

Laura Stigliano* and Ming Xiang

Experimental Evidence on Island Effects in Spanish Relative Clauses

<https://doi.org/10.1515/PRBS-2021-0008>

Published online November 11, 2021

Abstract: Research on islands has been central to linguistic theory for more than 50 years. Its importance relies on the theoretical consequences islands posit for movement and long distance dependencies. In this paper we aim to explore the contrast between a variety of islands in Spanish relative clauses to reveal whether there is any gradience in the strength of the island effects. In order to tease apart fine-grained contrasts we run an acceptability judgment study based on the factorial definition of island, an experimental paradigm that aims to isolate the various factors that can affect the acceptability of a sentence involving island violations. Overall, we found that the five constructions tested (embedded *wh*-questions, *whether*-clauses, adjuncts, complex NPs and relative clauses) show island effects in Spanish and that there are limited differences in the size of these effects, which points to a more categorical view of islands.

Keywords: islands, relative clauses, gradience, Spanish, experimental linguistics

1 Introduction

Research on islands has been central to linguistic theory since Ross (1967)'s work. Its importance relies on the theoretical consequences islands posit for movement and long distance dependencies. In this paper we report and discuss the findings from an acceptability judgment task that tested five different island structures in Spanish. We explore the contrast between five islands types where the extracted element is the head noun of a relative clause, and the variation of island effects across these constructions. Our main goal is to reveal which structures are indeed islands in this language and whether there is any evidence of gradience in the strength of island effects. In order to tease apart fine-grained contrasts we run an

*Corresponding author: Laura Stigliano, Department of Linguistics, The University of Chicago, Division of the Humanities, 115 E 58th St, Chicago, IL 60637-1511, USA, E-mail: laurastigliano@uchicago.edu

Ming Xiang, Department of Linguistics, The University of Chicago, Division of the Humanities, 115 E 58th St, Chicago, IL 60637-1511, USA, E-mail: mxiang@uchicago.edu

acceptability judgment study based on the factorial definition of island effects (Sprouse et al. 2012), an experimental paradigm that aims to isolate the various factors that can affect the acceptability of a sentence.

The structure of this paper is as follows. In the remainder of this Section we introduce the previous claims and proposals that motivate our study: in Section 1.1 we provide a brief background on island effects, in Section 1.2 we summarize the main features of the factorial definition of islands, and in Section 1.3 we outline previous claims and findings on islands in Spanish. In Section 2 we describe our study and in Section 3 we summarize the results obtained. In Section 4 we discuss our results and conclude.

1.1 Brief Background on Island Effects

The processing of long-distance dependencies requires relating two syntactic elements that are not in a local relation with each other. This demands different resources from the parser. For instance, the parser needs to maintain the uninterpreted element in the working memory and it also needs to determine the specific point in the sentence at which that element is interpreted. The relations between the moved phrases and their associated verbs are often referred to as ‘unbounded dependencies’. In English, for instance, a *wh*-phrase may be separated from the verb by any number of elements, as shown in (1). In the theoretical literature, these dependencies are called *A'*-dependencies, and the psycholinguistic literature refers to them as *filler-gap dependencies*. For illustrative purposes, in all the examples in this paper, the *filler* will be **bolded**, and the *gap* will be represented by the underscore ‘ ’.

- (1) a. **What** did Sonia buy ?
 b. **What** does Bruno think that Sonia bought ?
 c. **What** does Bruno think that Ana said that Sonia bought ?

Although *wh*-phrases can be extracted across multiple embedded clause boundaries, there are a number of syntactic environments where extraction is judged as unacceptable. These syntactic environments are known as *islands*, a term coined by Ross (1967). In this respect, island sensitivity is often considered a diagnostic of movement. The literature on this topic recognizes a variety of islands, which can be referred to based on the structure that creates them, as in (2a)–(2g), or based on the constraint they violate, as in (2h)–(2i):

- (2) a. *Whether*-islands
 ***What** did Sonia wonder [_{CP} whether Bruno bought]?

- b. Wh-islands
***What** does Sonia ask [_{CP} when Bruno bought ___]?
- c. Complex Noun Phrase (NP) islands
***What** did you hear [_{NP} the rumor that Sonia bought ___]?
- d. Subject islands
***Who** did [_{NP} pictures of ___] fell on Sonia's head?
- e. Adjunct islands
***What** does Sonia get happy [_{CP} when Bruno buys ___]?
- f. Relative Clause (RC) islands
***What** did Sonia meet [_{RC} the author who wrote ___]?
- g. Sentential Subject islands
***Who** did [_{CP} that Sonia help ___] surprised Bruno?
- h. Coordinate Structure Constraint violations
***What** did Sonia buy [_{ConjP} a book and ___]?
- i. Left Branch Extraction violations
***How tall** did Sonia meet [_{NP} ___ a woman]?

Importantly, in addition to wh-dependencies, island effects are observed in other constructions that involve movement such as relative clauses (3a), topicalizations (3b), and scrambling (in languages that allow it, such as Japanese):

- (3) a. *Sonia likes **the car** that Bruno wonders [_{CP} whether Ana bought ___].
 - b. *Sonia doesn't know who bought most of these cars, but **that car**, she wonders [_{CP} whether Ana bought ___].
- (Examples adapted from Sprouse et al. 2016)

For this reason, any successful account of the grammar, processing, or learning of island effects must be extended to account for all of them.

Different approaches have been proposed to account for the source of island effects since Ross (1967)'s seminal work. On the one hand, *grammatical* explanations postulate that syntactic constraints inherent to the grammatical system regulate the formation of long-distance dependencies (see e.g. Chomsky 1973, 1986; Huang 1982, among many others), and syntactic islands arise in situations where a certain type of syntactic movement (or extraction) is not licensed by the grammar. Under this approach, the existence of syntactic islands provide evidence for a rule-based grammatical system.¹ Grammatical explanations usually predict a categorical view of islands, that is, a construction is either an island or not; and island violation is also an across-construction generalization, i.e. all constructions that involve syntactic movement—such as wh-questions, relative clauses and

¹ For an overview on this topic, the reader can refer to Boeckx (2012) and references therein.

topicalizations—are predicted to exhibit similar island effects (although there has been some proposals that acknowledge a more gradient effect, as will be discussed below). On the other hand, *processing* explanations appeal to parsing constraints to explain islands, highlighting the presence of processing cost in structurally complex sentences (see, e.g. Hofmeister and Sag 2010; Hofmeister et al. 2013; Kluender and Gieselmann 2013; Kluender and Kutas 1993, among many others). This type of approach usually predicts a gradient effect of islands, since processing costs could vary based on specific constructions and contexts. A third type of account of island effects argues that the properties pertaining the information structure of a given construction play a crucial role in parsing sentences that contain an island structure (see e.g. Erteschik-Shir 1973; Takami 2012, among others). We can call these *discourse* explanations. For instance, Goldberg (2013) argues that islands arise when information that belongs to the background is extracted and is put in focus. This type of approach could predict gradience in island violations, depending on the fine-grained information structure properties of different constructions (e.g. island violations in *wh*-extractions are predicted to be more degraded than relativizations).

