L2 Representation and Processing of Spanish Focus

Tania Leal
University of Nevada, Reno

Bradley Hoot
DePaul University

Abstract

Research on second language (L2) acquisition has identified linguistic domains that appear to be especially difficult to learn—one such sticking point being syntactic structures that depend on the surrounding discourse. The Interface Hypothesis (IH) explains what makes such constructions problematic by appealing to a modular view of language, arguing that integrating knowledge from language-internal domains (e.g., syntax) with language-external domains (e.g., discourse) overwhelms the finite processing resources of L2 learners, especially when integration happens in real time. We test the IH with a syntax-discourse interface phenomenon in Spanish: information focus. The facts about information focus in L1 and L2 Spanish have been enthusiastically debated, but what is missing from these debates is evidence that directly indexes processing, which is essential to evaluate the IH. We use an offline forced-choice judgment task and an online self-paced reading task to provide a new source of evidence of L2 acquisition of Spanish focus. We find that L1-English/L2-Spanish learners largely resemble L1-Spanish natives in both their judgments and their processing of focus, contrary to the predictions of the IH.

Keywords: Information focus, Interface Hypothesis, syntax-discourse interface, self-paced reading, L2 Spanish

Acknowledgments: We are very grateful to Eve Edmunds-Díez and Heather Slawny for their assistance with data collection, and to DePaul's College of Liberal Arts and Social Sciences and University Research Council for financial support. Thank you to many colleagues who helped with participant recruitment and to everyone who participated in our study. We thank Jennifer Cabrelli for help with some data analysis questions and the audience at GASLA 15 (Reno, NV)

for their useful feedback. Finally, we would like to thank the reviewers of this article, who offered insightful and supportive critiques of our work.

1. Background & Motivation

1.1. *Second-language acquisition at the syntax/discourse interface*

Linguists studying second language (L2) acquisition have avidly described areas of learners' linguistic systems that appear to represent sticking points for acquisition. Under a modular view of language whereby each module operates under module- and interface-specific rules (e.g., Chomsky 1995; Jackendoff 2002), researchers have contended that difficulties stem from the specific linguistic modules involved. One influential account, the Interface Hypothesis (IH), proposes that structures at the syntax-discourse interface pose special challenges to bilinguals because syntax is *internal* to the grammar while discourse is grammar-*external* (Sorace 2011, 2012; Sorace & Serratrice 2009). Early theorizing on the IH suggested two potential explanations for the difficulty L2 learners have evinced with these structures (Sorace 2000, 2004; Sorace & Filiaci 2006). One explanation proposes difficulties stem from an emerging *representational* deficit or underspecification in the L2 grammar. The second explanation suggests that difficulties derive from *processing* deficits involving speakers' ability to coordinate different types of knowledge. A recent treatment (Sorace 2011, 2012) favors the latter explanation and proposes that the difficulties stem from resource constraints on the real-time integration of knowledge.

In our investigation, we concentrate on an external-interface property: the realization of information focus in L2 Spanish. Focus realization involves the integration of syntactic and discourse knowledge and, in Spanish, it has been claimed to trigger word-order alternations. These variations are rarely 'optional' and often involve subtle restrictions on pragmatics, semantics, or interpretation-related information structuring (Bresnan et al. 2007). Unsurprisingly,

it has been suggested that the complex process of integrating these disparate sources of information is at the center of the difficulty with the syntax-discourse interface, especially when this knowledge must be deployed in real time. Few studies, however, have used methods indexing processing, which provide essential evidence needed to evaluate interface vulnerability that appeal to resource limitations. We aim to fill this gap by using an instrument that allows us to measure reaction times.

1.2. *The Interface Hypothesis (IH): interfaces and other factors in acquisition*

To recapitulate, the IH proposes that structures involving external interfaces are challenging for bilinguals because of processing constraints rather than because of under-representation/under-specification in learners' grammars (Sorace 2011). Why should this be the case? Sorace offers two possible explanations. First, L2 learners' ability to manage computational constraints may be less detailed or automatic, resulting in reduced processing efficiency. Another possibility is that bilinguals work under reduced cognitive resources when integrating discourse and syntactic information in real time such that when learners are put under additional cognitive pressure, they might show different processing patterns. This would be the case even if learners can show nativelike responses in measures such as untimed judgment tasks. In methodological terms, the IH's focus on processing entails the use of methodologies which can either induce processing pressure (e.g., speeded tasks) or index it (e.g., methods measuring latencies). In the case of the acquisition of Spanish as an L2, the use of online methods is increasingly common, although much less so for external interface properties (but see Leal, Slabakova & Farmer 2017). To wit, although the IH has been challenged in several fronts (e.g., Ivanov 2009; Rothman 2009, a.o.), few of these investigations actually use methodologies that can speak to processing explanations because these investigations do not use methods that either impose time pressure or index

reaction times.

A host of factors can affect development and acquisition. Thus, even if the IH were to be found to be a reasonable explanation for external-interface difficulties, it is unlikely that it is the *only* explanation. As such, we consider the role of additional factors to determine whether these play a role in the processing of focus in L2 Spanish. Because *proficiency* has shown a positive relationship with a large gamut of linguistic outcomes in L2 acquisition, it seems only natural to investigate its role. Sorace (2011), however, contended that the IH pertains only to end-state learners, such that proficiency is not expected to play a role because the hypothesis is only meant to apply to near-native speakers. White (2011) challenged Sorace's reasoning, though, questioning why external-interface difficulties should arise in end-state learners but not in learners at intermediate stages. White notes that investigating the acquisition of syntax-discourse interface properties *throughout* development helps researchers establish whether the effects suggested by the IH are indeed long-term and residual. In fact, previous research has shown that difficulties at the syntax-discourse interface do affect intermediate learners (e.g., Hertel 2003; Pladevall Ballester 2010). In our investigation, we follow White's suggestion and include speakers with a broad range of proficiency scores to examine their development by treating proficiency as a continuous variable (see Leal 2018).

Like proficiency, *prior linguistic knowledge* is widely recognized to affect the acquisition of an L2/Ln. In this case, we operationalize previous knowledge as L1 transfer, which is the initial state of L2 acquisition. The IH is unique in this regard because it proposes that L1 transfer is *not* behind the difficulties at external interfaces, and divergences between native speakers and bilinguals stem from bilingualism itself instead. Sorace based her predictions on evidence showing that even when the L1 and L2 structures worked similarly (such that positive transfer

could be expected) bilinguals still exhibited non-native patterns, especially in the area of the acquisition of discourse-constrained null/overt subject distributions (Bini 1993; Lozano 2006a; Margaza & Bel 2006; Sorace & Serratrice 2009). In our case, we investigate English-Spanish bilinguals, where the learning task for information focus would include acquiring new syntactic constructions and their discourse distributions, including the possibility of focus-final orders in particular discourse contexts (see §1.4).

1.3. *Focus in Spanish*

When people communicate, they adapt their sentences to fit in with what has been said such that utterances contain less informative parts that connect it to the previous discourse—typically called *background*—and more informative parts that contribute new information—often called *focus*. Crosslinguistically, speakers make the focus more prominent than the background, although the mechanisms for marking prominence vary across languages and include particular word orders, pitch accents, or focus-marking morphemes. Additionally, these options can be used in tandem (i.e., movement to positions where focus receives a particular pitch accent). Focus thus fundamentally involves the interaction of sentence-level properties (largely syntax and prosody) with the existing discourse context (pragmatics/discourse), therefore constituting a syntax/discourse interface phenomenon (*pace* Tsimpli & Sorace 2006). Along with other elements of information structure, focus is of interest to linguists because it raises questions about how different parts of the grammatical system interact with one another and with the context (see Erteschik-Shir, 2007; López, 2009; Reinhart, 2006).

We concentrate here on the realization of *information focus*, as opposed to *contrastive focus*. *Contrastive* (or *correction*/*identificational*) focus has several properties that distinguish it from *information* (or *presentational*) focus. Focus can also be distinguished by its scope: whether

new information is provided by one constituent or a larger string (including whole sentences). Here, we only consider the former case, known as *narrow focus*. When we say "focus" for brevity's sake, we mean *narrow information focus*. See Büring (2009) and Krifka (2007) for a typology. While we recognize that intonation plays an essential role in focus realization, we concentrate our discussion on word order because that is the factor our experiments test.

In Spanish, it has long been observed that new information usually appears at the end of sentences (Bolinger 1954). This observation has led to widespread claims suggesting focus *must* appear sentence-finally (Büring 2009; Büring & Gutiérrez-Bravo 2001; Contreras 1978; Costa 2001; Domínguez 2004; Escandell Vidal & Leonetti 2019; Fábregas 2016; Gutiérrez-Bravo 2002, 2008; Leonetti 2014; Ortega-Santos 2006; Samek-Lodovici 2001; Zubizarreta 1998). Under this view, subject-final (1b) would be a felicitous answer to (1a), whereas stressing the subject in situ (1c) would be infelicitous. The same is claimed for focus on direct objects (2): movement to final position is considered obligatory (if there is material after the object), as in (2b), whereas stressing the object in its canonical position (2c) is claimed to not be possible.

(1) Subject focus

a. *¿Quién compró un erizo?*

'Who bought a hedgehog?'

b. *Compró un erizo* ***[Lori]***$_F$.

bought a hedgehog Lori

'Lori bought a hedgehog.'

c. *#* ***[Lori]***$_F$ *compró un erizo*.

(2) Object focus

a. *¿Qué compró Lori para su hermana?*

‘What did Lori buy for her sister?’

b. *Lori compró para su hermana [un **erizo**]$_F$.*

Lori bought for her sister a hedgehog

‘Lori bought a hedgehog for her sister.’

c. *# Lori compró [un **erizo**]$_F$ para su hermana.*

These descriptions, however, hardly constitute the whole story. In fact, the empirical picture is more nuanced because a growing number of quantitative experimental studies using different methodologies show that Spanish is more flexible than most syntactic proposals assert because both (1b) and (1c) are available subject-focus realizations, while (2b) and (2c) can realize object focus (see Hoot, Leal & Destruel 2020 for an overview of experimental results on Spanish focus).

Importantly, although numerous experiments have found a proclivity for in-situ marking, preferences are also modulated by other factors. For instance, Spanish speakers more strongly prefer in-situ focus marking with subjects than with objects (Hoot 2016). The details of focus realization also likely vary across dialects, but this factor has received little explicit attention. Hoot and Leal (2020) directly compared two varieties of Spanish and found that they largely behaved similarly, although they did note some evidence suggesting that Mexican Spanish may be more likely to resist discourse-conditioned word order alterations than Peninsular Spanish (see also Gutiérrez-Bravo 2020 for a similar claim). However, relatively few details are known about dialectal differences in information structure, and similar experimental results have been found across several varieties. Results also appear to be affected by design factors that are orthogonal to purely information-structural questions; in particular, the use of canonical word orders in the test sentences, presence or absence of other non-focal constituents, and syntactic

function of the argument in focus seem to affect results (Hoot, Leal & Destruel 2020). In our investigation, we control these factors where possible or address them explicitly where we cannot.

1.4. *Learning tasks*

Under the traditional syntactic view—i.e., focus marked in Spanish by syntactic movement and in English via in-situ stress—the learning task for L1-English/L2-Spanish learners is straightforward: learners must acquire the syntax of Spanish word order and recognize the discourse triggers that enable non-canonical orders. Indeed, many previous studies have assumed this learning task. However, predictions about L2 acquisition rest on descriptions of the target behavior, which have improved in recent years; thus, our description of the learning task reflects the experimental results discussed in the previous section. Further, we concentrate only on the specific structures we investigate.

Because of the role canonicity plays in judgments and processing, we chose to avoid canonical orders where possible. Crucially, Spanish has another word order with a post-verbal—but not final—subject: VSO. Noncanonical VSO is also contextually restricted: it can realize focus on the object but, crucially, not the subject (Domínguez 2004:74; Zubizarreta 1998:125). We thus compare VOS against VSO. Because English lacks verb-initial word orders, the task for L2 Spanish learners is to acquire, first, the syntax of these non-canonical constructions, then acquire the fact that, while the context never *requires* one or the other, each is only possible in one context (Table 1).

Table 1. Learning task, subject focus

	Spanish		English		
	Subject focus context	Object focus context	Subject focus context	Object focus context	Learning task
VSO	✗	✓	✗	✗	VSO syntax + contextual restrictions
VOS	✓	✗	✗	✗	VOS syntax + contextual restrictions

We expand beyond previous work by including factors relating to canonicity, clitics, and the argument in focus. First, we consider object focus when there is another constituent in the VP. Like VOS/VSO, VPPO word order is generally ungrammatical in English. Unlike for subject focus, though, there is no counterpart non-canonical word order. In this case, canonicity is an unavoidable confound that we recognize and incorporate into our analysis. The learning task for object focus is outlined in Table 2.

Table 2. Learning task, object focus

	Spanish		English		
	Object focus context	PP focus context	Object focus context	PP focus context	Learning task
(S)VOPP	✓	✓	✓	✓	Minimal[1]
(S)VPPO	✓	✗	✗	✗	VPPO syntax + contextual restriction

Finally, because previous studies suggest monolingual speakers are more accepting of focus-final word orders when objects are replaced with clitic pronouns (Gupton 2017), we included sentences with clitics in one task. As with the object focus case, though, the tradeoff is that canonical word orders are unavoidable: the only two options available are canonical SV and

[1] It is likely that the two languages differ prosodically, but the syntax and focus distribution is the same here.

non-canonical VS. For L1-English learners, the learning task entails acquiring the syntax of clitics, the syntax of subject-final orders, and the restriction that VS is only possible under subject focus (Table 3).

Table 3. Learning task, subject focus with clitic objects

	Spanish		English		
	Subject focus context	Object focus context	Subject focus context	Object focus context	Learning task
SclV	✓	✓	✗	✗	Syntax of clitics
clVS	✓	✗	✗	✗	Syntax of clitics, syntax of VS, contextual restrictions

In terms of the IH, then, processing signatures or judgments that mirror native speaker behavior should be taken as evidence against the hypothesis because learners are predicted to show residual optionality even at very advanced proficiency levels. Likewise, evidence that the learners have variable or optional judgements or processing patterns would constitute support for the IH, since information focus integrates syntactic and discourse knowledge. Finally, we point out that these learning tasks inform our experiment design, but our study is also informed by previous evidence on the L2 acquisition of focus, which we examine next.

