in honor of Olga Kagan* (pp. 191–215). Slavica Publishers.

Larsson, I., Tingsell, S., & Andréasson, M. (2015). Variation and change in American Swedish. In J. B. Johannessen & J. C. Salmons (Eds.), *Germanic heritage languages in North America: Acquisition, attrition and change* (pp. 359–388). John Benjamins. https://doi.org/10.1075/silv.18.16lar

Leal Méndez, T., Rothman, J., & Slabakova, R. (2015). Discourse-sensitive clitic-doubled dislocations in heritage Spanish. *Lingua*, **155**, 85–97. https://doi.org/10.1016/j.lingua.2014.01.002

Leal, T., Destruel, E., & Hoot, B. (2018). The realization of information focus in monolingual and bilingual native Spanish. *Linguistic Approaches to Bilingualism*, **8**(2), 217–251. https://doi.org/10.1075/lab.16009.lea

Leal, T., & Hoot, B. (2022). L2 representation and processing of Spanish focus. *Language Acquisition*, **29**(4), 410–440. https://doi.org/10.1080/10489223.2022.2049599

Leal, T., Rothman, J., & Slabakova, R. (2014). A rare structure at the syntax-discourse interface: Heritage and Spanish-dominant native speakers weigh in. *Language Acquisition*, **21**(4), 411–429. https://doi.org/10.1080/10489223.2014.892946

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

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (Second edition). Lawrence Erlbaum.

Meteyard, L., & Davies, R. A. I. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, **112**, 104092. https://doi.org/10.1016/j.jml.2020.104092

Miller, J., & Weinert, R. (2009). *Spontaneous spoken language: Syntax and discourse*. Oxford University Press.

Montrul, S. (2005). Second language acquisition and first language loss in adult early bilinguals: Exploring some differences and similarities. *Second Language Research*, **21**(3), 199–249. https://doi.org/10.1191/0267658305sr247oa

Montrul, S. (2010). How similar are adult second language learners and Spanish heritage speakers? Spanish clitics and word order. *Applied Psycholinguistics*, **31**(01), 167–207. https://doi.org/10.1017/S01427 1640999021X

Montrul, S. (2016). *The acquisition of heritage languages*. Cambridge University Press.

Montrul, S., Bhatt, R., & Girju, R. (2015). Differential object marking in Spanish, Hindi, and Romanian as heritage languages. *Language*, **91**(3), 564–610. https://doi.org/10.1353/lan.2015.0035

Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, **58**(3), 503–553. https://doi.org/10.1111/j.1467-9922.2008.00449.x

Montrul, S., & Polinsky, M. (2011). Why not heritage speakers? *Linguistic Approaches to Bilingualism*, **1**(1), 58–62. https://doi.org/10.1075/lab.1.1.07mon

Ocampo, F. (2009). El orden de palabras en el español hablado: La construcción sujeto verbo objeto directo. In M. Veyrat Rigat & E. Serra Alegre (Eds.), *La lingüística como reto epistemológico y como acción ocial: Estudios dedicados al Profesor Ángel López García con ocasión de su sexagésimo aniversario* (pp. 501–511). Arco Libros.

O'Grady, W., Yamashita, Y., & Lee, S.-Y. (2005). A note on canonical word order. *Applied Linguistics*, **26**(3), 453–458. https://doi.org/10.1093/applin/ami019

Olarrea, A. (2012). Word order and information structure. In J. I. Hualde, A. Olarrea, & E. O'Rourke (Eds.), *The handbook of Hispanic linguistics* (pp. 603–628). Wiley-Blackwell.

Polinsky, M. (2006). Incomplete acquisition: American Russian. *Journal of Slavic Linguistics*, **14**(2), 191–262.

