Claudia Pañeda* and Dave Kush

Spanish embedded question island effects revisited: an experimental study

Abstract: It is often reported that embedded questions (EQs) are not syntactic islands in Spanish. However, some authors have observed that the acceptability of filler-gap dependencies (FGDs) into Spanish EQs varies with the EQ-embedding verb: FGDs into EQs under *responsive* verbs (e.g., *know*) do not result in island effects, but FGDs into EQs under *rogative* verbs (e.g., *ask*) do yield island effects. One account attributes the contrast to a structural difference between the two EQs, due to which *ask*-EQs violate Bounding constraints, but *know*-EQs do not. In two acceptability studies we investigated the reliability of verb-dependent island effects in EQs introduced by *si* 'whether' and *cuándo* 'when'. We found no qualitative acceptability differences between *ask* and *know* EQ-island sentences, suggesting that the syntactic islandhood of Spanish EQs is not verb-dependent. Nevertheless, average island effects were numerically greater with *ask*, suggesting the presence of a non-syntactic constraint. In addition, FGDs into *whether*-EQs were generally acceptable, whereas FGDs into *when*-EQs obtained unacceptable average ratings and highly variable judgments. We argue that in neither case there is a Bounding constraint violation. Instead we explore alternative potential explanations for the differences in terms of features, presuppositions and processing pressures.

Keywords: acceptability judgments; *ask* verbs; *know* verbs; Spanish; *wh*-islands

*Corresponding author: Claudia Pañeda, Departamento de Filología Española, Universidad de Oviedo, Campus del Milán, C/Amparo Pedregal s/n, 33011, Oviedo, Spain,
E-mail: panedaclaudia@uniovi.es. https://orcid.org/0000-0002-1328-6605
Dave Kush, Department of Language Studies, University of Toronto Scarborough, Toronto, ON, Canada; and Department of Language and Literature, Faculty of Humanities, Norwegian University of Science and Technology, Trondheim, Norway, E-mail: dave.kush@utoronto.ca.
https://orcid.org/0000-0002-2567-6440

1 Introduction

Wh-questions involve dependencies between a filler (the *wh*-constituent) and a gap (the position where the *wh*-constituent is interpreted, represented by an underscore in [1]).

(1)   ***What*** *do you think that they brought __?*

Such filler-gap dependencies (FGDs) can span across several clauses, as in (1), but some structures, called *islands*, seem to block them (Ross 1967). Embedded questions (EQs) are often assumed to be islands in languages like English (Chomsky 1964): the attempt to establish a dependency between the *wh*-filler *what* and a gap inside the EQ gives rise to unacceptability or *island effects* as in (2).

(2)   ******What*** *do you wonder whether they brought __?*

The origin of these effects has been a point of considerable debate, both in EQs and in other island structures (e.g., Hofmeister and Sag 2010; Kluender and Kutas 1993; Phillips 2006; Sprouse et al. 2012; Yoshida et al. 2014). A recurrent question is whether they are *syntactic* in origin. The influential *Bounding* approach to island effects assumes that EQ-island effects arise when universal locality constraints on movement are violated (for different incarnations see Chomsky 1973, 1977, 1986, 2000). However, it is not always clear whether the source of observed unacceptability should be blamed on the syntax. For any given island effect it is, to some extent, an empirical question whether it is best explained by a syntactic or an extra-syntactic constraint. For instance, dependencies inside EQs could also fail for compositional or logical reasons, as argued by semantic accounts of islands (Abrusán 2014; Szabolcsi and Zwarts 1993), or because they overload working memory capacity, as proposed by processing accounts (Hofmeister and Sag 2010; Kluender and Kutas 1993).

Although the unacceptability of extraction from EQs is often assumed to follow from universal principles, the opacity of EQs to movement has been argued to be subject to significant cross-linguistic variation. One language in which EQs have been reported to allow extraction is Spanish, in cases like (3) (Suñer 1991; Torrego 1984).

(3)   *¿Qué   diccionario   no   sab-ías   si   Celia   hab-ía*
        what   dictionary   NEG   know-2SG.PST   whether   Celia   have-3SG.PST
        devuelto   __   ya?
        return.PTCP   __   yet
        'Which dictionary didn't you know whether Celia had returned __ yet?'
        (Torrego 1984: 115)

To accommodate cross-linguistic variation, it has been argued that parametric syntactic differences may allow some languages to circumvent the syntactic violations that FGDs into EQs incur in English (Suñer 1991; Torrego 1984; see also Rizzi 1982).