Various observations have also been made regarding *strong* and *weak* islands (see e.g. Cinque 1990; Postal 1998; Szabolcsi 2006). Descriptively, strong islands are domains in which no extraction can take place, and weak islands are domains from which some types of extractions, but not others, are possible. This is illustrated in (4). As these examples show, arguments can be extracted from a *whether* clause (4a), but adjuncts cannot (4b):

- (4) a. **Which topic** did John ask [whether to talk about ___]?
 b. ***How** did John ask [whether to behave ___]?
 (Examples adapted from Szabolcsi 2006)

The distinction between an argument versus an adjunct extraction from a weak island may be too coarse-grained since not all argument extractions from a weak island are considered completely acceptable. The acceptability judgments and the magnitude of the island violation effect could sometimes be very nuanced (i.e. Almeida 2014; Boeckx 2012; Cinque 1990). This is shown in the following examples of *whether* islands. According to Boeckx (2012), extraction of an argument gives rise to a mild degradation, represented by ‘??’, but extraction of an adjunct gives rise to full ungrammaticality, represented by ‘*’:

- (5) a. ??**Which student** do you wonder whether Bruno praised ___?
 b. ***How** do you wonder whether Bruno praised Sue ___?
 (Examples adapted from Boeckx 2012)

Almeida (2014) calls this kind of island sensitivity in which a given structure shows super-additive effects but it’s still not considered entirely ungrammatical

subliminal island effects (contrasting with more ‘traditional’ island effects, which are called *supraliminal island effects*). This distinction between different *degrees of deviance* has important consequences for a theory of islands (and for the overall architecture of grammar in general), and has led to different explanations for islandhood in each case. For instance, according to Rizzi (1990)’s *Relativized Minimality* analysis, a milder deviance is analyzed as an intervention effect, that is, the relation between a filler and its gap is interrupted by the intervention of a constituent of the same/similar type. What is more, to explain why some configurations give rise to a strong deviance in one language but only to a mild deviance in another language, Rizzi claimed that there were parametric differences between languages, and that the relevant syntactic constraints were ranked differently crosslinguistically. This variation in the strength of island effects is one of the main issues that yet to be fully addressed.

1.2 Testing Islandhood with a Factorial Design

Many theoretical proposals regarding the origin of islands were based on informal judgments, often times lacking systematic comparisons between the various structures under study. Additionally, some previous studies only considered the deviance of the island itself, informally comparing islands types without taking into consideration the base-line, non-island condition, which might show a difference to begin with. However, it has been shown that controlled experimental studies have the advantage of capturing fine-grained distinctions between different types of sentences. For example, Sprouse et al. (2012) used a factorial design to evaluate the contribution of different factors on acceptability, controlling for possible confounding factors that might obscure the judgments. This experimental paradigm has since then been adapted by studies on different languages such as Brazilian Portuguese (Almeida 2014), English (Michel 2014; Sprouse et al. 2012, 2016), Hebrew (Keshev and Meltzer-Asscher 2019), English and Italian (Sprouse et al. 2016), Japanese (Fukuda and Sprouse 2019), L1 and L2 English and Spanish (Ortega-Santos et al. 2018), Norwegian (Kush et al. 2018, 2019), Spanish (López Sancio 2015; Pañeda et al. 2020; Pañeda and Kush 2021), Slovenian (Stepanov et al. 2018), French and English (Abeillé et al. 2020), among many others. Since we will adapt this paradigm for the current study as well, we introduce the basic features of the paradigm below.

The study in Sprouse et al. (2012) manipulates two factors that can affect the acceptability of a sentence: the *distance* between a filler and a gap, and the *structural complexity* of the sentence containing the dependency. Each factor has two levels: SHORT and LONG for the former factor, and ISLAND and NON-ISLAND for the

latter factor. This gives rise to a 2×2 experimental design. An example for *whether-island* is given in (6), sentences are presented without acceptability judgments:

- | | | |
|-----|---|-------------------|
| (6) | a. Who __ thinks [that Sonia bought a car]? | NON-ISLAND, SHORT |
| | b. What does Bruno think [that Sonia bought __]? | NON-ISLAND, LONG |
| | c. Who __ wonders [whether Sonia bought a car]? | ISLAND, SHORT |
| | d. What does Bruno wonder [whether Sonia bought __]? | ISLAND, LONG |
- (Examples adapted from Sprouse et al. 2012)

In this *factorial design*, the first condition (6a) is the *baseline*, as it represents the combination of the “unmarked” level under the two factors. Conditions (6b)–(6c) deviates from the baseline by just one feature. The fourth condition (6d) is the combination of the marked levels in both factors. Note that although both (6c) and (6d) contain an island configuration, only (6d) contains an extraction out of an island (i.e. an island violation). Using this type of design, two effects can be obtained. The first one is simply *linear additivity*, which is the linear sum of individual costs. The second one is *super-additivity*, which is a combined effect that is greater than the linear sum of individual costs, in other words: $(6d - 6a) > (6b - 6a) + (6c - 6a)$. A super-additive effect is statistically evaluated by an interaction between the two factors *distance* and *structural complexity*. Additionally, a *Differences-In-Differences* (DD) score is calculated to assess the strength of the super-additive effect (see Section 3 for details on how to calculate DD scores). That is, a positive DD score reflects a super-additive effect, and a larger DD score is interpreted as a stronger island effect. On the other hand, a DD score of zero indicates no island effect.

The current study applies this paradigm to Spanish relative clauses, a language in which not many experimental studies have been conducted, specially in the context of relative clauses. We turn in the next section to briefly summarize the previous theoretical claims and experimental findings on islands in Spanish.

1.3 Islands in Spanish

Research on islands in Spanish was mostly based on informal judgments, which gave rise to a lot of variation with regard to the examples used and the judgments reported (see, e.g. Gallego 2011; Gallego et al. 2007; Haegeman 2014; Jiménez Fernández 2009; Suñer 1991; Torrego 1984). Most of the literature on this topic in this language analyze subject islands, discussing Torrego (1984)’s original observation that *wh*-movement out of subjects in Spanish is possible and interacts with the subject-inversion rule in this language. As for the discussion of other islands in Spanish, Torrego (1984) and Suñer (1991) claim that extractions out of

indirect/embedded questions are possible (contrasting with languages like English), as the following examples show (judgments are Torrego's):

- (7) a. **Quién** no sabes cuánto pesa ___?
 who not you.know how.much weights
 'Who don't you know how much weights?'
 b. **Quién** no sabes qué es __ en esta empresa?
 who not you.know what is in this firm
 'Who don't you know what is in this firm?'
 (Examples adapted from Torrego 1984)

What is more, Torrego claims that extraction out of non-selected questions (8a) or *whether*-clauses (i.e. embedded polar questions) (8b) are also possible:

- (8) a. **Qué** no te explicas por qué Juan habrá comprado ___?
 what not you understand why Juan will.have bought
 'What don't you understand why Juan will have bought ___?'
 b. **Qué** diccionario no sabías si Celia había devuelto ___?
 what dictionary not you.know if Celia had returned
 'What dictionary didn't you know whether Celia had returned ___?'
 (Examples adapted from Torrego 1984)

Additionally, Fábregas (2013) claims that relativization from a *whether*-clause (9a) is possible, and relativization of certain adjuncts, like temporal adjuncts, is also possible (9b), but relativization of cause adjuncts is not (9c):

- (9) a. **el juguete** que Ana se pondría contenta si su padre
 the toy that Ana se would.get happy if her father
 le compra ___
 to.her buys
 'the toy that Ana would get happy if her father buys for her'
 b. **el tipo** que los vecinos se enfadaron cuando __ tocó
 the guy that the neighbours got.mad when played
 la guitarra
 the guitar
 'the guy that the neighbours got mad when he played the guitar'
 c. ***el tipo** que María se puso contenta porque __ vino
 the guy that María se got happy because came
 'the guy that María got happy because he came'
 (Examples adapted from Fábregas 2013)

Although this is not an exhaustive summary of previous work on islands in Spanish, we can already observe a substantial degree of variation. First, there are cross-linguistic differences in the islandhood status of certain configurations (e.g. *wh*- and *whether* constructions are claimed to be islands in English, but not in

Spanish). Second, there are variations with respect to the judgments offered for various island types. And finally, there are also variations with respect to how the examples of islands are constructed, which makes it difficult to compare between various claims. A number of studies on Spanish have conducted controlled experiments to better assess the variations found in the syntax literature.