1.5. *L2 Spanish acquisition of information focus*

Early work supported the notion that the syntax-discourse constructions are especially problematic for acquisition. An influential study of L1-English/L2-Spanish learners by Hertel (2003) using a written production task where responses where elicited using an information-seeking question found that discourse restrictions were acquired relatively late, leading Hertel to argue that acquiring them is especially difficult. Subsequent work (Lozano 2006a, 2006b) with a similar design went further, finding persistent problems with the acquisition of focus with

advanced learners (L1-English/L1-Greek/L2-Peninsular Spanish). Pladevall Ballester's (2010) study of five-, ten-, and seventeen-year-old British children (L1-English) in a Peninsular Spanish immersion school also concluded that discourse restrictions presented persistent problems, even if they were eventually acquirable. The children successfully acquired the syntactic properties allowing subject postposing under focus and most of the discourse restrictions, yet even advanced learners displayed residual optionality on postverbal subjects.

Recent work, in contrast, has found relatively little difficulty with L2 focus acquisition. Gathering evidence from acceptability judgments, a study of beginner, intermediate, and advanced L1-English students in the U.K. (Domínguez & Arche 2008, 2014) attributed divergences to syntactic deficits, not to problems with the discourse context. Additionally, Gupton (2017) found that advanced L1-American English learners gave target-like judgments of subject focus, as did Leal Méndez and Slabakova (2011), leading the latter to claim that judging contextual felicity posed little problem even for learners at intermediate stages. These studies show, albeit indirectly, that factors beyond information structure can affect judgments. Gupton asked participants to rate sentences with transitive verbs where objects were replaced with clitics (S-clitic-V vs. clitic-V-S), while Leal Méndez and Slabakova gave participants a choice between VSO and VOS. Notably, many of the previously mentioned experiments tested SV/VS with *intransitive* verbs, which could account for some differences.

Recent work with oral tasks has also shown little difficulty with focus marking. Using a speeded production task, Leal, Destruel, and Hoot (2019) found target-like production of subject and object focus by L1-American English learners of Spanish. Intermediate and advanced learners almost exclusively produced in-situ subject focus (SVO), but so did two monolingual control groups. Learners also approached target-like production of object focus as proficiency

increased. Similarly, Kim (2016) tested perception and production of nuclear stress by L1-American English learners. Kim also found that L2 learners overwhelmingly produced in-situ subject focus, like native controls, although she did observe prosodic differences in the L2 learners.

In sum, although early studies supported the notion that focus posed special difficulty for learners, most work in the last decade has differed, such that the field has reached no consensus. One potential reason is that, depending how the data is analyzed, some of these conclusions are open to reinterpretation. For example, although Pladevall Ballester (2010) concluded that the most advanced child learners showed "residual optionality" because they accepted SV word order along with VS for subject focus, a growing experimental literature shows that L1-Spanish speakers accept and produce focused in-situ subjects. Likewise, Domínguez and Arche (2014) contended that difficulties with non-canonical word orders was purely syntactic. Yet the most advanced group rated VS word orders equal to or better than SV orders in all contexts, which could be taken to indicate that they have acquired the syntax, but not the discourse restrictions. Our intention here is not to criticize these authors, but to note that there remain open empirical questions for the field to address.

Other limitations should be considered. First, there is limited diversity in the sampling of learners and target varieties. Except for Lozano's (2006a, 2006b) L1-Greek speakers, all research on L2 acquisition of Spanish focus we are aware of concerns L1-English learners (U.S./U.K.). While we do not address this problem—our learners are L1-American English speakers—it is worth noting. Second, the typical control group and/or the presumed target is Peninsular Spanish, while some of the U.S. studies have a control group represented by a wide mixture of dialects (often college faculty). In our study, we introduce a different control: monolingual speakers of

Yucatecan Spanish (Mexico). Although Yucatecan Spanish is a unique dialect within Mexico (Gutiérrez-Bravo 2020), using a single-dialect control group provides a more uniform baseline for comparison, one that is likely closer to U.S. learners' input, given the prevalence of Mexican Spanish in the U.S. as a whole and in the Midwest (where data was collected) in particular. Third, most previous work examines only focus on subjects with a limited range of sentence types. Yet monolingual evidence suggests that syntactic function (subject/object) and the presence of other constituents in the sentence also affect focus marking. In our study, we address this limitation directly by comparing focus-marking on subjects and objects and by including sentences in which objects are replaced by clitic pronouns.

A fourth limitation concerns the methods. While we recognize the value of tasks such as acceptability judgments—a task we also employ here—we are aware of their drawbacks.[2] Three previous L2 studies use production: one written (Hertel 2003), one spoken (Leal, Destruel & Hoot 2019), and one focused on intonation (Kim 2016). Yet no previous study has investigated the L2 processing of Spanish focus. We address this limitation with a method (self-paced reading) that has been successfully employed to investigate similar phenomena in other languages (e.g., Hopp 2009; Kaiser & Trueswell 2004; Slioussar 2011; Weskott et al. 2011).

[2] For instance, some researchers have argued that, unlike other tasks, AJTs require metalinguistic awareness, which has been interpreted as a threat to external validity (see Schütze & Sprouse, 2013 for discussion). In our case, we triangulated our methods to alleviate some of these concerns, but we were interested to use AJTs partly to replicate previous findings but also because these have been argued to be especially helpful to unveil contrasts with relatively small sample sizes (see Marty et al. 2020).

1.6. *Research questions*

Our study was designed to address the following questions:

- Which word order do L2-Spanish learners and native speakers *prefer* in subject and object focus contexts?
- Which word order do L2-Spanish learners and native speakers *process faster* in subject and object focus contexts?
- What is the *role of L2 proficiency* when judging and processing Spanish subject and object focus?

To answer these questions, we carried out a judgment experiment and a processing experiment, described in the following sections in turn. We also measured learners' L2 proficiency.

2. Experiment 1: Forced-choice Task

2.1. *Participants*

Seventy-six L2-Spanish learners (L1-English; 62 female) completed this task in the Midwest region of the United States. Participants completed the Bilingual Language Profile (BLP; Birdsong, Gertken & Amengual 2012) and the LexTALE-ESP vocabulary size assessment as a measure of proficiency (Izura, Cuetos & Brysbaert 2014). Those reporting using Spanish more than 10% of the time with their family were excluded, as were those detailing significant contact with languages other than Spanish or English before age 12 (Table 4). Some participants report beginning to learn Spanish in elementary school, which may mean that the L2 group is composed of both earlier and later sequential bilinguals, but they were crucially not raised in homes where Spanish was spoken.

Table 4. Self-reported participant characteristics, L2 group.

	Mean	Range	SD
Age (years)	**20.4**	18 – 47	3.8
LexTALE Score[3]	**5.4**	-16 – 36	8.4
BLP Dominance Score (Higher = English-dominant)	**126.3**	57.7 – 178.1	23.5
Started learning Spanish (age)	**11.7**	5 – 19	3.0
Years classes (English)	**14.8**	4 – 19	2.3
Years classes (Spanish)	**6.6**	0 – 19	3.7
Weekly English use (family, friends, school/work)[4]	**93%**	30 – 100%	11.5
Weekly Spanish use (family, friends, school/work)	**8%**	0 – 70%	11.4
Weekly other language use (family, friends, school/work)	**0%**	0 – 30%	3.0
English proficiency, speaking (1 = low, 6 = high)	**5.9**	5 – 6	0.2
English proficiency, listening (1 = low, 6 = high)	**6.0**	5 – 6	0.1
English proficiency, reading (1 = low, 6 = high)	**5.9**	5 – 6	0.3
English proficiency, writing (1 = low, 6 = high)	**5.9**	4 – 6	0.2
Spanish proficiency, speaking (1 = low, 6 = high)	**3.2**	1 – 5	1.0
Spanish proficiency, listening (1 = low, 6 = high)	**3.8**	1 – 6	1.0
Spanish proficiency, reading (1 = low, 6 = high)	**3.9**	2 – 6	1.1
Spanish proficiency, writing (1 = low, 6 = high)	**3.4**	1 – 6	1.1

To obtain a baseline of behavior, we compare our experimental group with a control group of 42 monolingual native speakers of Yucatecan Spanish (25 female) tested in Merida, Mexico, whose results were previously reported (Hoot & Leal 2020; Hoot, Leal & Destruel 2020).[5] Control-group participants completed the BLP but not the proficiency test.[6] Although a

[3] For reference, the maximum score is 60. Izura et al.'s (2014) L1 Spanish group had a mean score of 53.9 and a range of 34 – 60, while their L2 Spanish (L1 English) group had a mean score of 11.9 and a range of -16 – 58.

[4] One participant did not answer all the usage questions. Their answers are included where we had them.

[5] The reader is referred to those publications for details.

[6] Because the present work is part of a larger research program including Yucatec Maya bilinguals, the BLP weighed their dominance in Spanish/Yucatec Maya (not Spanish/English).

few reported exposure to Yucatec Maya, they were all functionally monolingual. Therefore, we report their BLP responses about Spanish (Table 5), but not their dominance score. Participants who reported significant exposure to languages other than Spanish or Yucatec Maya before age 12 were excluded, as were participants who had not acquired Spanish monolingually before age 6, or speakers raised outside Yucatan.

Table 5. Self-reported participant characteristics, control group

	Mean	Range	SD
Age (years)	**21.8**	18 – 39	3.9
Years classes (Spanish)	**14.7**	3 – 18	2.6
Weekly Spanish use (family, friends, school/work)	**95%**	60 – 100%	8.1
Weekly other language use (family, friends, school/work)	**7%**	0 – 100%[7]	16.5

2.2. *Materials*

The forced-choice task required that participants choose the best response to an information-seeking *wh*-question (subject/object) in a context provided by an image. Figure 1 and Figure 2 show sample items.

2.2.1. Subject focus

Because we expected both subject position and the presence/absence of the object to affect subject focus realization (see §2.2), subject focus stimuli included **Word Order** (focus-final

[7] Some participants reported usage percentages that summed to more than 100%, including one who answered 100% to all the languages. Having talked with them about their language usage, however, we are reasonably confident that they were functionally monolingual Spanish speakers.

VOS/non-final VSO) and **Clitic presence** (full-DP object/clitic) as factors (Table 6).[8] Figure 1 shows a sample trial.

Table 6. Forced-choice task design: Subject focus

	Full DP Object	**Clitic Pronoun Object**
Subject-Final	VOS *Rompió el juguete el obrero.* broke the toy the worker 'The worker broke the toy.'	Cl_{itic}VS *Lo rompió el obrero.* it.ACC broke the worker 'The worker broke it.'
Subject-Non-Final	VSO *Rompió el obrero el juguete.* broke the worker the toy	SCl_{itic}V *El obrero lo rompió.* the painter it.ACC broke

Figure 1. Sample forced-choice item: Subject focus

[8] Because we chose to include the Clitic condition, we were not able to also test these stimuli with object focus questions (as we do in Experiment 2) because the experiment was already quite long. We acknowledge this limitation in our design.

We created sixteen lexicalizations distributed across two lists such that participants saw either the full-DP version or the clitic version of any given item. Participants saw sixteen total trials (8 DP and 8 with clitics).[9] Full-DP objects did not include canonical orders: choices were between focus-final VOS and non-final VSO. Spanish VSO can express sentence-wide focus (out-of-the-blue contexts) as well as object focus, while, critically, being incompatible with subject focus (Domínguez 2004:74; Zubizarreta 1998:125). This design avoids the confound between non-final focus and canonical word order, which can affect judgments independent of information structure concerns.

For the stimuli with clitics, the only grammatical options were SV or VS; it was impossible to avoid canonical orders. Even with this caveat, we expected increased acceptance of VS when objects were cliticized, following Gupton (2017).

Stimuli included words among the 5,000 most common (Davies 2006). Subjects/objects were controlled for number of syllables to avoid phonological-weight confounds. Subjects and objects were always definite; subjects were always human and objects always inanimate. Verbs were core transitive verbs in the preterit (e.g., *compró* 'bought'). The full set of stimuli is available via OSF (https://osf.io/f6u4c/).

2.2.2. Object focus

This task included three word-order options (instead of two): object-final focus (VPPO), PP-final

[9] A technical error caused one lexicalization to appear on both lists with the object replaced by a pronoun. Thus, no participants saw the full-DP version of that lexicalization, and half the participants in fact saw 7 full DP trials and 9 clitic pronoun trials. This error never resulted in a given participant seeing any lexicalization twice.

focus (VOPP), and fronting (OVPP/PPVO) (Table 7). Figure 2 shows a sample trial.