Polinsky, M. (2018). *Heritage Languages and Their Speakers*. Cambridge University Press. https://doi.org/10.1017/9781107252349

Polinsky, M., & Scontras, G. (2020). Understanding heritage languages. *Bilingualism: Language and Cognition*, **23**, 4–20. https://doi.org/10.1017/S1366728919000245

Poplack, S. (2001). Variability, frequency, and productivity in the irrealis domain of French. In J. L. Bybee & P. J. Hopper (Eds.), *Frequency and the Emergence of Linguistic Structure* (pp. 405–428). John Benjamins. https://doi.org/10.1075/tsl.45.20pop

Putnam, M. T., & Sánchez, L. (2013). What's so incomplete about incomplete acquisition?: A prolegomenon to modeling heritage language grammars. *Linguistic Approaches to Bilingualism*, **3**(4), 478–508. https://doi.org/10.1075/lab.3.4.04put

Qian, D. D., & Lin, L. H. F. (2020). The relationship between vocabulary knowledge and language proficiency. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 66–80). Routledge.

Rizzi, L. (1997). The Fine Structure of the Left Periphery. In L. Haegeman (Ed.), *Elements of Grammar: Handbook in generative syntax* (pp. 281–337). Springer Netherlands. https://doi.org/10.1007/978-94-011-5420-8_7

Rohde, D. (2003). *Linger* (2.94). ~http://tedlab.mit.edu/~dr/Linger/

Schütze, C. T., & Sprouse, J. (2013). Judgment data. In R. J. Podesva & D. Sharma (Eds.), *Research methods in linguistics* (pp. 27–50). Cambridge University Press.

Selkirk, E. (1995). Sentence prosody, intonation, stress, and phrasing. In J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 550–569). Blackwell.

Sequeros-Valle, J., Hoot, B., & Cabrelli, J. (2020). Clitic-doubled left dislocation in heritage Spanish: Judgment versus production data. *Languages*, **5**(4), 47. https://doi.org/10.3390/languages5040047

Shin, N. L., & Cairns, H. S. (2009). Subject pronouns in child Spanish and continuity of reference. In J. Collentine, M. García, B. Lafford, & F. M. Marín (Eds.), *Selected proceedings of the 11th Hispanic linguistics symposium* (pp. 155–164). Cascadilla Proceedings Project. https://www.lingref.com/cpp/hls/11/abstract2210.html

Slabakova, R. (2011). Which features are at the syntax–pragmatics interface? *Linguistic Approaches to Bilingualism*, **1**(1), 89–93. https://doi.org/10.1075/lab.1.1.14sla

Slioussar, N. (2011). Processing of a free word order language: The role of syntax and context. *Journal of Psycholinguistic Research*, **40**(4), 291–306. https://doi.org/10.1007/s10936-011-9171-5

Song, M., O'Grady, W., Cho, S., & Lee, M. (1997). The learning and teaching of Korean in community schools. *The Korean Language in America*, **2**, 111–127.

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism*, **1**(1), 1–33. https://doi.org/10.1075/lab.1.1.01sor

Sorace, A., Serratrice, L., Filiaci, F., & Baldo, M. (2009). Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua*, **119**(3), 460–477. https://doi.org/10.1016/j.lingua.2008.09.008

Stadthagen-González, H., López, L., Parafita Couto, M. C., & Párraga, C. A. (2018). Using two-alternative forced choice tasks and Thurstone's law of comparative judgments for code-switching research. *Linguistic Approaches to Bilingualism*, **8**(1), 67–97. https://doi.org/10.1075/lab.16030.sta

Tsimpli, I. M., & Sorace, A. (2006). Differentiating interfaces: L2 performance in syntax-semantics and syntax-discourse phenomena. In *Proceedings of the 30th annual Boston University Conference on Language Development (BUCLD 30)* (pp. 653–664). Cascadilla Press.

van Osch, B., & Sleeman, P. (2018). Subject position in Spanish as a heritage language in the Netherlands: External and internal interface factors. In A. Gavarró (Ed.), *On the Acquisition of the Syntax of Romance* (pp. 187–214). John Benjamins. https://doi.org/10.1075/lald.62.09osc

Weskott, T., Hörnig, R., Fanselow, G., & Kliegl, R. (2011). Contextual licensing of marked OVS word order in German. *Linguistische Berichte*, **225**, 3–18.

West, B. T., Welch, K. B., Gałecki, A. T., & Gillespie, B. W. (2015). *Linear mixed models: A practical guide using statistical software (Second edition)*. CRC Press.

Zubizarreta, M. L. (1998). *Prosody, focus, and word order*. MIT Press.