López Sancio (2015) used a factorial design to test four islands with both wh-extractions (i.e. wh-questions) and extractions out of relative clauses. An example of the study's stimuli is provided in (10) only for relative clause extractions (judgments based on the study's results). The islands tested in this study were wh-islands (10a) with *por qué* 'why', *cuándo* 'when' and *cómo* 'how', complex NP islands (10b), subject islands (10c) and adjunct islands (10d). The study was divided into four smaller studies, and participants saw only one island type for the wh-extraction condition and one island type for the relative clause condition.

- (10) a. *Aurora adora a **la cantante** con la que Patri
 Aurora loves DOM the singer with the that Patri
 se pregunta cuándo se casará el rapero __.
 SE wonders when SE will.marry the rapper
 'Aurora loves **the singer** that Patri wonders when the rapper will
 marry __.'
- b. *Conozco **al entrenador** con el que el jugador oyó
 I.know DOM.the coach with the that the player heard
 los rumores sobre que Ana está saliendo __.
 the rumors about that Ana is.dating
 'I know **the coach** who the player heard the rumours about the fact
 that Ana is dating __.'
- c. He visto **al político** del que crees que varios
 I.have seen DOM.the politician of.the that think that some
 escándalos __ han suscitado rumores sobre el candidato.
 scandals have fuelled rumours about the candidate
 'I've seen **the politician** who you think that some scandals of __
 have fueled rumours about the candidate.'
- d. *He visto **al cliente** que el dependiente llamará
 I.have seen DOM.the customer that the shop.assistant will.call
 a su jefa si viene __
 DOM his boss if comes
 'I have seen **the customer** who the shop assistant will call his boss if
 __ comes.'

(Examples adapted from López Sancio 2015)

The results of this study show island effects for all the wh-dependencies tested. For relative-clause dependencies, island effects were only found for wh-, complex NP

and adjunct islands, but not for subject islands. DD-scores are reported in this study but no statistical analysis of them is done to see whether there's a significant difference in the strength of the effects for each island.

Ortega-Santos et al. (2018) also applied the factorial design and tested extraction of *wh*-phrases out of embedded interrogative clauses (i.e. *wh*-islands) with *por qué* 'why', as shown in (11) (judgment based on the study's result). Sentences were preceded by a context (not included in the example below). This study tested L1 Spanish and English speakers, and L2 Spanish and English speakers. With regard to the L1 Spanish speakers, the group relevant for our study, the study found island effects.

- (11) ***Quién** no sabes por qué __ escribió el informe?
 who not you.know why wrote the report
 'Who don't you know why __ wrote the report?'
 (Example adapted from Ortega-Santos et al. 2018)

Pañeda et al. (2020) examined islands using speeded acceptability judgment task, also with a factorial design. Their study tested *wh*-extractions (i.e. *wh*-questions) out of four islands: subject islands (12a), complex NP islands (12b), adjunct islands (12c) and *whether* islands (12d):

- (12) a. ***De quién** crees que el discurso __ofendió tanto a
 of who you.believe that the discourse offended so.much DOM
 Julia?
 Julia
 'Of who do you believe that the discourse __ offended Julia so much?'
 b. ***Qué** has hecho la petición de que resolvamos __
 what you.have made the petition of that we.solve __
 el viernes?
 the Friday
 'What have you made the petition that we solve __ on Friday?'
 c. ***Qué** protestaste cuando anunciamos __en la
 what you.complaint when we.announced in the
 reunión de improviso?
 meeting of unexpected
 'What did you complain when we unexpectedly announced __
 at the meeting?'
 d. ***Qué** preguntas si hemos encontrado __ por casualidad?
 what you.asked if we.have found by chance
 'What do you ask if we have found __ by chance?'

The study tested these structures using a word-by-word presentation procedure followed by with a forced-choice task (i.e. participants had to choose between *acceptable* and *unacceptable*). Using a Bayesian statistical analysis, they found

that all four constructions tested display super-additive effects. They further analyzed the strength of the island effect for each construction and found that it varied in a way that complex NP islands showed the weakest effect, subject islands showed the strongest effect, and *whether* and adjunct islands showed an intermediate effect.

Finally, in a recent study, Pañeda and Kush (2021) examined island effects in embedded questions. In their first experiment they used a factorial design to test the effects of responsive verbs like *know* (13a) and rogative verbs like *ask* (13b) in embedded *whether* questions (i.e. embedded polar questions), including a preceding context to pragmatically motivate the target sentence (context omitted in the examples below):

- (13) a. **Qué paquete** no sabía el funcionario si habíamos
 what package NEG knew the officer whether had
 recogido __ ?
 pick.up
 'Which package didn't the officer know whether we had picked up __ ?'
- b. **Qué paquete** preguntó el funcionario si habíamos
 what package asked the officer whether had
 recogido __ ?
 pick.up
 'Which package did the officer ask whether we had picked up __ ?'
- (Examples adapted from Pañeda and Kush 2021)

Their results showed that these structures were rated in the acceptable range with mostly high scores. This result contrasts with adjunct and relative clause islands, also tested in this study as a control, which were rated low and were taken to be ungrammatical. Pañeda and Kush concluded that *whether* structures are not islands in Spanish. Crucially, these results differ from those reported by Suñer (1991) and also those found in López Sancio (2015) and Pañeda et al. (2020). They attributed this difference to the presence of the context preceding the target sentence.

In their second experiment they tested the same verbs in embedded wh-questions with *when* (14a)–(14b) (judgments based on the study's results):

- (14) a. ***Qué paquete** no sabía el funcionario cuándo habíamos
 what package NEG knew the officer when had
 recogido __ ?
 pick.up
 'Which package didn't the officer know when we had picked up __ ?'

- b. ***Qué paquete** preguntó el funcionario cuándo habíamos
 what package asked the officer when had
 recogido __ ?
 pick.up
 ‘**Which package** did the officer ask when we had picked up __ ?’
 (Examples adapted from Pañeda and Kush 2021)

In this case, the structures tested did show island effects and were considered ungrammatical, but the effect was smaller compared to adjunct and relative clause islands.

To sum up, the four studies discussed above have tested a variety of islands for different types of extractions in Spanish. Table 1 summarizes their findings.^{2,3}

Table 1: Summary of experimental studies on islands in Spanish.