Table 7. Forced-choice task design: Object focus

	Object focus context	**PP focus context**
	¿Qué plantó en el jardín? what planted in the garden	*¿Dónde plantó los árboles?* where planted the trees
PP-Final VOPP	*Plantó los árboles en el jardín.* planted the trees in the garden 'He planted the trees in the garden.'	*Plantó los árboles en el jardín.* planted the trees in the garden
Object-Final VPPO	*Plantó en el jardín los árboles.*	*Plantó en el jardín los árboles.*
Fronting OVPP / PPVO	*Los árboles plantó en el jardín.*	*En el jardín plantó los árboles.*

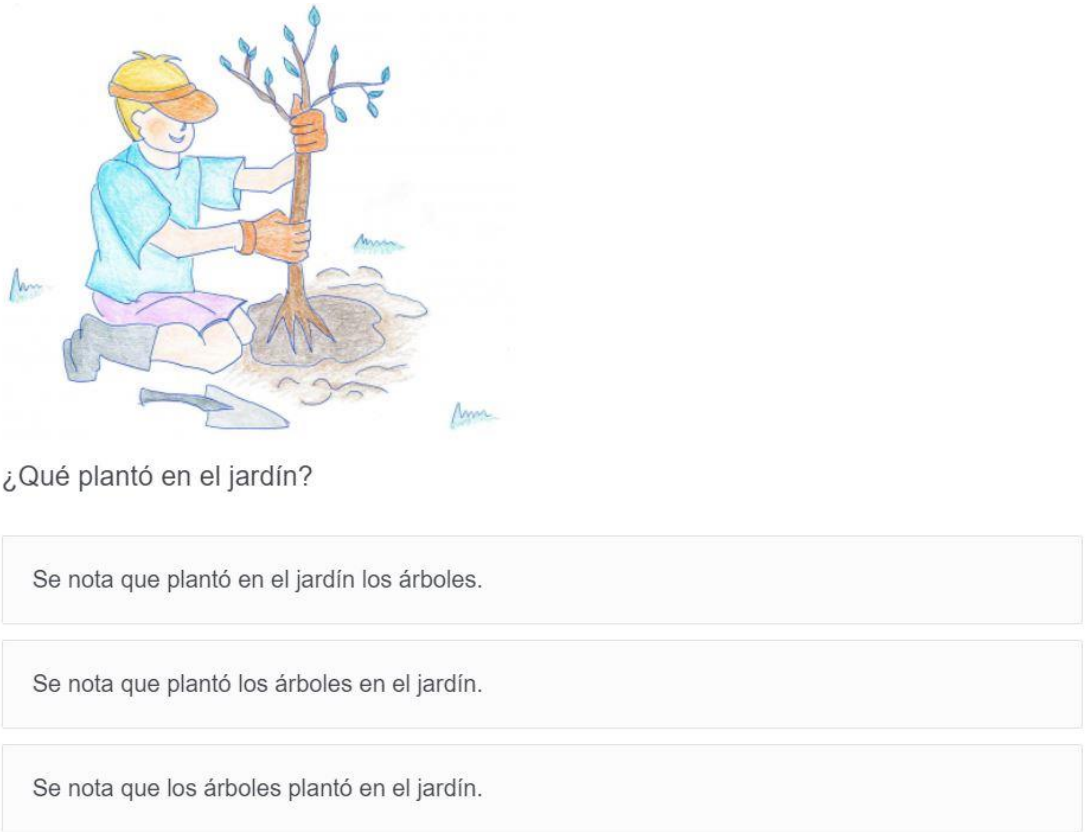

Figure 2. Sample forced-choice item: Object focus

Because canonical orders could not be avoided, canonical VOPP could be preferred *independent* of focus marking. However, comparing preferences for the two word orders across two contexts in a crossed design is ideal because if we see that VPPO becomes *relatively* better under object focus, despite the canonicity preference, that constitutes evidence that object-focus contexts favor object-final word orders (see Schütze and Sprouse 2013 for argumentation favoring crossed designs).

The third option—fronting—was included because this study was part of a project investigating cross-linguistic influence wherein one group was bilingual in Yucatec Maya, which marks focus via fronting (Verhoeven & Skopeteas 2015). While we do not consider Fronting in detail, fronting focal material is available in both Spanish and English as an instantiation of contrastive, rather than information, focus (López 2009). The fronted option was thus expected to be infelicitous in both focus contexts for all speakers.

We created sixteen lexicalizations distributed across two lists, so that each participant saw eight trials in each focus context. Stimuli were controlled for frequency: objects and PPs were syllable-matched for any given trial to avoid phonological-weight effects. Objects were plural, definite, and inanimate. PPs were temporal or locative adjuncts. Subjects, always singular and human, were overt in the context and null in the answer (the most natural option in Spanish). Because subjects were singular, agreement on the verb avoided possible confusion of the (plural) object for the subject.

2.3. *Procedure*

Participants generally completed both the self-paced reading task and then the forced-choice task in one session (with a few exceptions for scheduling conflicts). Testing was completed face-to-face, under supervision, either in a laboratory (L2ers) or in public places (NSs).

We administered the forced-choice task via Qualtrics. The instructions included two practice items with no feedback. The L2 group judged 48 trials: 16 subject-focus, 16 object-focus, and 16 fillers. As is common in many experimental designs, the items from each experiment served as fillers for the other, because the subject-focus and object-focus items were distinct in format (different questions, different number and type of options, different constituents present in the sentences), thus serving to obscure the purpose of each part. The additional fillers followed the same format but tested an unrelated grammatical structure (i.e., the use of preterit vs. present perfect for past tense reference). Trials and answers were randomized by participant, and participants were randomly assigned to one of two lists. Completing this task took between 10 and 20 minutes.

The control group judged 64 trials. In addition to the materials described above, they judged 16 additional trials testing subject and object fronting as part of a larger project, as mentioned above. We do not report those results here.

2.4. *Results*

2.4.1. Data analysis

Each focus type was analyzed separately. Because the dependent variable was categorical, we used logistic regressions via the GENLINMIXED procedure for generalized linear mixed-effects models (GLMM) in SPSS. For **Subject** focus, we analyzed the two **sentence Types** (full-DP object/clitic pronoun) separately because we did not have a fully crossed (factorial) design. Therefore, these models had only one fixed factor: **Group** (L2/native). The outcome was binary (VOS or not), thus analyzed with binomial logistic regressions. For **Object** focus, fixed factors were **focus Type** (object/PP focus), **Group** (L2/native), and their interaction. Because this test

had three outcome levels (VOPP, VPPO, and Fronting), we analyzed it with a multinomial logistic regression and then performed follow-up binomial tests to explore specific effects. We also conducted a separate analysis per focus type including proficiency as a continuous factor to examine development. Fixed factors were effect-coded (-0.5, +0.5).

For the two subject focus models, we used the maximal random effects structure (RES), which includes only random intercepts by participant and item because these tests had no within-participant fixed effects. For the object focus tests, models with the full RES with random by-participant and by-item slopes over Type did not converge; we reduced the models successively by eliminating first the by-item random slope, then the by-participant slope, and then, in one case, the by-item random intercept, until finding the maximal model that converged. In both cases, "by-item" random effects correspond to the different lexicalizations or token sets (viz., a particular set of words used to instantiate the linguistic structures across conditions), rather than each individual token sentence, given that each individual token sentence appears in only one level of each fixed factor. The RES for each model is presented in its output table.

As Meteyard and Davies (2020) point out in the context of linear mixed-effects models, GLMMs are flexible analysis tools that can be conceptualized as a regression or as similar to an ANOVA (or a nonparametric equivalent); the choice of which output to report depends on the aims of the experiment. Because we are fundamentally interested in determining whether the groups and/or conditions differ, we chose to report the *F* statistic produced by the Type III test of fixed effects, along with its associated *p* value, which best represents how we interpreted the statistical tests. For the subject focus results, which had binary outcomes, the *p* value associated with the *F* statistic is the same as the one associated with the *t* statistic produced for that fixed effect. For the object focus results with three outcomes, though, the omnibus *F* test for each fixed

effect is not the same as the simple comparisons made by the *t*-tests. Given our aims, we found it more sensible to report the omnibus test and then explore that result using follow-up tests. We also adopt Meteyard and Davies's (2020) suggestion of reporting the full results in tables.

2.4.2. Subject focus results

For subject focus, we first tested cases with the full-DP objects. We found no significant difference by Group (Table 8). Mean preferences, comparing between VSO and focus-final VOS, are shown in Figure 3.

Table 8. Binomial logistic regression, subject focus, full-DP object (target outcome = VOS)

Fixed Effect	**Coefficient**	**SE**	**95% CI**	**Odds Ratio**	**95% CI of Odds Ratio**	***F***	***p***
Group	0.396	0.253	-0.107 – 0.899	1.486	0.899 – 2.457	2.442	.121
Random Effects				**Variance**		**SE**	
By-Participant Intercept				1.084		0.245	
By-Item Intercept				0.067		0.058	

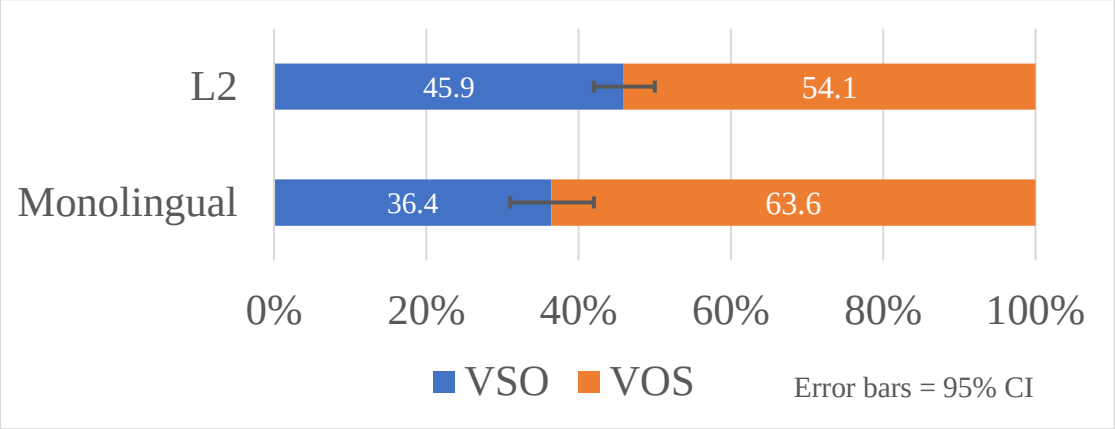

Figure 3. Preferences by group, subject focus, full-DP object

For the clitic object condition, on the other hand, we found a significant difference by Group (Table 9). The odds of the monolingual group choosing focus-final order are 2.9 times

larger than those of the L2 group. Mean preferences between canonical SclV and focus-final clVS are presented in Figure 4.

Table 9. Binomial logistic regression, subject focus, clitic object (target outcome = clVS)

Fixed Effect	**Coefficient**	**SE**	**95% CI**	**Odds Ratio**	**95% CI of Odds Ratio**	***F***	***p***
Group	1.057	0.299	0.464 – 1.650	2.879	1.591 – 5.209	12.515	0.001
Random Effects				**Variance**		**SE**	
By-Participant Intercept				1.650		0.338	
By-Item Intercept				0.031		0.051	

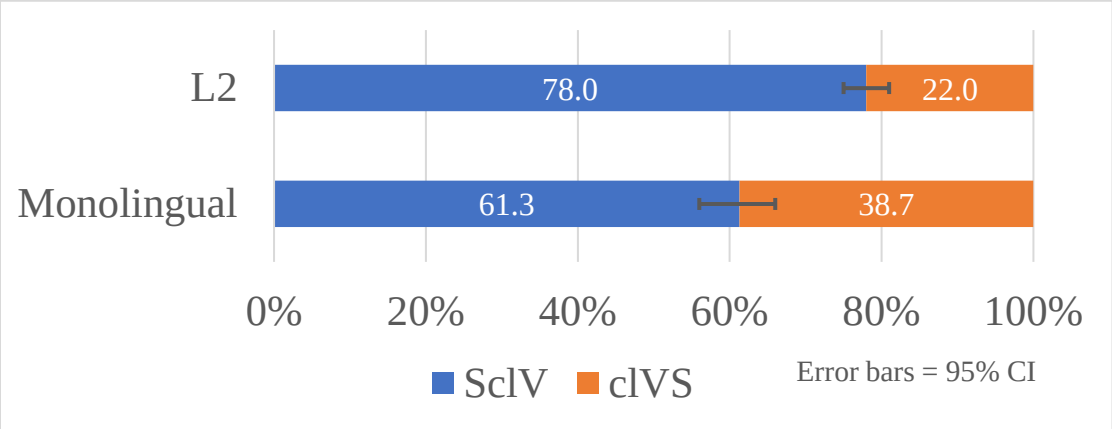

Figure 4. Preferences by group, subject focus, clitic object

To investigate proficiency effects, two follow-up binomial logistic regressions were conducted on the L2 data with mean-centered Proficiency as a continuous covariate. Neither test revealed an effect of Proficiency (Table 10). These results indicate that L2 participants did not vary in their preferences as a function of their proficiency.

Table 10. Binomial logistic regressions, subject focus, by proficiency (target outcome = focus final)

Condition	**Fixed Effect**	**Coefficient**	**SE**	**95% CI**	**Odds Ratio**	**95% CI of Odds Ratio**	***F***	***p***
Full-DP object	Proficiency	0.010	0.022	-0.033 – 0.053	1.010	0.968 – 1.054	0.217	.643
	Random Effects				**Variance**		**SE**	

	By-Participant Intercept				1.700		0.432	
	By-Item Intercept				0.079		0.085	
Condition	**Fixed Effect**	**Coefficient**	**SE**	**95% CI**	**Odds Ratio**	**95% CI of Odds Ratio**	***F***	***p***
Clitic object	Proficiency	-0.010	0.027	-0.063 – 0.043	0.990	0.939 – 1.044	0.134	.715
	Random Effects				**Variance**		**SE**	
	By-Participant Intercept				2.670		0.643	
	By-Item Intercept				0.109		0.111	

2.4.3. Object focus results

We report mean percentages for each outcome in the object focus condition in Figure 5 and the PP focus condition in Figure 6.