Study	Sentence type	Island type	Island?	DD-scores
López Sancio (2015)	wh-question	wh-	Yes	1.65
		complex NP	Yes	1.41
		subject	Yes	0.80
		adjunct	Yes	1.71
	relative clause	wh-	Yes	1.18
		complex NP	Yes	2.11
		subject	No	0.42
		adjunct	Yes	1.52
Ortega-Santos et al. (2018)	wh-question	wh-	Yes	–
Pañeda et al. (2020)	wh-question	subject	Yes	–
		complex NP	No	–
		adjunct	Yes	–
		<i>whether</i>	Yes	–
Pañeda and Kush (2021)	wh-question	<i>whether</i>	No	0.22/0.38
		wh-	Yes	1.09/1.39
		adjunct	Yes	1.46/1.78
		relative clauses	Yes	1.82/1.79

Most of the studies tested wh-questions, only one examined islands in which the extracted element was the head noun of a relative clause. It is important to test this type of extraction because, as Abeillé et al. (2020) argues, it can provide evidence for discourse-based explanations of island effects if they are shown to be

² The island ‘names’ used in Table 1 might differ from the one used by the researchers in each study. We decided to standardize all the island names them to make comparisons more clear.

³ The two DD scores for Pañeda and Kush (2021)’s results for *whether* and wh- correspond to extractions out of clauses with *know* and *ask* respectively. The two DD scores for adjunct and relative clause islands correspond to Experiment 1 and 2 respectively.

less degraded than *wh*-extractions. In addition, we are particularly interested in testing whether there is any difference in the strength of island effects, which can help refine a theory of the source of island effects. In this respect, inconsistent results were reported across studies. For instance, while Pañeda et al. (2020) report that *whether* and adjunct islands show an intermediate effect, Pañeda and Kush (2021) found that the former are not actually islands in Spanish and the latter exhibit a stronger effect. As for relative clauses, although DD-scores are reported in López Sancio (2015), no analysis was performed on these data points. Finally, in all the studies cited, the existence and strength of the island effects were measured using different baselines for each island type, which might make comparisons between islands less warranted. The current study contributes to the existing body of empirical work by testing five different islands against the same baseline condition.

2 The Current Study

The current study⁴ tested and compared extractions out of relative clauses for six different constructions in Spanish, which included a non-island structure and five islands types: *wh*-islands, *whether* islands, complex NP islands, adjunct islands and relative clause islands. As mentioned above, relative clauses have not been widely tested in this language, and they were only examined in López Sancio (2015). We were interested, first, in replicating his results for *wh*-, complex NP, and adjunct islands, but also in testing other two structures: *whether* islands (i.e. embedded polar questions) and relative clause islands—the former subject to debate in terms of its island status (Pañeda and Kush 2021; Suñer 1991; Torrego 1984). We aimed to determine which structures give rise to island effects in Spanish and to tease apart fine-grained distinctions to reveal whether there is evidence of any gradability in the judgments obtained for these constructions. Crucially, our study differs from all the other studies reported in the literature on Spanish in that we used the same baseline for all the islands tested, and we tested all island types in the same population. We describe further differences between our study and previous studies in Section 2.2 below.

An acceptability judgment experiment was designed following the factorial definition of island effects (Sprouse et al. 2012). The factorial design incorporated two factors that are known to affect the acceptability of a sentence independently of grammatical constraints: the distance between a gap and its antecedent ('Length') and the structural complexity of the sentence containing the

⁴ Additional files, including materials, data and analysis code, can be found at the following Open Science Framework website <https://osf.io/t9vpq/>.

dependency ('Structure'). We tested five construction types and two extraction types for relative clauses (see experimental design and sample stimuli below).

2.1 Participants

Eighty participants were recruited on social media. Three participants were removed since they reported to be bilingual speakers (two participants reported being Spanish-English bilinguals and one participant reported being a Spanish-French bilingual). The remaining 77 participants were native speakers of Spanish: 58 participants from Buenos Aires, Argentina; 11 participants from other Argentinian provinces, four participants from Spain, six participants from other Latin American countries (two participants from Mexico, two participants from Ecuador, one participant from Chile, and one participant from Venezuela). Furthermore, two participants were excluded due to their failure in following the instructions. The analysis was carried over the data from the remaining 75 participants, which had a mean age of 30.85 years old (range: 18–57 years). Fifty seven participants were self-identified as female, and 24 were self-identified as male. All participants provided their consent to participate in the study. As compensation, each participant received a \$5 gift-card.

2.2 Materials

We tested six construction types (non-islands, wh-islands, *whether* islands, complex NP islands, adjunct islands and relative clause islands) and two length variations (short and long), resulting in a 6×2 design (12 conditions). There were 36 sentence sets of items (for a total of 432 sentences), and each participant saw three items per condition, from different sets. That is, each participant saw three sentences that belonged to the non-island condition, three sentences that belonged to the wh-island condition, three sentences that belonged to the complex NP island condition, and so on. The main difference between our experimental design and previous studies is that we compared all island structures to the same non-island condition. This ensures that our inferences about the strength of island effects based on comparisons between each island type to the same baseline. As we will show below, each condition minimally differ from the baseline and from other conditions (although they didn't form strict minimal pairs), which allowed us to carry the desired comparisons between these structures. Experimental items were intermixed with a set of 36 fillers that contained an equal number of acceptable and unacceptable sentences. The order of presentation of fillers and experimental trials was randomized on a by-participant basis and experimental items were distributed across Latin square lists, such that each participant only saw one condition from each 12-condition experimental item-set. Each participant rated a total of 72 items

(36 experimental and 36 fillers),⁵ and the task took around 20 min. Sample stimuli are shown below for each construction type (examples are provided without a grammatical judgment). All the sentences were relative clauses, where the direct object of the main verb was the (extracted) relativized argument, which was also a direct object of the embedded verb. This differs from the stimuli in López Sancio (2015) whose relativized argument was either a direct object or a prepositional phrase, which was either the direct object, indirect object or prepositional phrase of the embedded verb. In addition, all the extracted DPs were human DPs, which differs from the stimuli used in Pañeda et al. (2020) and Pañeda and Kush (2021), who intermixed human and non-human DPs. All our sentences also had the same initial structure, schematically shown in (15):

(15) DP_{Subject} V_{Matrix} DP_{DirectObject} that...

The Subject DP was always a proper name, and half the sentences contained typically female names and half the sentences contained typically male names. The verb was always a transitive verb. The extracted argument was always a DP, and half the sentences contained a DP with feminine inflection and half the sentences contained a DP with masculine inflection. The complementizer used was always a bare *que* ‘that’.⁶

In the examples below ((16)–(21)), the examples in (a) are the *short* conditions for each structure, and the examples in (b) are the *long* condition for each structure. Following previous studies, the short condition was generated by extracting the subject of the first embedded verb, and the long condition was generated by extracting the object of the second embedded verb. Therefore only the *long* condition involves extraction out of an island. The *non-island* condition (16) was the baseline of comparison for *all* the other island structures and involved a bridge verb (i.e. *escuchar* ‘to hear’) as the first embedded verb:

(16) Non-island (baseline for all the other islands)

- a. Sonia vio **al** **profesor** que __ escuchó que Bruno contrató
 Sonia saw DOM.the professor that heard that Bruno hired
 a Ana.
 DOM Ana
 ‘Sonia saw the professor that heard that Bruno hired Ana.’

⁵ As an anonymous reviewer points out, fillers were evenly distributed between grammatical and ungrammatical sentences, but target conditions were not. This didn’t lead to a 1:1 ratio of grammatical and ungrammatical sentences.

⁶ This differs from the stimuli in López Sancio (2015) in various ways. I refer the reader to that study for further information.

- b. Sonia vio **al** **profesor** que Ana escuchó que Bruno
 Sonia saw _{DOM.}the professor that Ana heard that Bruno
 contrató __ .
 hired
 ‘Sonia saw the professor that Ana heard that Bruno hired.’

In order to create a *whether*-island (17), we minimally modified the sentences in (16) in that the first embedded verb was changed to *preguntar* ‘to ask’, followed by the complementizer *si* ‘if/whether’:

(17) *Whether* island

- a. Sonia vio **al** **profesor** que __ preguntó si Bruno contrató
 Sonia saw _{DOM.}the professor that asked if Bruno hired
 a Ana.
_{DOM} Ana
 ‘Sonia saw the professor that asked whether Bruno hired Ana.’
- b. Sonia vio **al** **profesor** que Ana preguntó si Bruno contrató __ .
 Sonia saw _{DOM.}the professor that Ana asked if Bruno hired
 ‘Sonia saw the professor that Ana asked whether Bruno hired.’

As for *wh*-islands, the first embedded verb was also *preguntar* ‘to ask’, as in the examples above, but in these cases it was followed by a *wh*-word *quién* ‘who’, which was the subject of the second embedded verb:

(18) *Wh*-island

- a. Sonia vio **al** **profesor** que __ preguntó quién contrató a
 Sonia saw _{DOM.}the professor that asked who hired _{DOM}
 Ana.
 Ana
 ‘Sonia saw the professor that asked who hired Ana.’
- b. Sonia vio **al** **profesor** que Ana preguntó quién contrató __ .
 Sonia saw _{DOM.}the professor that Ana asked who hired
 ‘Sonia saw the professor that Ana asked who hired.’

With respect to complex NP islands (19), the verb *escuchar* ‘to hear’ was used, as in the non-island condition, but it was followed by *el rumor de que* ‘the rumor that’, which was the NP from which the argument was extracted. As in the examples above, the short condition involved extraction from the subject of the first embedded verb (19a), and the long condition involved extraction from the second embedded verb (19b):

(19) Complex NP island

- a. Sonia vio **al profesor** que __ escuchó el rumor de que
 Sonia saw _{DOM}.the professor that heard the rumor of that
 Bruno contrató a Ana.
 Bruno hired _{DOM} Ana
 ‘Sonia saw the professor that heard the rumor that Sonia hired Ana.’
- b. Sonia vio **al profesor** que Ana escuchó el rumor
 Sonia saw _{DOM}.the professor that Ana heard the rumor
 de que Bruno contrató __.
 of that Bruno hired
 ‘Sonia saw the professor that Ana heard the rumor that Bruno
 hired.’

To create adjunct islands we used an intransitive verb as the first embedded (*alegrarse* ‘get.happy’, in this case), followed by a temporal adjunct headed by *cuando* ‘when’. The extracted argument was extracted from that adjunct:

(20) Adjunct island

- a. Soniavio **al profesor** que __ se alegró cuando Bruno contrató.
 Sonia saw _{DOM}.the professor that got.happy when Bruno hired
 a Ana
_{DOM} Ana
 ‘Sonia saw the professor that got happy when Bruno hired Ana.’
- b. Soniavio **al profesor** que Anase alegró cuando Bruno contrató __
 Sonia saw _{DOM}.the professor that Ana got.happy when Bruno hired
 ‘Sonia saw the professor that Ana got happy when Bruno hired.’

Finally, to create relative clause islands we relativized the subject of the second embedded verb; this relative clause was the one from which the filler was extracted:

(21) Relative Clause island

- a. Sonia vio **al profesor** que __ escuchó que los
 Sonia saw _{DOM}.the professor that heard that the
 empleados que contrataron a D. renunciaron ayer.
 employees that hired _{DOM} D. resigned yesterday
 ‘Sonia saw the professor that heard that the employees that hired D.
 resigned yesterday.’
- b. Sonia vio **al profesor** que D. escuchó
 Sonia saw _{DOM}.the professor that D. heard
 que los empleados que contrataron renunciaron ayer __.
 that the employees that hired resigned yesterday
 ‘Sonia saw the professor that Ana heard that the employees that hired
 resigned yesterday.’

To sum up, we used the same baseline (non-island sentences) for all the islands tested. We minimally modified that structure to obtain the different island constructions. Although not strict minimal pairs, different conditions are nonetheless similar to each other, which allowed us to run comparisons between islands to assess whether there is any evidence of gradient effects.

2.3 Procedure

The acceptability judgment task was run on IbexFarm (Drummond 2013), a web-based, online tool designed for linguistics studies. Each participant used their own computer. Participants first signed a consent form and completed a short demographic questionnaire. Then, they rated the acceptability of a set of sentences on a 1–7 scale, with 1 indicating that the sentence was completely *unacceptable*, and seven indicating that it was completely *acceptable*. Participants were instructed to base their judgments on their intuitions as native speakers of Spanish, and not to base their judgments on any prescriptive rules or the plausibility of the situations described. After reading the instructions, participants were given three practice trials. Participants were allowed to take one or two breaks during the task. Participants saw one sentence at a time, providing their rating either by entering the number with the keyboard or by clicking on the number on the screen. Once they chose a number, the sentence disappeared and the next sentence was displayed.

3 Analysis and Results

The critical dependent measure was the acceptability response given for each sentence, which range from 1 (completely unacceptable) to 7 (completely acceptable). Before data analysis, we transformed each individual's raw acceptability ratings into *z*-scores. Mean acceptability ratings and *z*-scores by condition are reported in Table 2:

Table 2: Mean acceptability rating ('rating') and mean normalized score ('z-score') by condition. Standard error (SE) shown in parenthesis.

	SHORT		LONG	
	Rating	<i>z</i> -score	Rating	<i>z</i> -score
Non-islands	5.41 (0.14)	0.44 (0.05)	4.63 (0.14)	0.14 (0.05)
Adjunct islands	6.13 (0.11)	0.71 (0.04)	2.41 (0.12)	−0.72 (0.04)
Wh-islands	6.15 (0.11)	0.72 (0.04)	2.91 (0.13)	−0.52 (0.05)
Complex NP islands	5.80 (0.12)	0.59 (0.04)	2.12 (0.1)	−0.84 (0.03)
Rel clause islands	4.74 (0.16)	0.17 (0.05)	1.68 (0.09)	−1.01 (0.03)
<i>Whether</i> islands	6 (0.11)	0.67 (0.04)	3.48 (0.14)	−0.30 (0.05)

The mean rating for the grammatical fillers was 6.78 (SE = 0.02), and the mean rating for the ungrammatical fillers was 1.74 (SE = 0.04). The results on the fillers confirm that participants understood the task. Figure 1 shows the ratings by island type, displaying z-scores. Each island type is plotted together with the baseline (the non-island conditions). Visual inspection on Figure 1 suggests that all islands showed some degree of super-additive effects (i.e. the two lines in each plot are not parallel):