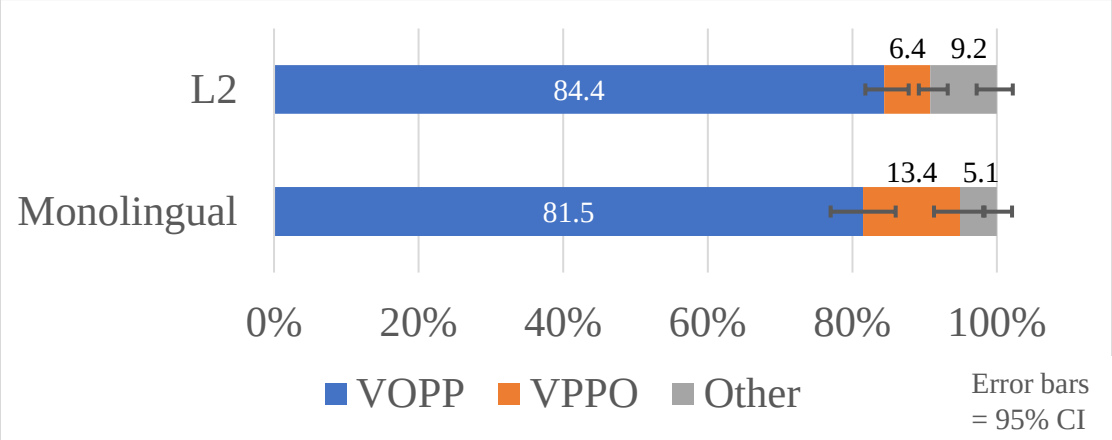

Figure 5. Preferences by group, object focus

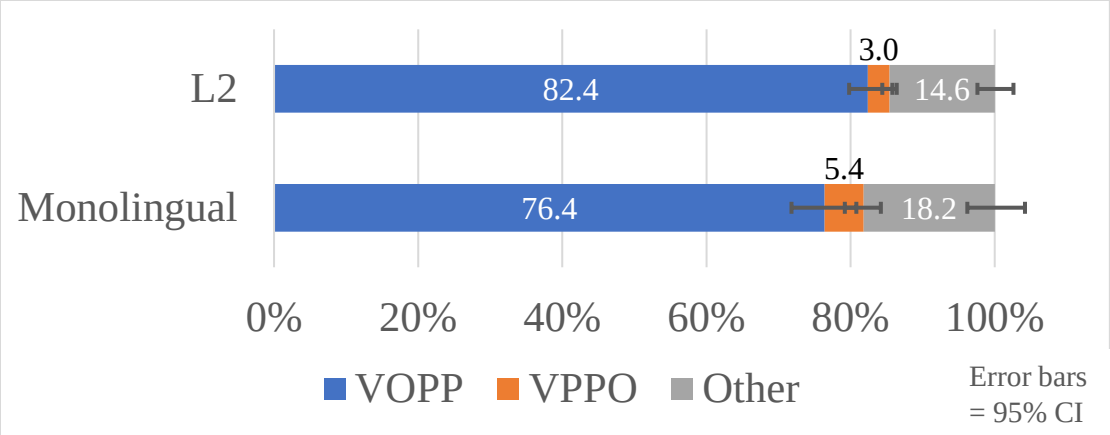

Figure 6. Preferences by group, PP focus

The overall results of the multinomial logistic regression (Table 11) tell us whether the distribution of the likelihoods across the three possible answers differs according to our independent variables. They suggest that the outcome distributions vary by group, that the outcomes differ by type, and, most importantly, that the groups responded to each focus type differently.

Table 11. Multinomial logistic regression, object focus[10]

Fixed Effects	***F***	***p***
Group	4.344	.015
Type	21.430	<.001
Group*Type	3.357	.035
Random Effects	**Variance**	**SE**
By-Participant Intercept	0.648	0.171

The omnibus test does not show us how the answers' distributions differ, though, so, to explore these results, we conducted a series of binomial logistic regressions. First, we compared

[10] As discussed in §3.4.1, we first report only the result of the omnibus test, without the regression coefficients, which we find less clear to interpret and clutter the reporting.

only canonical (VOPP) against non-canonical orders (VPPO/Fronting). The test revealed no significant differences by Group or Type, nor a significant interaction between them (Table 12).

Table 12. Binomial logistic regression, object focus, VOPP vs VPPO/Fronting

Fixed Effect	**Coefficient**	**SE**	**95% CI**	**Odds Ratio**	**95% CI of Odds Ratio**	***F***	***p***
Group	-0.288	0.210	-0.705 – 0.129	0.750	0.494 – 1.138	1.873	.174
Type	0.231	0.165	-0.113 – 0.575	1.260	0.893 – 1.778	1.963	.177
Group*Type	0.268	0.279	-0.285 – 0.821	1.307	0.752 – 2.273	0.922	.339
Random Effects				**Variance**		**SE**	
By-Participant Intercept				0.705		0.161	
By-Participant Slope over Type				0.243		0.256	
By-Item Intercept				0.209		0.099	
By-Item Slope over Type				0.123		0.130	

Second, we attempted to compare VPPO only against the other two orders (canonical VOPP and Fronting) collapsed together. However, no model, regardless of random effects structure, converged. We suspect that this is because the small number of VPPO cases left insufficient variation to model.

Third, we compared only the cases of Fronting against the other two orders, revealing that the groups did not differ overall in their use of Fronting, but the amount of Fronting differed significantly by Type, and moreover that there was an interaction between Group and Type (Table 13).

Table 13. Binomial logistic regression, object focus, Fronting vs VPPO/VPPO

Fixed Effect	**Coefficient**	**SE**	**95% CI**	**Odds Ratio**	**95% CI of Odds Ratio**	***F***	***p***
Group	-0.198	0.238	-0.668 – 0.273	0.821	0.512 – 1.314	0.687	.408
Type	-1.058	0.260	-1.594 – -0.522	0.347	0.203 – 0.593	16.597	< .001
Group*Type	-1.097	0.394	-1.875 – -0.319	0.334	0.153 – 0.727	7.773	.006
Random Effects				**Variance**		**SE**	
By-Participant Intercept				0.631		0.174	
By-Participant Slope over Type				0.665		0.410	
By-Item Intercept				0.322		0.156	

By-Item Slope over Type	0.429	0.295

Both groups were more likely to choose fronting under PP focus: taking the reciprocal of the odds ratio for the effect by Type, for ease of interpretation, we observe that the odds of choosing Fronting are 2.9 times (1/0.347) higher under PP focus. The interaction indicates that the effect of focus type varied by group. To explore the interaction, we conducted pairwise post hoc comparisons with the sequential Bonferroni correction for multiple comparisons and calculated odds ratios for them individually. The monolingual group's odds of choosing fronting under PP focus were nearly five times higher than under object focus (p = .001, odds ratio = 4.958), while the L2 group did not differ significantly by Type (p = .083, odds ratio = 1.650), suggesting that the main effect by Type was due mainly to the monolingual group's responses.

This finding suggests that the interaction observed both here and in the multinomial regression may be driven by differences in the use of Fronting, although that is not the outcome of most interest here. Given that fact, we removed Fronting from the data set as a follow up and compared only VOPP/VPPO in a final binomial test (Table 14), revealing a significant difference by Type and by Group, but no interaction, suggesting that the L2 group was more likely to choose VOPP orders than the control group (odds 2.1 times higher) and that both groups were more likely to prefer VOPP orders under PP focus than under object focus (odds 2.3 times higher).

Table 14. Binomial logistic regression, object focus, VOPP vs. VPPO (Fronting removed)

Fixed Effect	**Coefficient**	**SE**	**95% CI**	**Odds Ratio**	**95% CI of Odds Ratio**	***F***	***p***
Group	0.748	0.266	0.222 – 1.274	2.113	1.249 – 3.575	7.913	.006
Type	0.815	0.209	0.404 – 1.225	2.258	1.498 – 3.404	15.144	< .001
Group*Type	0.068	0.419	-0.753 – 0.889	1.070	0.471 – 2.432	0.026	.871

Random Effects	**Variance**	**SE**
By-Participant Intercept	0.671	0.217

Therefore, we conclude that the Type*Group interaction in our overall model is likely driven by a difference in the use of Fronting, not in the use of VOPP/VPPO. Taking a big-picture view, it is also worth noting that the two groups' patterns are largely the same overall.

Finally, to investigate the effects of proficiency, a second multinomial mixed-effects logistic regression was conducted on the L2 data with mean-centered Proficiency as a continuous covariate plus the same Type variable and their interaction, revealing significant differences by Type but no effect by Proficiency, and no interaction between them (Table 15). These results suggest that L2 participants did not vary in their preferences as a function of their proficiency; that is, the distribution of likelihoods among the three answers did not change as proficiency increased.

Table 15. Multinomial logistic regression, object focus, by proficiency

Fixed Effects	***F***	***p***
Proficiency	1.028	.361
Type	5.867	.003
Proficiency*Type	1.017	.362
Random Effects	**Variance**	**SE**
By-Participant Intercept	0.688	0.200
By-Item Intercept	0.683	0.323

2.5. *Interim discussion: Experiment 1*

The Forced-Choice Task reveals four main findings. First, we observed a preference for canonical word order whenever it was available. For object focus, both native speakers and learners displayed a strong preference for VOPP to realize both object and PP focus. For subject focus, where canonical order was available in the clitic object condition, we see a similar preference for canonical SclV order, although the L2 group showed a stronger preference than

the native group, even though their L1 (English) does not have clitics.

Second, we do observe preferences for focus-final realizations when canonical orders are removed. Namely, when comparing VSO against VOS (subject focus), participants exhibit a preference for focus-final VOS, and when comparing between object and PP focus, despite the overwhelming effect of canonical VOPP, we observe an increase in preference for VPPO under object focus, indicating an association between focus and final position. These results are important in the context of the focus literature in Spanish (see §1.2) because they show that although non-canonical focus-final orders can be used to mark focus, as suggested by the syntactic literature, canonical orders eclipse all other choices. This result aligns with previous work arguing that canonical orders can realize any information structure, while non-canonical orders are contextually restricted (Hoot & Leal 2020; Muntendam 2009). Because our language combination cannot allow us to determine whether this effect is driven by canonicity or by frequency (SVO in both languages is the canonical order *and* the most frequent), we leave this issue for further research. The main difference we observe between the groups is their relative canonical-order preference. While both groups preferred canonical SclV over non-canonical clVS, learners showed a stronger preference for canonical order (83%) than natives did (63%). Similarly, both groups preferred canonical VOPP over VPPO, but learners preferred it slightly more strongly (3% more for object focus and 6% more for PP focus).

Finally, in all other respects, L2 learners resemble the control group: their overall preference patterns are the same, and we find no effects by proficiency. Both these findings align with other recent judgment studies of Spanish focus marking (Domínguez & Arche 2008, 2014; Gupton 2017; Leal Méndez & Slabakova 2011) and suggest that the correlation between context and word order have not presented substantial acquisition difficulties for these L2 speakers.

3. Experiment 2: Self-paced reading

3.1. *Participants*

The same participants completed both tasks. However, because a technical error caused 34 L2 participants' SPR data to be lost, we report SPR data from 42 L2 participants, who are a subset of the 76 L2 speakers described in §3.1. Of the 42 participants, 33 were female. The L1 control group remained intact.

3.2. *Materials*

The self-paced reading task used a non-cumulative moving window design. Target sentences were preceded by non-moving contexts establishing focus (subject/object). Participants read target sentences segment by segment by pressing the spacebar. The critical region was embedded in a longer carrier phrase. Figure 7 shows a sample trial. All trials ended with a true/false comprehension question. Half these questions targeted the context and half the target sentence, but none targeted the item in focus. Half the questions were true and half false.

3.2.1. Subject focus

This experiment had a 2x2 factorial design (Table 16) with **Word Order** (VSO/VOS) and **Focus** (subject/object) as factors. As with the forced-choice task, we avoided canonical SVO order, instead choosing to contrast subject-final VOS with VSO because it can realize narrow object focus (Domínguez 2004:74; Zubizarreta 1998:125). Sample contexts are provided in (3) and (4); a representation of the procedure is in Figure 7.

Table 16. Self-paced reading task design: Subject focus

	Subject focus context **Who distracted the worker?**	**Object focus context** **Whom did the apprentice distract?**
Subject-Non-Final VSO	*Distrajo el aprendiz al*[11] *obrero.* distracted the apprentice ACC.the worker Prediction: Slower	*Distrajo el aprendiz al obrero.* distracted the apprentice ACC.the worker Prediction: Faster
Subject-Final VOS	*Distrajo al obrero el aprendiz.* distracted ACC.the worker the apprentice Prediction: Faster	*Distrajo al obrero el aprendiz.* distracted ACC.the worker the apprentice Prediction: Slower

(3) Subject Focus Context

Hubo un accidente en la fábrica porque alguien distrajo al obrero. ¿Sabes quién lo distrajo?

'There was an accident at the factory because someone distracted the worker. Do you know who distracted him?'

(4) Object Focus Context

Hubo un accidente en la fábrica porque el aprendiz distrajo a alguien. ¿Sabes a quién distrajo?

'There was an accident at the factory because the apprentice distracted someone. Do you know whom he distracted?'

[11] In Spanish, [+specific] and [+animate] objects are marked with the accusative marker *a* (Differential Object Marking). It is contracted to *al* before the masculine definite article *el* 'the.'

```
Hubo un accidente en la fábrica porque alguien distrajo al obrero.
¿Sabes quién lo distrajo?

---- -- ---- --- -------- -- -------- -- ------- ------ ----- ------------
Pues yo ---- --- -------- -- -------- -- ------- ------ ----- ------------
---- -- creo que -------- -- -------- -- ------- ------ ----- ------------
---- -- ---- --- distrajo -- -------- -- ------- ------ ----- ------------
---- -- ---- --- -------- el aprendiz -- ------- ------ ----- ------------
---- -- ---- --- -------- -- -------- al obrero, ------ ----- ------------
---- -- ---- --- -------- -- -------- -- ------- aunque ----- ------------
---- -- ---- --- -------- -- -------- -- ------- ------ puedo equivocarme.
```

Figure 7. Sample self-paced reading trial: Subject focus

Subjects and objects were masculine, animate nouns denoting professions or occupations (e.g., *worker*); verbs were core transitive verbs in the past tense. Subjects and objects were length-matched in two ways (they were always three syllables; always containing 25-32 characters in entire critical region). The five-word critical region was embedded in a carrier phrase because it has been argued that verb-initial orders are ungrammatical in Mexican Spanish (Gutiérrez-Bravo 2008). To avoid conflating the critical region with wrap-up effects, critical regions were followed by an adjunct (e.g., *aunque puedo equivocarme* 'although I might be wrong.')

In order to avoid surprisal reactions due to semantic plausibility (e.g., *The student taught the teacher* could produce a slowdown whereas *The teacher taught the student* would not), we conducted a norming test to ensure that the semantic roles of subjects and objects were reversible (i.e., neither the worker not the apprentice is *a priori* more likely to be distracting). In the norming test, native Spanish speakers (n = 26) rated the plausibility of 54 candidate lexicalizations with SVO word order (to make semantic roles maximally unambiguous) on a

scale from 0 (“totally impossible”) to 100 (“completely possible”). Two items were produced per lexicalization by swapping the subject for the object (5); items were distributed across two lists so that participants saw only one form.

(5) Norming Pre-test

a. Durante su viaje a Paris…

‘During his trip to Paris…’

b. …el artista visitó al poeta.

‘…the artist visited the poet.’

c. …el poeta visitó al artista.

‘…the poet visited the artist.’

Lexicalizations with mean ratings below the scale’s midpoint were discarded. From the remaining items, we selected the 32 items for which ratings were the closest to each other, maximizing reversibility so that participants’ RTs would reflect information-structural factors rather than semantic plausibility.