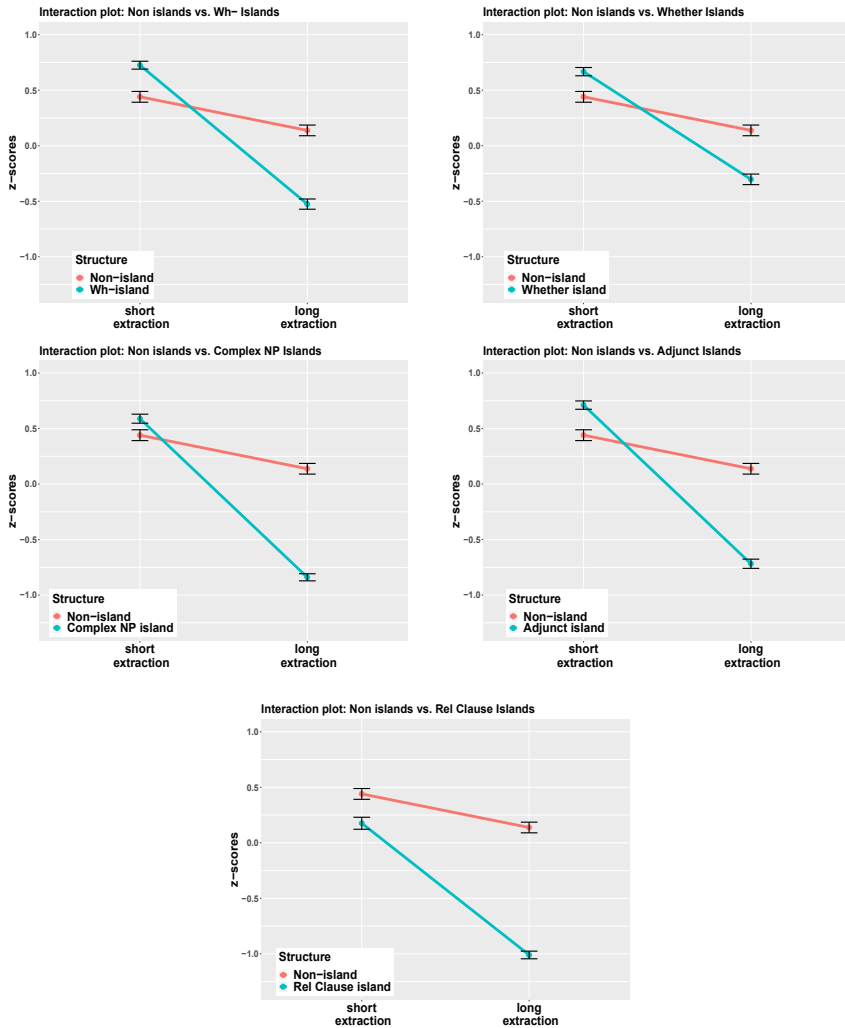


Figure 1: Plots by Island Type. Error bars correspond to ± 1 SE.

For the statistical analysis, we ran linear-mixed effects models on z -scores, using the “lmerTest” package (Kuznetsova et al. 2017) in the R software (R Core Team 2013). We constructed five separate models, one for each island type. For each model, we included the data from a given island and the shared baseline non-island conditions (four conditions total for each model). All the models contained the predictor ‘Length’ (short vs. long) and ‘Structure’ (non-island vs. island), as well as their interaction as the fixed effects. We also included by-participant random intercept. More complex models did not successfully converge. The two predictors were sum-coded before entering into the models. According to the current design, island effects should emerge as a significant interaction between ‘Length’ and ‘Structure’. We report the results in Table 3.

Table 3: Results of linear mixed models per island type compared to the baseline.

	β	SE	t	p	
Adjunct islands					
Structure	0.15	0.02	6.92	<0.0001	***
Length	0.43	0.02	20.48	<0.0001	***
Structure \times length	-0.28	0.02	-13.32	<0.0001	***
Complex NP islands					
Structure	0.21	0.02	10.18	<0.0001	***
Length	0.43	0.02	21.25	<0.0001	***
Structure \times length	-0.28	0.02	-13.83	<0.0001	***
Wh-islands					
Structure	0.1	0.02	4.41	<0.0001	***
Length	0.39	0.02	18.09	<0.0001	***
Structure \times length	-0.24	0.02	-11.05	<0.0001	***
Whether islands					
Structure	0.05	0.02	2.53	<0.0001	***
Length	0.32	0.02	15.02	<0.05	*
Structure \times length	-0.17	0.02	-7.88	<0.0001	***
Relative clause islands					
Structure	0.35	0.02	15.81	<0.0001	***
Length	0.37	0.02	16.66	<0.0001	***
Structure \times length	-0.22	0.02	-9.89	<0.0001	***

To summarize, for all island types we found a significant effect of Structure, a significant effect of Length, and most importantly, all island types showed a significant interaction ($p < 0.0001$).

Following Sprouse et al. (2012), we measured the strength of islands effects with *Differences-in-Differences* (DD) scores (Maxwell et al. 2017). The DD scores is a measure that captures the strength of the super-additive interaction. To calculate it, first, we calculated the difference (D1) between the NON-ISLAND/LONG condition and the ISLAND/LONG condition. Second, we calculated the difference (D2) between the

scores in the other two conditions (i.e. NON-ISLAND/SHORT and ISLAND/SHORT). Finally, we calculated the difference between these two difference scores. According to Sprouse et al., a positive DD score reflects a super-additive interaction, and the larger the DD, the larger the interactions, which is interpreted as a stronger island effect; on the contrary, a DD score of zero represents no interaction at all (i.e. no island effect). Our DD scores are summarized in Table 4, and show that there is a super-additive effect in all constructions, henceforth, we can conclude that all constructions are indeed islands.

Table 4: DD scores.

	DD SCORE
DD adjunct islands	1.13
DD complex NP islands	1.13
DD wh-islands	0.95
DD relative clause islands	0.88
DD <i>Whether</i> islands	0.67

We were interested in testing whether there was a difference in the strength of the island effect across constructions, i.e. whether some islands would give rise to a stronger effect. That is, we were interested in testing if some kind of gradability in the island effects can be supported statistically. To do this, we ran two models on the DD scores. The first model included the structure type based on which the DD scores were computed as a fixed effect predictor and a by-participant random intercept. The second model is a null model that only contained the by-participant random intercept.⁷ A model comparison between these two showed that there is a significant effect of the structure type based on which the DD scores were computed ($X^2(4) = 30.578, p < 0.0001$). That is to say, there is a difference regarding the strength of the island effect across the constructions tested. To better understand the source of this difference, we ran pair comparisons on DD scores for every pair of constructions, using the *glht* function from the “multcomp” package in R (Hothorn et al. 2008). The pair comparisons revealed a significant difference between *whether* islands and three other islands: adjunct islands ($\beta = 0.46, SE = 0.1, p < 0.001$), complex NP islands ($\beta = -0.46, SE = 0.1, p < 0.001$), and wh-islands ($\beta = -0.28, SE = 0.1, p < 0.05$), but not relative clause islands ($\beta = -0.22, SE = 0.1, p = 0.16$). All the other islands didn’t show a significant difference compared to each other. These results are schematized in (22):

⁷ The model with the structure type predictor `lmer(DDscore ~ DD.structure.Type + (1|Participant), data)` and the null model `lmer(DDscore ~ 1 + (1|Participant), data)`. More complex models (e.g. including Item as random intercept) did not converge successfully.

- (22) *Whether* < (Adjunct = CNP = Wh)
Whether = RC
 (where ‘<’ means that the island effect size is *smaller* than, and ‘=’ means that the island effect size is *equal to*)

These results therefore suggests a selective gradability regarding the strength of the island effect across different island constructions—only the *whether*-island showed weaker island effect relative to most other islands tested here.

4 Discussion & Conclusions

Our study tested six constructions in Spanish. They all involved extractions out of a relative clause, and one of them was considered the baseline, non-island condition. We aimed to show whether the other five constructions show island effects and whether there was a difference in the strength of this effect. Our results show that the five constructions tested exhibit super-additive effects, which means that the five constructions should be considered islands in Spanish under a factorial definition of islands. We repeat Table 1 below, adding the results obtained in this current study (Table 5).

Table 5: Summary of experimental studies on islands in Spanish.

Study	Sentence type	Island type	Island?	DD-scores
López Sancio (2015)	wh-question	wh-	Yes	1.65
		complex NP	Yes	1.41
		subject	Yes	0.80
		adjunct	Yes	1.71
	relative clause	wh-	Yes	1.18
		complex NP	Yes	2.11
		subject	No	0.42
		adjunct	Yes	1.52
Ortega-Santos et al. (2018)	wh-question	wh-	Yes	–
Pañeda et al. (2020)	wh-question	subject	Yes	–
		complex NP	No	–
		adjunct	Yes	–
		<i>whether</i>	Yes	–
Pañeda and Kush (2021)	wh-question	<i>whether</i>	No	0.22/0.38
		wh-	Yes	1.09/1.39
		adjunct	Yes	1.46/1.78
		relative clauses	Yes	1.82/1.79
Current study	relative clause	<i>whether</i>	Yes	0.67
		wh-	Yes	0.95
		adjunct	Yes	1.13
		relative clauses	Yes	0.88
		complex NP	Yes	1.13

Our results are only partially in line with previous experimental studies in Spanish (in particular, López Sancio 2015; Pañeda and Kush 2021; Pañeda et al. 2020). In line with López Sancio (2015), the only study who tested islands out of relative clauses, we also found an interaction between ‘Length’ and ‘Structure’ for extractions from embedded *wh*-questions, complex NPs, and adjuncts, which confirms that those structures should be considered islands in Spanish. In addition, we also found island effects for *whether* clauses and relative clauses (not tested in López Sancio 2015). For the comparison between the current study and other studies, which tested only *wh*-extractions, Pañeda et al. (2020) found island effects for extractions out of adjuncts and *whether* clauses, but not for complex NPs, and Pañeda and Kush (2021) found island effects for *wh*-, adjunct and relative clause islands, but not for *whether* islands.

The discrepancies between different studies call for more rigorous and systematic experimental investigations. An important difference between the current study and previous studies is that we used the same baseline condition for all island types, and within each item-set, we also tried to maintain similarities across all conditions (while keeping the critical differences). This designed feature allowed us to have a more consistent standard of comparison across island types when the island-hood question was being tested, and a shared baseline also made it easier to compare the strength of island effect across different island types.

For instance, extractions out of relative clauses, which were only tested in López Sancio (2015), showed a DD score of 2.11 for complex NP islands, whereas our DD score was 1.13. This led to different conclusions with regard to the strength of the island effects. Upon a closer look, however, there are important differences in the stimuli used to test these constructions. For example, the island and non-island conditions from López Sancio (2015) are not minimally different, as shown in (23) below:

- (23) a. Conozco **al entrenador** con el que el jugador
 I.know the trainer with that the player
 oyó los rumores sobre. que Ana está saliendo __.
 heard the rumors about that Ana is going.out
 ‘I know **the trainer** with whom the player heard the rumors that Ana was going out __.’
- b. Conozco **al entrenador** con el que dice el jugador
 I.know DOM.the trainer with the that says the player
 que sale Ana __.
 that goes.out Ana
 ‘I know **the trainer** with whom the player says that Ana goes out __.’
 (Examples adapted from López Sancio 2015)

In particular, the first embedded clause didn't contain the same verb nor the same tense (i.e. *oyó* 'heard' vs. *dice* 'says'), the embedded subjects were not in the same position with respect to the verb (i.e. preverbal vs. postverbal), the second embedded verb was not in the same tense/aspect (i.e. *sale* 'dates' vs. *está saliendo* 'is dating'), and the subjects of these second embedded verbs were also not in the same position with respect to the verb. Similar issues could be raised for the other island conditions tested. These differences might be responsible for showing a bigger difference between the baseline and the island. On the other hand, our stimuli compared minimally different sentences, as the examples in (16b)–(19b), repeated below in (24), which allowed for testing only the effect of the complex NP:

- (24) a. Sonia vio **al** **profesor** que Ana escuchó que Bruno
 Sonia saw DOM.the professor that Ana heard that Bruno
 contrató __.
 hired
 'Sonia saw the professor that Ana heard that Bruno hired.'
- b. Sonia vio **al** **profesor** que Ana escuchó el rumor
 Sonia saw DOM.the professor that Ana heard the rumor
 de que Bruno contrató __.
 of that Bruno hired
 'Sonia saw the professor that Ana heard the rumor that Bruno hired.'

Based on the analysis of DD scores, we found some limited differences in the size of island effects. In particular, we only found that *whether*-islands are different from adjunct, complex NP and *wh*-islands, but not relative clause islands. No other significant differences were found with regard to the strength of the island effects. Since the majority of the constructions we tested show island effects with similar strength, our findings seem to present a more categorical picture of island effects than some previous studies have suggested.⁸ More future research is clearly needed. But it is worth pointing out that in the current study, the judgments on the island constructions are always compared with their matched baseline conditions. This is important because a direct comparison between island constructions may lead to erroneous conclusions about island strength. For instance, while extraction out of a relative clause island gives rise to lower acceptability judgments, the strength of the island effect (measured by the DD scores) is not greater than the other islands. Looking at

⁸ As an anonymous reviewer points out, there are a number of factors that might have contributed to our results, which don't show clear gradient effects, such as, the fact that relative clause dependencies are known to show smaller gradient effects than *wh*-dependencies and the fact that it's not possible to manipulate D-linking in relative clause dependencies. The exploration of these factors exceeds the aims of this paper.

only the acceptability ratings, therefore, could lead to the ‘illusion’ that some islands show a stronger degradation than others. When careful statistical comparisons with a baseline are carried out, it’s possible to observe that the islands that showed the mildest degradation in acceptability rating (i.e. extraction out of *whether* islands, with a mean rating of 3.48) and the islands that shows the strongest degradation (i.e. extraction out of relative clause islands, with a mean rating of 1.68) don’t actually differ in the strength of their island effect.

As an anonymous reviewer points out, although we only found evidence for a very limited gradient effects, there is evidence of a more general gradient effects in islands in the literature (see, e.g. Atkinson et al. 2016; Goodall 2015; Kush et al. 2019; Sprouse et al. 2016; Villata et al. 2016, among many others). We acknowledge that a more detailed theory of gradience is needed in order to capture the complete set of observations across multiple constructions in different languages. It may be possible that the categorical distinction could be maintained in the grammar, and the observed gradience would arise from extra-grammatical constraints. It is also possible that the grammar-internal constraints are applied in a probabilistic fashion depending on the context, leading to gradient outcomes. The current study remains inconclusive on this question.

In conclusion, we investigated a variety of island constructions in the context of relative clauses in Spanish. All the constructions tested here are shown to be islands in Spanish. Furthermore, we did not find an across-the-board gradient effect in terms of the island strength, although *whether* islands are shown to be weaker than other types of islands. The current study made use of a factorial design to examine the island status of different constructions and their respective island strength, minimizing the risk of making erroneous conclusions based on only the acceptability rating of the island-violation target. In this respect, This work contributes to a growing body of work that highlights the role of formal experimental techniques in linguistic theory.

Acknowledgements: We wish to thank the anonymous *Probus* reviewers for their insightful comments and suggestions. We are also grateful to the audiences at the Language, Evolution, Acquisition and Processing Workshop at the University of Chicago and the 50th Linguistic Symposium on Romance Languages for feedback and discussion. Thanks, also, to Eszter Ronai for all her help. All mistakes and shortcomings are our own.