We conducted a pilot version of the self-paced reading task with native Spanish speakers (n = 9), who reviewed items for comprehensibility, naturalness, and word choice. Pilot participants reported the sentences were comprehensible and made only minor suggestions, which we adopted.

3.2.2. Object focus

The object-focus stimuli also had a 2x2 factorial design (Table 17), with **Word Order** (VOPP/VPPO) and **Focus** (object/PP) as factors. Sample contexts are provided in examples (6) and (7).

Table 17. Self-paced reading task design: Object focus

	Object focus context **What did he install?**	**PP focus context** **Where did he install it?**
VOPP	*Instaló el mosaico en el patio.* installed the mosaic in the patio Prediction: Slower	*Instaló el mosaico en el patio.* installed the mosaic in the patio Prediction: Faster
VPPO	*Instaló en el patio el mosaico.* Installed in the patio the mosaic Prediction: Faster	*Instaló en el patio el mosaico.* Installed in the patio the mosaic Prediction: Slower

(6) Object Focus Context

Contrataron a un artista para instalar una obra en el patio del museo. ¿Sabes qué instaló?

'They hired an artist to install a work in the patio of the museum. Do you know what he installed?'

(7) PP Focus Context

Contrataron a un artista para instalar un mosaico en alguna parte del museo. ¿Sabes dónde lo instaló?

'They hired an artist to install a mosaic in some part of the museum. Do you know where he installed it?'

Stimuli were controlled as follows. PPs and direct objects were matched for syllables, and the critical region overall was always 26-32 characters. Verbs were always three syllables and were core transitive verbs in third person singular past tense (e.g., *construyó* 'built'). Direct objects were always inanimate, indefinite, and masculine (e.g., *un ensayo* 'an essay'). The PP was either a temporal or locative adjunct (preposition + definite DP) (e.g., *por la noche* 'at night'). Like the subject-focus stimuli, these tokens were embedded in a carrier phrase and followed by an adjunct. To meet these requirements, it was not possible to control for frequency.

3.3. *Procedure*

The self-paced reading task was conducted in Linger (Rohde 2003). The experiment began with instructions in Spanish, followed by practice items. Stimuli were presented in three blocks with optional breaks. For each type of stimulus, we created 32 lexicalizations which we distributed so that each participant saw eight stimuli in each cell of the design. Each participant read 32 sentences from the subject-focus set and 32 from the object-focus set, along with 32 fillers. As is common in many experimental designs, the items from each experiment served as fillers for the other, because the subject-focus and object-focus items were distinct in format (different questions, different constituents present in the sentences, different lexical items, different contexts and context types), thus serving to obscure the purpose of the experiment. The additional fillers followed the same format but tested an unrelated grammatical structure (e.g., copula choices, *ser* vs. *estar*). Overall, participants read 96 context/stimulus dyads, which generally took 35-60 minutes. As mentioned in §3.3, participants generally completed the self-paced reading task first, then the forced-choice task, then the background questionnaire.

3.4. *Results*

3.4.1. Data analysis

Reading times (RTs) were trimmed at 100 and 10,000 milliseconds and then log-transformed because RT data is generally skewed. Two participants (both L2 speakers) whose overall log-transformed RTs (logRTs) were more than 2.5 standard deviations above the mean were excluded. We also used overall accuracy (70%) on the comprehension questions to exclude participants who were distracted during the task; eleven people were excluded for this reason (all but one were L2 speakers). These exclusions do not affect the participants numbers reported in

§4.1 (total number of participants was calculated after exclusions). To account for potential differences in word length and reading speeds, logRTs were length-adjusted using Fine et al.'s (2013) procedure. This procedure has previously been used on self-paced reading data (e.g., Kim 2018; Lee & Fraundorf 2019). We analyzed the length-adjusted logRTs for the three (aggregated) regions of interest and for the post-critical region.

Each focus type was analyzed separately. For each, we fit a linear mixed-effects model (LMM) via the MIXED procedure in SPSS with three fixed factors: **Focus context** (subject/object or object/PP), **Word Order** (VSO/VOS or VOPP/VPPO), and **Group** (monolingual/L2), plus all possible interactions. Fixed factors were effect-coded (-0.5, 0.5) to make the regression coefficients more easily interpretable. To investigate the role of proficiency, we ran follow-up tests on the L2 data, with Focus and Order as fixed factors plus addition of Proficiency as a continuous covariate, along with interactions.

To specify the random effects structure (RES), we followed the procedure used by Barr et al. (2013): we began with random intercepts and all possible random slopes by participant and by item, then, if that model was not supported by the data (i.e., did not converge), we simplified it by removing the random slope associated with the smallest variance, proceeding successively until arriving at the maximal RES supported by the data. As before, "by-item" random effects correspond to the different lexicalizations or token sets (viz., a particular set of words used to instantiate the linguistic structures across conditions), rather than each individual token sentence, given that each individual token sentence appears in only one level of each fixed factor. We report each model's RES alongside its output below.

LMMs are flexible tools that can be conceptualized as regressions or as tests similar to ANOVA, and they produce both regression coefficients with their corresponding *t*-tests and

omnibus *F* tests, both of which can be reasonable to report, depending on the aims of the experiment (Meteyard & Davies 2020). We chose to report the omnibus *F* tests produced by the Type III test of fixed effects because this method corresponds to how we understand the purpose of the experiment and because these omnibus tests can easily be followed with post hoc pairwise comparisons by condition to explore significant results. Nevertheless, because our fixed effects have only two levels and were effect-coded, the *p* values associated with the *F* test for a given fixed effect are identical to those associated with the *t*-test for the same effect, so this choice in reporting does not alter any outcomes. SPSS produces *p* values by using the Satterthwaite approximation to calculate denominator degrees of freedom to produce "an accurate F-test approximation, and hence accurate *p*-values for the F-test" (Ender, 2011:20), which more effectively avoids Type I error than some common alternatives for calculating *p* values in mixed-effects models (Luke 2017). As before, we follow Meteyard and Davies's (2020) advice and present each test's full set of outcomes is in a table.

Whenever relevant, we conducted pairwise post hoc comparisons with the Bonferroni correction, although we recognize that these comparisons are generally overly conservative (Larson-Hall 2010). To give a measure of effect size, we also calculated Cohen's *d* for these comparisons between means. Following advice from an anonymous reviewer, we calculated *d* directly from the data rather than from the estimated marginal means produced by the LMM. Cohen's *d* is a standardized measure of effect size that states the mean difference in standard deviation units, but there are several ways to calculate it (see Cumming 2012; Lakens 2013). For the sake of comparability, we chose to calculate the denominator using the pooled standard deviation, following what Avery and Marsden (2019) used in their meta-analysis of SPR effect

sizes.[12,13] Note that Avery and Marsden found that effect sizes for SPR are smaller than those observed for L2 research using other methods, presumably because the variability in RTs relative to the mean is inherently large for SPR. They found that L2 learners' sensitivity to morphosyntactic information during reading was d = .20, with a 95% CI of .15-.25, while for native speakers it was d = .28, with a 95% CI of .21-.35.

3.4.2. Subject focus results

Figure 8 plots the residual logRTs by sentence region for the VSO/VOS orders in subject-focus contexts (right panels) and object focus contexts (left panels), for native speakers (top) and L2 learners (bottom).

[12] We recognize that the pooled standard deviation, perhaps the most common way to calculate Cohen's *d*, assumes independent samples. Hence, Cumming (2012) recommends using the average standard deviation or an alternative calculation that considers the correlation between measures for repeated measures data, which is the type we employ. We calculated these as well, but given our equal sample sizes there is little difference between the calculations, so followed Avery and Marsden for ease of comparison.

[13] We also recognize that Cohen's *d* is biased at small sample sizes and the bias-corrected Hedges' *g* should generally be preferred (Lakens 2013). However, our samples were large enough that the effect sizes were always the same, so they could be reported as either one, Again, we follow Avery and Marsden for the sake of comparability.

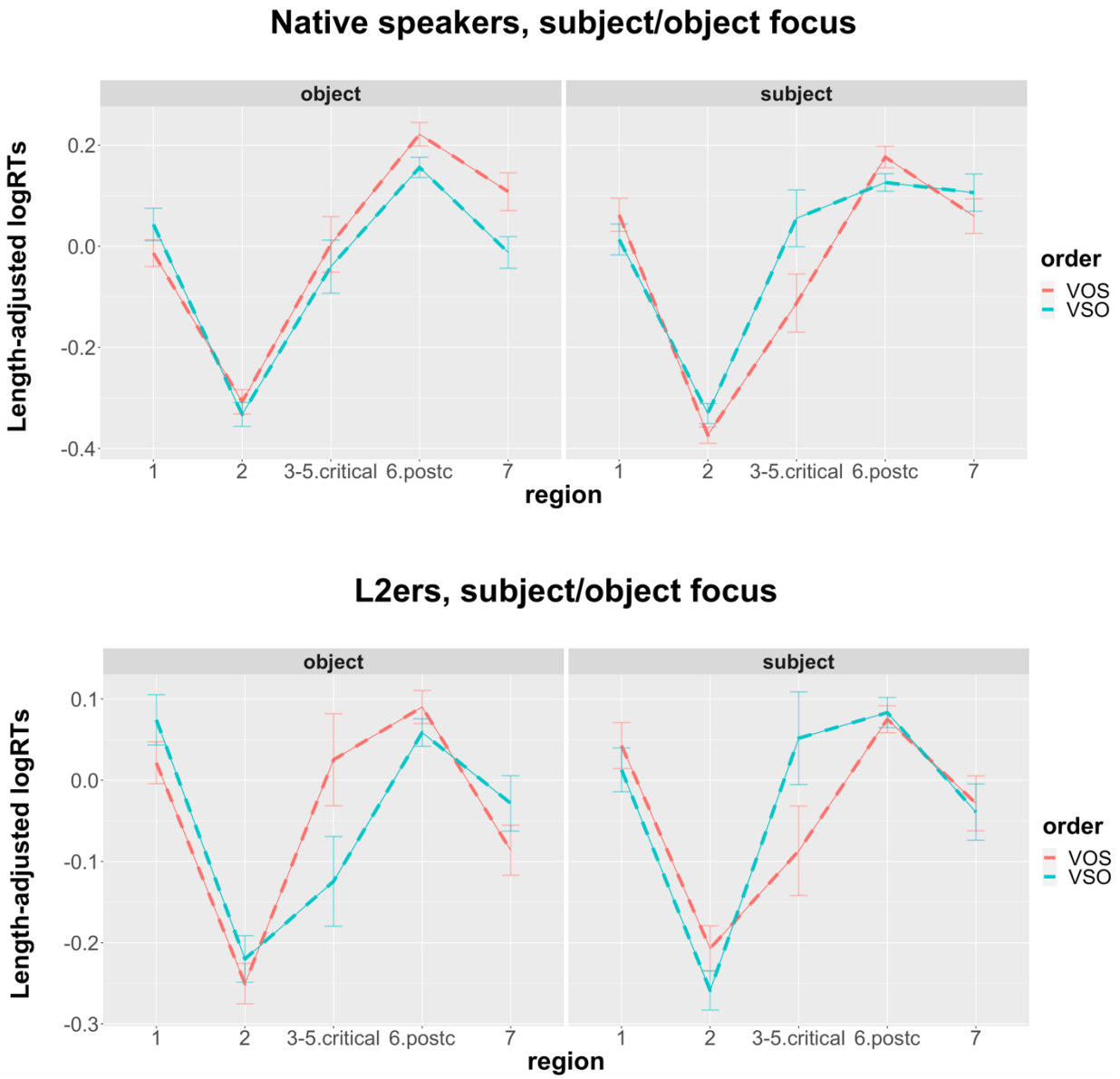

Figure 8. Length-adjusted log-transformed reading times in subject focus experiment by word order, focus context, and group

For the critical region (the aggregate of regions 3-5), the test of fixed effects revealed a significant Focus*Order interaction, but no other effects reached significance (Table 18). Crucially, we did not find evidence for differences by group.

Table 18. Linear mixed-effects model, subject focus, critical region

Fixed Effect	**Coefficient (β)**	**SE**	**95% CI**	***F***	***p***
Focus	0.010	0.040	-0.069 – 0.091	0.073	.787
Order	0.029	0.042	-0.058 – 0.115	0.470	.499
Focus*Order	0.250	0.077	0.098 – 0.402	10.510	.001
Group	0.010	0.053	-0.095 – 0.116	0.036	.849
Focus*Group	-0.042	0.080	-0.203 – 0.118	0.276	.601
Order*Group	0.067	0.080	-0.091 – 0.225	0.707	.403
Focus*Order*Group	-0.076	0.154	-0.379 – 0.226	0.245	.621
Random Effects		**Variance**		**SE**	
By-Participant Intercept		0.028		0.009	
By-Participant Slope over Focus		0.011		0.022	
By-Participant Slope over Order		0.007		0.021	
By-Item Intercept		0.008		0.005	
By-Item Slope over Order		0.005		0.014	

Bonferroni-corrected post-hoc pairwise comparisons conducted on the interaction revealed that VOS was read significantly faster in subject focus contexts than in object focus contexts (p = .041, d = .11) and VSO was read significantly faster in object focus contexts than subject focus contexts (p = .015, d = .13). Comparing within contexts, VOS was read significantly faster than VSO under subject focus (p = 0.008, d = .15), while under object focus the difference failed to reach significance but was trending toward VSO being read faster (p = 0.094, d = .10). Effect sizes in terms of Cohen's d (see §3.4.1) were slightly smaller than the effects observed for SPR studies by Avery and Marsden's (2019) meta-analysis.

We also examined region 6, immediately after the experiment manipulation, because some RT effects are known to spill over into the post-critical region (Marsden, Thompson & Plonsky 2018). We observe only a significant effect for group, such that the native speakers read this region faster overall (p < .001, d = .26), but we observe no differences in how the groups react to the context or word order (Table 2019).

Table 19. Linear mixed-effects model, subject focus, post-critical region

Fixed Effect	**Coefficient (β)**	**SE**	**95% CI**	***F***	***p***
Focus	-0.016	0.014	-0.044 – 0.012	1.353	.248
Order	-0.033	0.017	-0.067 – 0.001	3.813	.060
Focus*Order	0.027	0.027	-0.026 – 0.081	1.057	.307
Group	0.093	0.026	0.041 – 0.146	12.685	.001
Focus*Group	-0.042	0.028	-0.098 – 0.014	2.241	.138
Order*Group	-0.048	0.028	-0.104 – 0.007	2.966	.089
Focus*Order*Group	-0.024	0.053	-0.131 – 0.082	0.207	.650
Random Effects		**Variance**		**SE**	
By-Participant Intercept		0.011		0.002	
By-Participant Slope over Focus		0.003		0.003	
By-Participant Slope over Order		0.002		0.003	
By-Participant Slope over Focus*Order		0.004		0.010	
By-Item Intercept		0.002		0.001	
By-Item Slope over Order		0.003		0.002	

Finally, to examine the role of proficiency, we fit an LMM with only the L2 data for the critical region and proficiency as a continuous covariate, as described above. The test of fixed effects revealed a three-way interaction of Focus*Order*Proficiency (Table 20), suggesting that the way that L2 learners reacted to the four Focus*Order combinations (i.e., to contextual felicity) during online processing changed as proficiency increased. To visualize this interaction, Figure 9 plots how the difference in learners' logRTs between the felicitous and infelicitous word orders changes as proficiency increases. The positive slope observed in the figures indicates that the RT difference between the felicitous (focus-final) order and the infelicitous order in each context increases as proficiency increases, suggesting that more proficient learners make a greater distinction between orders by contextual felicity than less proficient learners.

Table 20. Linear mixed-effects model, subject focus, critical region, by proficiency

Fixed Effect	**Coefficient (β)**	**SE**	**95% CI**	***F***	***p***
Focus	-0.007	0.067	-0.142 – 0.128	0.010	.920
Order	-0.012	0.065	-0.147 – 0.123	0.032	.861
Focus*Order	0.191	0.117	-0.042 – 0.424	2.596	.107
Proficiency	0.004	0.005	-0.006 – 0.013	0.649	.425

Focus* Proficiency	-0.015	0.010	-0.034 – 0.004	2.439	.126
Order* Proficiency	-0.002	0.009	-0.020 – 0.016	0.068	.796
Focus*Order* Proficiency	-0.038	0.017	-0.071 – -0.004	4.849	.028
Random Effects		**Variance**		**SE**	
By-Participant Intercept		0.008		0.009	
By-Participant Slope over Focus		0.034		0.037	
By-Participant Slope over Order		0.013		0.032	
By-Item Intercept		0.011		0.009	
By-Item Slope over Order		0.012		0.028	

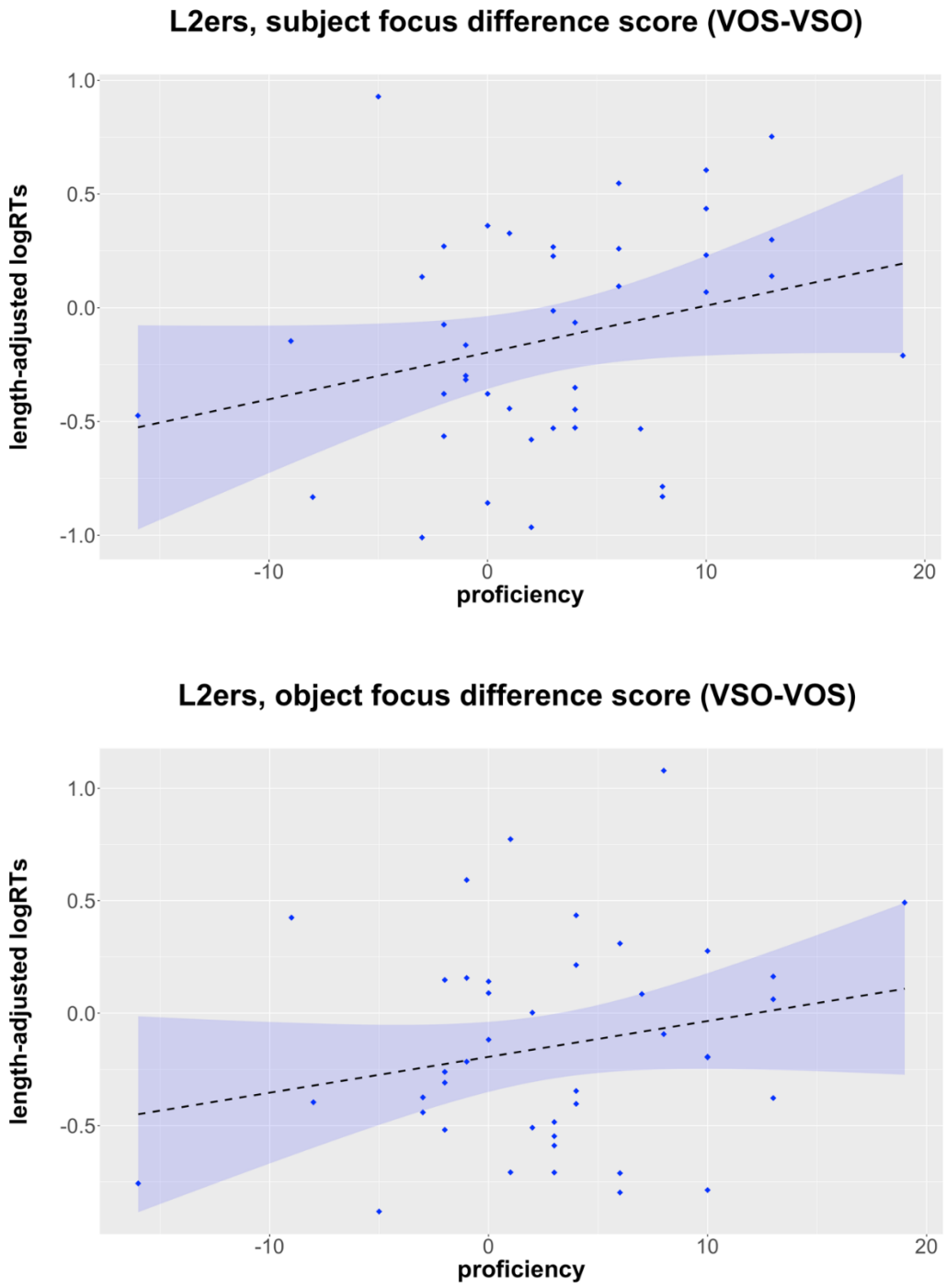

Figure 9. L2 learners' differences in length-adjusted log-transformed reading times between felicitous and infelicitous word orders in each context by proficiency

3.4.3. Object focus results

Figure 10 plots the logRTs by sentence region for the VOPP/VPPO orders in PP-focus contexts (right panels) and object focus contexts (left panels), for native speakers (top) and L2 learners (bottom).

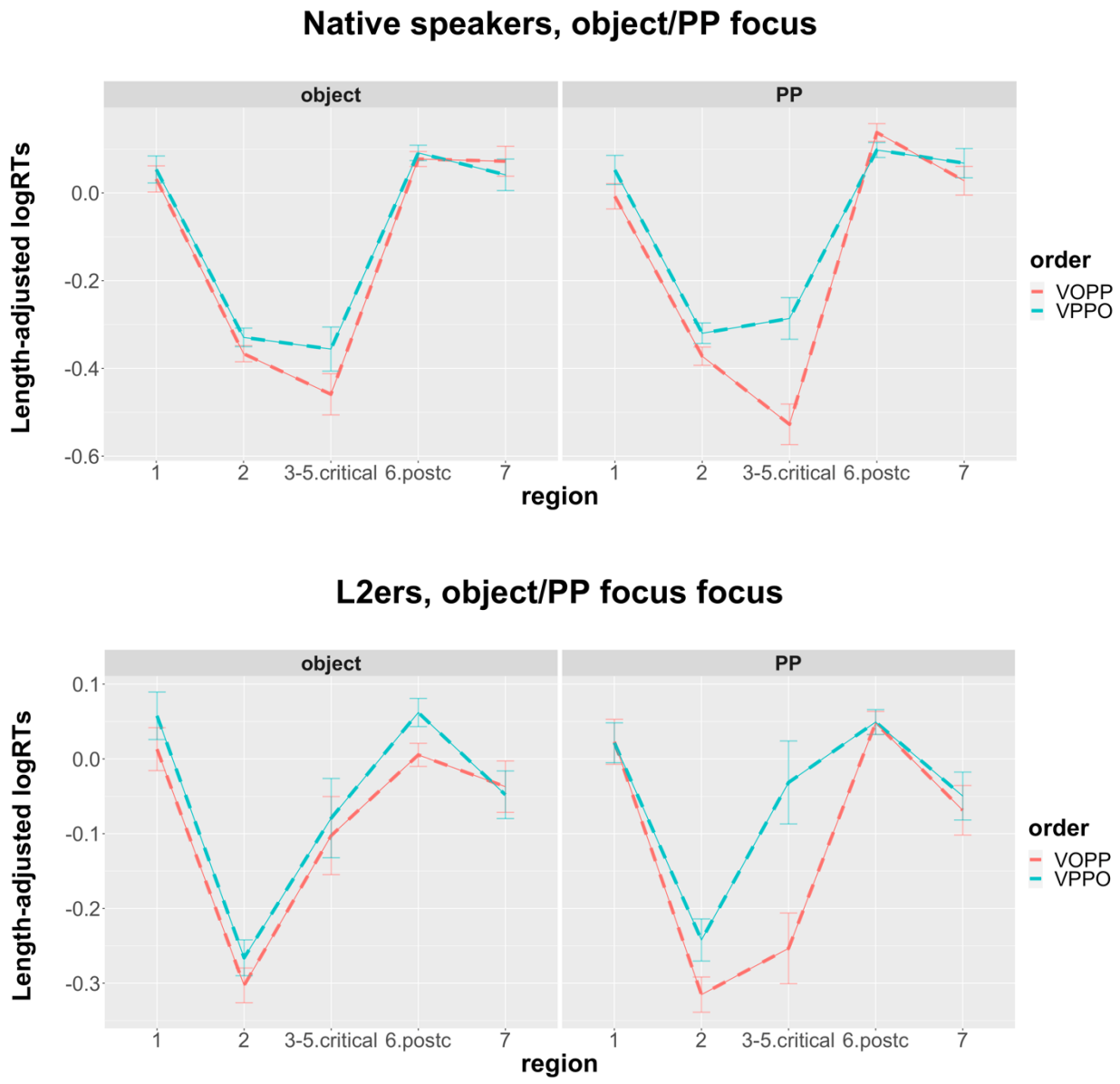

Figure 10. Length-adjusted log-transformed reading times in object focus experiment by word order, focus context, and group

For the critical region (the aggregate of regions 3-5), the test of fixed effects revealed significant main effects of Order and Group, plus a significant Focus*Order interaction; no other effects reached significance (Table 21). Crucially, we did not find evidence for differences by group.

Table 21. Linear mixed-effects model, object focus, critical region

Fixed Effect	**Coefficient (β)**	**SE**	**95% CI**	***F***	***p***
Focus	0.010	0.040	-0.069 – 0.091	0.073	.787
Order	0.029	0.042	-0.058 – 0.115	0.470	.499
Focus*Order	0.250	0.077	0.098 – 0.402	10.510	.001
Group	0.010	0.053	-0.095 – 0.116	0.036	.849
Focus*Group	-0.042	0.080	-0.203 – 0.118	0.276	.601
Order*Group	0.067	0.080	-0.091 – 0.225	0.707	.403
Focus*Order*Group	-0.076	0.154	-0.379 – 0.226	0.245	.621
Random Effects		**Variance**		**SE**	
By-Participant Intercept		0.028		0.009	
By-Participant Slope over Focus		0.011		0.022	
By-Participant Slope over Order		0.007		0.021	
By-Item Intercept		0.008		0.005	
By-Item Slope over Order		0.005		0.014	

Bonferroni-corrected post-hoc pairwise comparisons conducted on the interaction revealed that VOPP was read faster than VPPO under PP focus ($p < .001$, $d = .25$) and that VOPP was read faster in PP focus contexts than in object focus contexts ($p = .026$, $d = .12$). We observe no significant difference between word orders under object focus ($p = .258$, $d = .07$) nor between contexts for VPPO order ($p = .232$, $d = .06$). Effect sizes in terms of Cohen's *d* (see §3.4.1) for the two significant effects were similar to or slightly smaller than the effects observed for SPR studies by Avery and Marsden's (2019) meta-analysis.

We also examined region 6, immediately after the experiment manipulation, because some RT effects are known to spill over into the post-critical region (Marsden, Thompson & Plonsky 2018). We observe a significant effect for Focus and a significant effect for Group.

Native speakers read this region faster overall (p = .007, d = .19), and everyone read it slightly faster under object focus (p = .042, d = .08), but we observe no differences in how the groups react to the context or word order (Table 22).

Table 22. Linear mixed-effects model, object focus, post-critical region

Fixed Effect	**Coefficient (β)**	**SE**	**95% CI**	***F***	***p***
Focus	-0.025	0.012	-0.049 – -0.001	4.288	.042
Order	-0.008	0.014	-0.036 – 0.020	0.301	.585
Focus*Order	-0.055	0.029	-0.114 – 0.004	3.588	.069
Group	0.060	0.021	0.017 – 0.103	7.739	.007
Focus*Group	-0.018	0.024	-0.065 – 0.030	0.566	.454
Order*Group	0.041	0.028	-0.014 – 0.097	2.207	.141
Focus*Order*Group	0.002	0.050	-0.097 – 0.100	0.001	.974
Random Effects		**Variance**		**SE**	
By-Participant Intercept		0.007		0.002	
By-Participant Slope over Focus		0.001		0.002	
By-Participant Slope over Order		0.005		0.003	
By-Participant Slope over Focus*Order		0.006		0.008	
By-Item Intercept		0.001		0.001	
By-Item Slope over Focus*Order		0.007		0.006	

Finally, to examine the role of proficiency, we fit an LMM with only the L2 data for the critical region and proficiency as a continuous covariate, as described above. The test of fixed effects revealed a main effect of order, suggesting that L2 learners always read the canonical order faster, but no significant effects of proficiency, suggesting that RTs did not change as proficiency increased (Table 23).