References

- Abeillé, Anne, Barbara Hemforth, Elodie Winckel & Edward Gibson. 2020. Extraction from subjects: Differences in acceptability depend on the discourse function of the construction. *Cognition* 204. 104293.
- Almeida, Diogo. 2014. Subliminal wh-islands in Brazilian Portuguese and the consequences for syntactic theory. *Revista da ABRALIN* 13. <https://doi.org/10.5380/rabl.v13i2.39611>.
- Atkinson, Emily, Aaron Apple, Kyle Rawlins & Akira Omaki. 2016. Similarity of wh-phrases and acceptability variation in wh-islands. *Frontiers in Psychology* 6. 2048.
- Boeckx, Cedric. 2012. *Syntactic islands*. Cambridge: Cambridge University Press.
- Chomsky, Noam. 1973. Conditions on transformations. In Stephen R. Anderson & Paul Kiparsky (eds.), *A festschrift for Morris Halle*, 232–286. New York: Holt, Rinehart and Winston.
- Chomsky, Noam. 1986. *Barriers*, vol. 13. Cambridge, Massachusetts: MIT Press.
- Cinque, Guglielmo. 1990. *Types of \bar{a} -dependencies*. Cambridge, Massachusetts: MIT Press.
- Drummond, Alex. 2013. Ibx farm. Online server. Available at: <http://spellout.net/ibxfarm>.
- Erteschik-Shir, Nomi. 1973. *On the nature of island constraints*. Massachusetts Institute of Technology Doctoral dissertation.
- Fábregas, Antonio. 2013. Nota sobre unas construcciones que temblaría la gramática si fueran extracciones de isla. *Signo y seña* 24. 175–188.
- Fukuda, Shin & Jon Sprouse. 2019. Islandhood of Japanese complex NPs and the factorial definition of island effects. In *Poster presentation. The 27th Japanese Korean Linguistics Conference (JK 27)*. October 18–20. Seoul, South Korea: Sogang University.
- Gallego, Ángel J. 2011. Successive cyclicity, phases, and CED effects. *Studia Linguistica* 65. 32–69.
- Gallego, Ángel J., Juan Uriagereka, José Camacho, Nydia Flores-Ferrán, Liliana Sánchez, Viviane Déprez & María José Cabrera. 2007. *Sub-extraction from subjects: A phase theory account. Amsterdam studies in the theory and history of linguistic science series 4*, vol. 287. Amsterdam: John Benjamins Publishing.
- Goldberg, Adele E. 2013. *10 Backgrounded constituents cannot be “extracted”. Experimental syntax and island effects*. Cambridge: Cambridge University Press.
- Goodall, Grant. 2015. The d-linking effect on extraction from islands and non-islands. *Frontiers in Psychology* 5. 1493.
- Haegeman, Liliane, Ngel Jimnez-Fernández & Andrew Radford. 2014. Deconstructing the subject condition in terms of cumulative constraint violation. *The Linguistic Review* 31. 73150.
- Hofmeister, Philip, Laura Staum Casasanto, Ivan A Sag, Jon Sprouse & Norbert Hornstein. 2013. *Islands in the grammar? Standards of evidence. Experimental syntax and island effects*, 42.
- Hofmeister, Philip & Ivan A. Sag. 2010. Cognitive constraints and island effects. *Language* 86. 366.
- Hothorn, Torsten, Bretz Frank & Peter Westfall. 2008. Simultaneous inference in general parametric models. *Biometrical Journal* 50. 346–363.
- Huang, C-T James. 1982. *Logical relations in Chinese and the theory of grammar*. PhD diss.
- Jiménez Fernández, Ángel Luis. 2009. On the composite nature of subject islands: A phase-based approach. *SKY Journal of Linguistics* 22. 91–138.
- Keshev, Maayan & Aya Meltzer-Asscher. 2019. A processing-based account of subliminal wh-island effects. *Natural Language & Linguistic Theory* 37. 621–657.
- Kluender, Robert & Simone Gieselman. 2013. *What’s negative about negative islands? A re-evaluation of extraction from weak island contexts. Experimental syntax and island effects*, 186–207.

- Kluender, Robert & Marta Kutas. 1993. Subacency as a processing phenomenon. *Language & Cognitive Processes* 8. 573–633.
- Kush, Dave, Terje Lohndal & Jon Sprouse. 2018. Investigating variation in island effects: A case study of Norwegian wh-extraction. *Natural Language & Linguistic Theory* 36. 743–779.
- Kush, Dave, Terje Lohndal & Jon Sprouse. 2019. On the island sensitivity of topicalization in Norwegian: An experimental investigation. *Language* 95. 393–420.
- Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82. 1–26.
- López Sancio, Sergio. 2015. *Testing syntactic islands in Spanish*. Vitoria-Gasteiz: Universidad del País Vasco/Euskal Herriko Unibertsitatea master thesis.
- Maxwell, Scott E., Harold D. Delaney & Ken Kelley. 2017. *Designing experiments and analyzing data: A model comparison perspective*. New York, NY: Psychology Press.
- Michel, Daniel. 2014. *Individual cognitive measures and working memory accounts of syntactic island phenomena*. UC San Diego Doctoral dissertation.
- Ortega-Santos, Iván, Lara Reglero & Jon Franco. 2018. Wh-islands in L2 Spanish and L2 English: Between poverty of the stimulus and data assessment. *Fontes Linguae Vasconum* 126. 435–471.
- Pañeda, Claudia & Dave Kush. 2021. Spanish embedded question island effects revisited: An experimental study. *Linguistics*. <https://doi.org/10.1515/ling-2020-0110>.
- Pañeda, Claudia, Sol Lago, Elena Vares, João Veríssimo & Claudia Felser. 2020. Island effects in Spanish comprehension. *Glossa: a Journal of General Linguistics* 5(1). 21.
- Postal, Paul Martin. 1998. *Three investigations of extraction*, vol. 29. Cambridge, Massachusetts: MIT Press.
- R Core Team. 2013. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: <http://www.R-project.org/>.
- Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, Massachusetts: The MIT Press.
- Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, Mass.: MIT Doctoral Dissertation.
- Sprouse, Jon, Ivano Caponigro, Ciro Greco & Carlo Cecchetto. 2016. Experimental syntax and the variation of island effects in English and Italian. *Natural Language & Linguistic Theory* 34. 307–344.
- Sprouse, Jon, Matt Wagers & Colin Phillips. 2012. Working-memory capacity and island effects: A reminder of the issues and the facts. *Language* 88. 401–407.
- Stepanov, Arthur, Manca Mušič & Penka Stateva. 2018. Two (non-) islands in Slovenian: A study in experimental syntax. *Linguistics* 56. 435–476.
- Suñer, Margarita. 1991. Indirect questions and the structure of CP: Some consequences. In Héctor Campos & Fernando Martínez-Gil (eds.), *Current studies in Spanish linguistics*, 283–312. Washington, DC: Georgetown University Press.
- Szabolcsi, Anna. 2006. Strong vs. weak islands. In Everaert Martin & van Riemsdijk Henk (eds.), *The Blackwell companion to syntax*, vol. 4, 479–531. Malden, MA: Blackwell Publishing.
- Takami, Ken-ichi. 2012. *Preposition stranding*. Berlin: De Gruyter Mouton.
- Torrego, Esther. 1984. On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15. 103–129.
- Villata, Sandra, Luigi Rizzi & Julie Franck. 2016. Intervention effects and relativized minimality: New experimental evidence from graded judgments. *Lingua* 179. 76–96.