Table 23. Linear mixed-effects model, object focus, critical region, by proficiency

Fixed Effect	**Coefficient (β)**	**SE**	**95% CI**	***F***	***p***
Focus	0.062	0.059	-0.060 – 0.185	1.116	.302
Order	-0.143	0.069	-0.284 – -0.003	4.290	.046
Focus*Order	0.182	0.118	-0.056 – 0.420	2.382	.131
Proficiency	-0.010	0.005	-0.021 – 0.001	3.502	.069
Focus* Proficiency	0.004	0.008	-0.012 – 0.020	0.274	.604
Order* Proficiency	-0.009	0.008	-0.024 – 0.006	1.423	.233
Focus*Order* Proficiency	-0.006	0.017	-0.041 – 0.028	0.143	.707
Random Effects		**Variance**		**SE**	
By-Participant Intercept		0.025		0.012	
By-Participant Slope over Focus		0.003		0.025	
By-Participant Slope over Focus*Order		0.080		0.114	
By-Item Intercept		0.020		0.010	
By-Participant Slope over Focus		0.016		0.026	
By-Item Slope over Order		0.060		0.036	

3.5. *Interim discussion: Experiment 2*

The self-paced reading results reveal three main conclusions. First, when canonical order was not available, both groups were faster when reading focus-final sentences than when reading sentences in which the final constituent did not correspond to the contextually appropriate focus. Specifically, VOS was read faster in subject-focus contexts and VSO in object-focus contexts, with no overall difference by word order and or in the pattern between groups.

Second, both focus types reveal the expected effects of context. For subject focus, when the canonical order is removed, non-canonical word orders are read faster in the appropriate context (i.e., VOS for subject focus and VSO for object focus). For object focus, canonical order could not be removed. Tellingly, canonical VOPP was read relatively fast in all contexts. In contrast, non-canonical VPPO differed by context: object-final orders were read more slowly outside the appropriate context. Nonetheless, object-final VPPO order was never read faster than canonical VOPP; at best, the object-focus context eliminated the reading time penalty for VPPO.

We take this to show that canonical VOPP is contextually appropriate in both contexts, while VPPO is restricted to object focus.

Finally, the L2 learners patterned with the native speakers. Although the native speakers, as is typical, read the post-critical regions faster *overall*, both groups display the same processing patterns in both the critical and post-critical regions. We thus conclude that these L2 learners can process the contextual felicity of non-canonical word orders in a nativelike way. The effect of proficiency on processing was split. For subject focus, learners' processing approximated nativelike patterns as proficiency increased, whereas for object focus we observed no proficiency effects. The lack of a developmental trend for object focus may suggest that nativelike processing patterns are achieved even at lower proficiency levels, rather than only by top scorers, but it may also be related to canonicity: when analyzed alone, the L2 learners read canonical VOPP faster overall, which may have obscured possible proficiency effects. Nevertheless, as we've pointed out, the L2 learners display the same processing patterns as the native speakers for both focus types, leading us to conclude that the contextual felicity of focus does not appear to present a processing challenge for these L2 learners.

6. General Discussion

Overall, our data is problematic for the Interface Hypothesis (IH) in its current articulation because we found no differences between L2ers and native speakers in terms of their overall preferences in judgments or, crucially, in processing information focus. In the Forced-Choice Task, our learners did show that they preferred canonical orders even more strongly than native speakers, who showed the same preference, yet, overall, our L2 learners could both judge focus acceptability in context and process focus in a nativelike manner. Crucially, the learners showed that they could *process* focus in an online task without the residual optionality predicted by the

IH. In this regard, our investigation aligns with recent results showing that information focus in L2 Spanish is not especially vulnerable to optionality (Leal, Destruel & Hoot 2019). This is akin to what other researchers have found for different structures at the interface that pertain to information structure, such as CLLD (e.g., Leal, Slabakova & Farmer 2017).

Nevertheless, the question of whether the syntax-pragmatics interface is especially problematic for end-state learners is still an open one, and the IH is ripe for further investigation. First, as noted by Jin and Ke (2020), only a limited number of language combinations and structures have been tested. Second, it seems to be the case that certain properties are more problematic than others—even within the same interface, raising the question of what additional factors play a role in the evinced difficulties (Leal, Destruel & Hoot 2019; Rothman & Slabakova 2011; White 2011). In our case, the L2 learners' pronounced preference for canonical SclV raises the possibility that canonicity or even phonological factors may play a role in the acquisition of focus.

Furthermore, because the IH, in essence, predicts *relative* differences (between internal and external interfaces), the question of what structures should be compared and how this comparison can come about are still open for debate (see Hopp 2009 for discussion). Finally, many of the studies testing the IH, including our previous work, use offline measures, which complicates the falsification of the hypothesis. In this regard, our work represents a useful contribution, given that self-paced reading data allows us to weigh in directly on the hypothesis's claims about processing. Because we did not find any qualitative differences in this task, our results cannot support the predictions of the IH.

One result in our data that merits further discussion is the lack of proficiency effects in all but one analysis, even though proficiency, as a factor, has frequently been associated with a host

of linguistic outcomes in L2 studies across methodologies and frameworks. In methodological terms, measuring and representing proficiency can be a complex endeavor given that learners can differ in terms of their competence, use, dominance, and exposure to the language. In L2 acquisition research, representing L2 competence is typically done via standardized or in-house tests. We are no exception. In our case, given the length of the testing materials, we chose a standardized test: the LexTALE_Esp (Izura, Cuetos & Brysbaert 2014), which is a short test in which learners must choose among real words (n=60) and non-words (n=30). As evident from Table 4, our sample included a fair amount of variability, so the lack of proficiency effects is likely not due to sampling issues. However, it should be noted that although the LexTALE has been shown to have agreement with other measures, such as interviews, self-rated proficiency, picture-naming tests, and multi-dimensional survey instruments like the BLP, (Bonvin, Brugger & Berthele 2021; de Bruin, Carreiras & Duñabeitia 2017), it is, after all, a measure of lexical knowledge. It is thus possible that because our measurement of proficiency comes from a single source, we could have missed proficiency effects that might have been captured using the multi-measure approach advocated by several L2 researchers (de Bruin, Carreiras & Duñabeitia 2017; Gollan et al. 2012).

Irrespective of considerations of the measurement of proficiency, however, our results show unambiguous evidence that focus is acquirable by L2 Spanish learners in the long run. This adds to existing challenges to the hypothesis (e.g., Slabakova 2011; White 2011) although, as we mentioned earlier, it has the advantage of providing experimental evidence from an online task. An open question in the research of the L2 acquisition of Spanish focus, however, is the point in L2 development where acquisition takes place. What our data suggests is that the acquisition of focus does not present unusual difficulty in L2 acquisition, even at the early stages of

acquisition. While further research might uncover more about the development paths in the acquisition of focus in L2 Spanish, our results align with previous L2 research using different methodologies in finding that L1-English/L2-Spanish learners can perform like native speakers with regard to focus (Gupton 2017; Kim 2016; Leal Méndez & Slabakova 2011; Leal, Destruel & Hoot 2019).

Another notion that merits discussion is related to canonicity, and indirectly, to frequency. Although we chose to avoid canonical orders wherever possible (see §3.2.1 and §4.2.1), our data makes clear that canonicity plays an important role in information focus marking and that, in some cases (i.e., when an object is cliticized), learners might show an especial proclivity for canonical orders. Previous research has shown that native Spanish speakers can realize different information structure notions using canonical orders, while non-canonical orders are reserved for specific discourse contexts (Hoot & Leal 2020; Muntendam 2009). The present study shows that this behavior is echoed in the L2 data. However, because SVO is the canonical and most frequent order in both English and in Spanish, puzzling out the contributions of canonicity and frequency is not possible within our design. Yet because we chose to avoid canonical orders in two of our conditions in both experiments, we can say that high frequency is not a *necessary* condition for the successful acquisition of discourse restrictions of infrequent word orders. Recall that the two V-initial orders we use in the study (VSO/VOS), while grammatical in all Spanish varieties, are quite rare. Ocampo (2009), for instance, analyzed a corpus of 20 hours of informal conversation and found only two instances of VSO and zero examples of VOS. To complicate matters further, only one of the VSO instances instantiated a focus-final realization—the second appears to be realizing focus on the subject,

instead.[14] Thus, although Slabakova (2015) has suggested that the difficulties predicted by the IH could be exacerbated by differences in L1-L2 relative frequency, we do not appear to find those effects in this data set.

The rarity of verb-initial orders does highlight the improbable feat that the learners in our study achieved: How are L2 learners able to acquire the association between final sentential position and focus marking given its infrequency? Perhaps the crosslinguistic association of focus with prominence (§1.3) can help guide acquisition. Although languages vary in how a constituent is made prominent, there are certain correlates that learners could identify, such as higher pitch or particular word orders. What is surprising in our case, is that both VSO and VOS are quite rare in the input, so this possibility seems unlikely. Similarly, learners could benefit from the basic word-order principles, such as the Verb-Object Bonding principle (Tomlin 1986: verbs tend to be more tightly associated with Objects than with Subjects), but this principle would only help in cases of subject focus (VOS), not object focus (VSO). In this context, we believe that the best explanation for the learner's behavior falls under the Full Transfer/Full Access Hypothesis, because Universal Grammar principles can best explain successful acquisition given that the input underdetermines the knowledge our learners demonstrated (see Schwartz & Sprouse 1996).

The last issue we will discuss here involves our choice of methods. Our investigation contributes to the documentation of the processing of information focus in Spanish. In the case of the IH, this data is vital because Sorace (2011) proposes that difficulties will arise during real-

[14] In his analysis, Ocampo (2009: 504) groups information focus with the notions of "*contraste y realce*" ('contrast and highlighting'), which could mean, in principle, that even this one occurrence of VSO instantiated contrastive, rather than information, focus.

time processing. While there are two previous investigations of information structure in L2 Spanish that test the IH by using time pressure in their design (Leal, Destruel & Hoot 2019; Sequeros-Valle, Hoot & Cabrelli 2020), ours is the first investigation of the acquisition of Spanish information focus that can index the speakers' reaction times more directly. In terms of the literature on Spanish focus, we have initial evidence that learners align with native Spanish speakers in their processing signatures of focus.

Investigating properties at the interfaces is, intrinsically, a complex endeavor, given the modules involved. In this regard, we must mention that one of the limitations of our research is that neither of our tasks included prosody in its design. While it is unlikely that our participants read our SPR items with a special intonation (given the self-paced nature of the task), the items in the FCT where not timed and could have been susceptible to such a confound. On the other hand, the piecemeal delivery of the SPR could also have unknown prosodic effects. Thus, although our data contributes important information to the investigation of focus, especially as it concerns its processing, more research with methodologies such as forced-choice listening tasks, cross-modal priming, eye-tracking, or ERPs might be able to explore this particular structure in more detail, by including prosodic considerations in the design.

7. Conclusion

Our study investigated the acquisition of narrow information focus in L2 Spanish. We tested the Interface Hypothesis (Sorace 2011), which predicts that structures that involve external modules, such as information focus, will result in optionality in L2 grammars at near-native level due to resource constraints on real-time processing. Our results, however, did not provide unequivocal evidence of optionality, either in our offline task (forced-choice) or in our online task (self-paced reading). Instead, L1-English/L2-Spanish learners largely performed like native speakers,

including on the task measuring processing. Given that our design avoided, when possible, focus realizations via canonical SVO, we note that frequency does not seem to complicate the acquisition of this particular structure. Although further research is needed to disentangle the effects of prosody, the role of canonicity, and the developmental path for acquiring discourse restrictions, our data shows that information focus does not appear to pose special difficulties for L2 learners, even at lower proficiency levels, a finding that is in line with other recent experimental studies (Domínguez & Arche 2008, 2014; Gupton 2017; Kim 2016; Leal Méndez & Slabakova 2011; Leal, Destruel & Hoot 2019).

Data availability statement

The data that support the findings of this study are openly available via OSF.io at https://osf.io/f6u4c/ (DOI: 10.17605/OSF.IO/F6U4C), along with the instruments used to collect the data.

References

Avery, Nick & Emma Marsden. 2019. A meta-analysis of sensitivity to grammatical information during self-paced reading: Towards a framework of reference for reading time effect sizes. *Studies in Second Language Acquisition* 41(5). 1055–1087. doi:10.1017/S0272263119000196.

Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278. doi:10.1016/j.jml.2012.11.001.

Bini, Milena. 1993. La adquisición del italiano: Más allá de las propiedades sintácticas del parámetro pro-drop en el español no nativo. In Juana M. Liceras (ed.), *La lingüística y el análisis de los sistemas no nativos*, 126–139. (Ottawa Hispanic Studies 12). Ottawa: Dovehouse Ed.

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

Bolinger, Dwight. 1954. Meaningful word order in Spanish. *Boletín de Filología* 8. 45–56.

Bonvin, Audrey, Ladina Brugger & Raphael Berthele. 2021. Lexical measures as a proxy for bilingual language dominance? *International Review of Applied Linguistics in Language Teaching*. De Gruyter Mouton. doi:10.1515/iral-2020-0093. https://www.degruyter.com/document/doi/10.1515/iral-2020-0093/html (29 July, 2021).

Bresnan, Joan, Anna Cueni, Tatiana Nikitina & R. Harald Baayen. 2007. Predicting the dative alternation. In Gerlof Bouma, Maria Aloni & Koninklijke Nederlandse Akademie van Wetenschappen (eds.), *Cognitive foundations of interpretation*, 69–94. Amsterdam: Royal Netherlands Academy of Arts and Sciences.

de Bruin, Angela, Manuel Carreiras & Jon Andoni Duñabeitia. 2017. The BEST dataset of language proficiency. *Frontiers in Psychology* 8. 522. doi:10.3389/fpsyg.2017.00522.

Büring, Daniel. 2009. Towards a typology of focus realization. In Malte Zimmermann & Caroline Féry (eds.), *Information Structure: Theoretical, Typological, and Experimental Perspectives*, 177–205. Oxford: Oxford University Press.

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

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Contreras, Heles. 1978. *El orden de palabras en español*. Madrid: Cátedra.

Costa, João. 2001. The emergence of unmarked word order. In Géraldine Legendre, Jane Barbara Grimshaw & Sten Vikner (eds.), *Optimality-Theoretic syntax*, 171–204. Cambridge, Mass.: MIT Press.

Cumming, Geoff. 2012. *Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. London: Taylor & Francis. http://ebookcentral.proquest.com/lib/depaul/detail.action?docID=957018.

Davies, Mark. 2006. *A Frequency dictionary of Spanish: core vocabulary for learners*. New York/London: Routledge.

Domínguez, Laura. 2004. The effects of phonological cues on the syntax of focus constructions in Spanish. In Reineke Bok-Bennema, Bart Hollebrandse, Brigitte Kampers-Manhe & Petra Sleeman (eds.), *Romance Languages and Linguistic Theory 2002: Selected Papers*

from "Going Romance," Groningen, 28-30 November 2002, 69–81. Amsterdam: John Benjamins.
Domínguez, Laura & María J. Arche. 2008. Optionality in L2 grammars: the acquisition of SV/VS contrast in Spanish. In Harvey Chan, Heather Jacob & Enkeleida Kapia (eds.), *Proceedings of the 32nd Annual Boston University Conference on Language Development*, 96–107. Somerville, Mass.: Cascadilla Proceedings Project.
Domínguez, Laura & María J. Arche. 2014. Subject inversion in non-native Spanish. *Lingua* 145. 243–265. doi:10.1016/j.lingua.2014.04.004.
Ender, Phil. 2011. xtmixed and denominator degrees of freedom: Myth or magic. Paper presented at the 2011 Stata Conference Chicago, Chicago. https://www.stata.com/meeting/chicago11/materials/chi11_ender.pdf.
Erteschik-Shir, Nomi. 2007. *Information structure: the syntax-discourse interface*. Oxford: Oxford University Press.
Escandell Vidal, Victoria & Manuel Leonetti. 2019. Una nota sobre el foco informativo en español. In Ramón Gonzáles Ruiz, Inés Olza Moreno & Óscar Loureda Lamas (eds.), *Lengua, cultura, discurso: Estudios ofrecidos al profesor Manuel Casado Velarde*, 207–224. Pamplona: EUNSA.
Fábregas, Antonio. 2016. Information structure and its syntactic manifestation in Spanish: facts and proposals. *Borealis – An International Journal of Hispanic Linguistics* 5(2). 1–109. doi:10.7557/1.5.2.3850.
Fine, Alex B., T. Florian Jaeger, Thomas A. Farmer & Ting Qian. 2013. Rapid expectation adaptation during syntactic comprehension. *PLOS ONE* 8(10). e77661. doi:10.1371/journal.pone.0077661.
Gollan, Tamar H., Gali H. Weissberger, Elin Runnqvist, Rosa I. Montoya & Cynthia M. Cera. 2012. Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingualism: Language and Cognition* 15(3). 594–615. doi:10.1017/S1366728911000332.
Gupton, Timothy. 2017. Early minority language acquirers of Spanish exhibit focus-related interface asymmetries: Word order alternation and optionality in Spanish-Catalan, Spanish-Galician, and Spanish-English bilinguals. In Fraser Lauchlan & María Carmen Parafita Couto (eds.), *Bilingualism and Minority Languages in Europe*, 212–239. Newcastle upon Tyne: Cambridge Scholars Publishing.
Gutiérrez-Bravo, Rodrigo. 2002. Structural markedness and syntactic structure: A study of word order and the left periphery in Mexican Spanish. Santa Cruz, Calif.: University of California - Santa Cruz Dissertation.
Gutiérrez-Bravo, Rodrigo. 2008. La identificación de los tópicos y los focos. *Nueva revista de filología hispánica* 56(2). 363–401.
Gutiérrez-Bravo, Rodrigo. 2020. La sintaxis del español de México: Un esbozo. *Cuadernos de la ALFAL* 12(2). 44–70.
Hertel, Tammy Jandrey. 2003. Lexical and discourse factors in the second language acquisition of Spanish word order. *Second Language Research* 19(4). 273–304. doi:10.1191/0267658303sr224oa.
Hoot, Bradley. 2016. Narrow presentational focus in Mexican Spanish: Experimental evidence. *Probus* 28(2). 335–365. doi:10.1515/probus-2014-0004.
Hoot, Bradley & Tania Leal. 2020. Processing subject focus across two Spanish varieties. *Probus* 32(1). 93–127. doi:10.1515/probus-2019-0004.

Hoot, Bradley, Tania Leal & Emilie Destruel. 2020. Object focus marking in Spanish: An investigation using three tasks. *Glossa: a journal of general linguistics* 5(1). 70. doi:10.5334/gjgl.1160.
Hopp, Holger. 2009. The syntax–discourse interface in near-native L2 acquisition: Off-line and on-line performance. *Bilingualism: Language and Cognition* 12(04). 463–483. doi:10.1017/S1366728909990253.
Ivanov, Ivan Prodanov. 2009. Second language acquisition of Bulgarian object clitics: A test case for the interface hypothesis. Iowa City: University of Iowa Dissertation.
Izura, Cristina, Fernando Cuetos & Marc Brysbaert. 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.
Jackendoff, Ray. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.
Jin, Jing & Sihui Ke. 2020. Second language acquisition at interfaces: A study of the word order variation in the Chinese nominal domain. *Journal of Psycholinguistic Research* Online ahead of print. doi:10.1007/s10936-020-09707-6.
Kaiser, Elsi & John C. Trueswell. 2004. The role of discourse context in the processing of a flexible word-order language. *Cognition* 94(2). 113–147. doi:10.1016/j.cognition.2004.01.002.
Kim, Hyunwoo. 2018. Second language processing of Korean floating numeral quantifiers. *Journal of Psycholinguistic Research* 47(5). 1101–1119. doi:10.1007/s10936-018-9581-8.
Kim, Ji Young. 2016. The perception and production of prominence in Spanish by heritage speakers and L2 learners. Urbana, Ill.: University of Illinois at Urbana-Champaign Dissertation.
Krifka, Manfred. 2007. Basic notions of information structure. *Interdisciplinary Studies on Information Structure* 6. 16–55.
Lakens, Daniël. 2013. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology* 4. 863. doi:10.3389/fpsyg.2013.00863.
Larson-Hall, Jenifer. 2010. *A guide to doing statistics in second language research using SPSS*. New York/London: Routledge.
Leal Méndez, Tania & Roumyana Slabakova. 2011. Pragmatic Consequences of P-movement and Focus Fronting in L2 Spanish: Unraveling the Syntax-Discourse Interface. In Julia Herschensohn & Darren Tanner (eds.), *Proceedings of the 11th Generative Approaches to Second Language Acquisition Conference (GASLA 2011)*, 63–75. Somerville, Mass.: Cascadilla Proceedings Project.
Leal, Tania. 2018. Data analysis and sampling. In Aarnes Gudmestad & Amanda Edmonds (eds.), *Critical Reflections on Data in Second Language Acquisition*, 63–88. Amsterdam: John Benjamins.
Leal, Tania, Emilie Destruel & Bradley Hoot. 2019. The acquisition of Focus in L2 Spanish. *Second Language Research* 35(4). 449–477. doi:10.1177/0267658318784343.
Leal, Tania, Roumyana Slabakova & Thomas A. Farmer. 2017. The fine-tuning of linguistic expectations over the course of L2 learning. *Studies in Second Language Acquisition* 39(3). 493–525. doi:10.1017/S0272263116000164.

Lee, Eun-Kyung & Scott Fraundorf. 2019. Native-like processing of prominence cues in L2 written discourse comprehension: Evidence from font emphasis. *Applied Psycholinguistics* 40(2). 373–398. doi:10.1017/S0142716418000619.
Leonetti, Manuel. 2014. Gramática y pragmática en el orden de palabras. *Linred: Lingüística en la Red* 12. 1–25.
López, Luis. 2009. *A derivational syntax for information structure*. Oxford: Oxford University Press.
Lozano, Cristóbal. 2006a. The development of the syntax-information structure interface: Greek learners of Spanish. In Vincent Torrens & Linda Escobar (eds.), *The acquisition of syntax in Romance languages*, 371–399. Amsterdam: John Benjamins.
Lozano, Cristóbal. 2006b. Focus and split-intransitivity: The acquisition of word order alternations in non-native Spanish. *Second Language Research* 22(2). 1–43.
Luke, Steven G. 2017. Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods* 49(4). 1494–1502. doi:10.3758/s13428-016-0809-y.
Margaza, Panagiota & Aurora Bel. 2006. Null subjects at the syntax-pragmatics interface: Evidence from Spanish interlanguage of Greek speakers. In Mary Grantham O'Brien, Christine Shea & John Archibald (eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)*, 88–97. Somerville, Mass.: Cascadilla Proceedings Project.
Marsden, Emma, Sophie Thompson & Luke Plonsky. 2018. A methodological synthesis of self-paced reading in second language research. *Applied Psycholinguistics* 39(5). 861–904. doi:10.1017/S0142716418000036.
Meteyard, Lotte & Robert A.I. Davies. 2020. Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language* 112. 104092. doi:10.1016/j.jml.2020.104092.
Muntendam, Antje. 2009. Linguistic transfer in Andean Spanish: Syntax or pragmatics? Urbana, Ill.: University of Illinois at Urbana-Champaign Dissertation.
Ocampo, Francisco. 2009. El orden de palabras en el español hablado: La construcción sujeto verbo objeto directo. In Montserrat Veyrat Rigat & Enrique Serra Alegre (eds.), *La lingüística como reto epistemológico y como acción Social: Estudios dedicados al Profesor Ángel López García con ocasión de su sexagésimo aniversario*, 501–511. Madrid: Arco Libros.
Ortega-Santos, Iván. 2006. On new information focus, sentence stress assignment conditions and the copy theory: a Spanish conspiracy. *University of Maryland Working Papers in Linguistics* 14. 188–212.
Pladevall Ballester, Elisabet. 2010. Child L2 development of syntactic and discourse properties of Spanish subjects. *Bilingualism: Language and Cognition* 13(02). 185–216. doi:10.1017/S1366728909990447.
Reinhart, Tanya. 2006. *Interface strategies: optimal and costly computations*. Cambridge, Mass.: MIT Press.
Rohde, Doug. 2003. *Linger*. http://tedlab.mit.edu/~dr/Linger/.
Rothman, Jason. 2009. Pragmatic deficits with syntactic consequences?: L2 pronominal subjects and the syntax–pragmatics interface. *Journal of Pragmatics* 41(5). (Pragmatic Markers). 951–973. doi:10.1016/j.pragma.2008.07.007.
Rothman, Jason & Roumyana Slabakova. 2011. The Mind-Context Divide: On acquisition at the linguistic interfaces. *Lingua* 121(4). 568–576. doi:10.1016/j.lingua.2011.01.003.

Samek-Lodovici, Vieri. 2001. Crosslinguistic typologies in Optimality Theory. In Géraldine Legendre, Jane Barbara Grimshaw & Sten Vikner (eds.), *Optimality-Theoretic syntax*, 315–354. Cambridge, Mass.: MIT Press.
Schütze, Carson T. & Jon Sprouse. 2013. Judgment data. In Robert J. Podesva & Devyani Sharma (eds.), *Research methods in linguistics*, 27–50. Cambridge: Cambridge University Press.
Schwartz, Bonnie D. & Rex A. Sprouse. 1996. L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* 12(1). 40–72. doi:10.1177/026765839601200103.
Sequeros-Valle, Jose, Bradley Hoot & Jennifer Cabrelli. 2020. Clitic-doubled left dislocation in L2 Spanish: The effect of processing load at the syntax-discourse interface. *Language Acquisition* 27(3). 306–330. doi:10.1080/10489223.2020.1769628.
Slabakova, Roumyana. 2011. Which features are at the syntax–pragmatics interface? *Linguistic Approaches to Bilingualism* 1(1). 89–93. doi:10.1075/lab.1.1.14sla.
Slabakova, Roumyana. 2015. The effect of construction frequency and native transfer on second language knowledge of the syntax–discourse interface. *Applied Psycholinguistics* 36(3). 671–699. doi:10.1017/S0142716413000386.
Slioussar, Natalia. 2011. Processing of a Free Word Order Language: The Role of Syntax and Context. *Journal of Psycholinguistic Research* 40(4). 291–306. doi:10.1007/s10936-011-9171-5.
Sorace, Antonella. 2000. Syntactic optionality in non-native grammars. *Second Language Research* 16(2). 93–102. doi:10.1191/026765800670666032.
Sorace, Antonella. 2004. Native language attrition and developmental instability at the syntax-discourse interface: Data, interpretations and methods. *Bilingualism: Language and Cognition* 7(2). 143–145. doi:10.1017/S1366728904001543.
Sorace, Antonella. 2011. Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism* 1(1). 1–33. doi:10.1075/lab.1.1.01sor.
Sorace, Antonella. 2012. Pinning down the concept of interface in bilingual development: A reply to peer commentaries. *Linguistic Approaches to Bilingualism* 2(2). 209–217. doi:10.1075/lab.2.2.04sor.
Sorace, Antonella & Francesca Filiaci. 2006. Anaphora resolution in near-native speakers of Italian. *Second Language Research* 22(3). 339–368. doi:10.1191/0267658306sr271oa.
Sorace, Antonella & Ludovica Serratrice. 2009. Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism* 13(2). 195–210. doi:10.1177/1367006909339810.
Tomlin, Russell S. 1986. *Basic word order: Functional principles*. London: Croom Helm.
Tsimpli, Ianthi Maria & Antonella Sorace. 2006. Differentiating interfaces: L2 performance in syntax-semantics and syntax-discourse phenomena. *Proceedings of the 30th annual Boston University Conference on Language Development (BUCLD 30)*, 653–664. Somerville, Mass.: Cascadilla Press.
Verhoeven, Elisabeth & Stavros Skopeteas. 2015. Licensing focus constructions in Yucatec Maya. *International Journal of American Linguistics* 81(1). 1–40.
Weskott, Thomas, Robin Hörnig, Gisbert Fanselow & Reinhold Kliegl. 2011. Contextual Licensing of Marked OVS Word Order in German. *Linguistische Berichte* 225. 3–18.
White, Lydia. 2011. The interface hypothesis: How far does it extend? *Linguistic Approaches to Bilingualism* 1(1). 108–110. doi:10.1075/lab.1.1.19whi.

Zubizarreta, Maria Luisa. 1998. *Prosody, focus, and word order*. Cambridge, Mass.: MIT Press.