The simple parameterization approach predicts, all else equal, that Spanish should allow FGDs into all EQs. However, there is evidence that island effects arise with a select subset of EQs. Some authors (Suñer 1991; Torrego 1984) report that the acceptability of FGDs into EQs depends on the (semantic) class of the embedding predicate: FGDs are supposedly allowed into EQs selected by *responsive* verbs, but not into EQs selected by *rogative* verbs (terms after Lahiri 2002). Responsive verbs are verbs that can take both interrogative and declarative complements (e.g., *saber* 'to know' or *decir* 'to say' [4]). Rogative verbs can only take interrogative complements (e.g., *preguntar* 'to ask' or *preguntarse* 'to wonder', literally 'to ask oneself') [5], [6]).[1]

(4)   *¿A   cuántos   te   dij-eron   si   hab-ía*
       to   how.many   you.DAT say-3PL.PST   whether have-3SG.PST
       invit-ado   __   Carlos?
       invite-PTCP   __   Carlos
       'How many (people) did they tell you whether Carlos had invited __?'
       (Suñer 1991: 302)

(5)   **¿A   cuántos   pregunt-aron   que   si   hab-ía*
       to   how.many   ask-3PL.PST   that   whether have-3SG.PST
       invit-ado   __   Carlos?
       invite-PTCP   __   Carlos
       'How many (people) did they ask whether Carlos had invited __?'
       (Suñer 1991: 303)

(6)   **¿Qué   pregunt-aste   si   comer   __?*
       what   ask-2SG.PST   whether   eat.INF __
       'What did you ask whether to eat __?'
       (Suñer 1991: 304)

---

1 Complements to rogative verbs are sometimes preceded by the complementizer *que* ('that' [5]), but, according to Suñer (1991), unacceptability arises whether *que* is present (5) or not (6). See Villa-García (2015) and references therein for more information about this and other types of *que*.

In this paper we investigate (i) the reliability of verb-dependent EQ-island effects and (ii) what the source of such selective island effects might be. Specifically, we consider whether the verb-dependent effects are likely to reflect a violation of a syntactic Bounding constraint.

We address these questions with formal acceptability judgment studies, as part of a general program for identifying constraints and various sources of island effects (Sprouse et al. 2011 and subsequent work). We test sentences in context and with complex fillers, in order to reduce the burden of extra-syntactic factors on acceptability and tease them apart from syntactic factors. To preview our conclusions: Our results speak against a syntactic origin for verb-dependent selective EQ-island effects, particularly one based on Bounding-style locality constraints. The rest of the introduction provides an overview of the Bounding account and an overview of previous experiments.

1.1 Bounding and EQ-islands in Spanish

Bounding approaches to island effects (Chomsky 1973, 1977, 1986, 2000) are based on the idea that movement must be local. For instance, the principle of Subjacency (Chomsky 1973) bars movement in one fell swoop across more than one bounding node, i.e., S (currently TP) or NP/DP in English. Movement across several bounding nodes is only possible if it proceeds successive-cyclically, making intermediate stops at unfilled spec,CP positions. From this perspective, island effects arise because extractees are forced to cross more than one bounding node without making intermediate stops. This occurs in English *wh*-questions with extraction from EQs because the *wh*-word crosses two TPs and the intermediate spec, CP position is already filled by the *wh*-phrase or operator introducing the EQ (see Figure 1).

Sentences like (3) motivated the idea that EQ-island sensitivity could vary across languages (Suñer 1991; Torrego 1984). This variation was squared with Subjacency in languages like Spanish via the proposal that allowed bounding nodes to vary parametrically (Rizzi 1982; Torrego 1984). If S′ (currently CP) is a bounding node in Spanish, as opposed to S (TP) in English, A′-movement from EQs would only cross a single bounding node. Such movement would comply with the Subjacency condition.

As discussed above Torrego (1984) and Suñer (1991) report that some Spanish EQs nevertheless disallow extraction: FGDs into EQs under rogative verbs like *ask* or *wonder* are supposedly unacceptable (5), (6). Suñer (1991) explains the supposed islandhood of rogative-EQs as a Bounding/Subjacency effect: EQs embedded under rogative verbs contain additional functional structure in their left
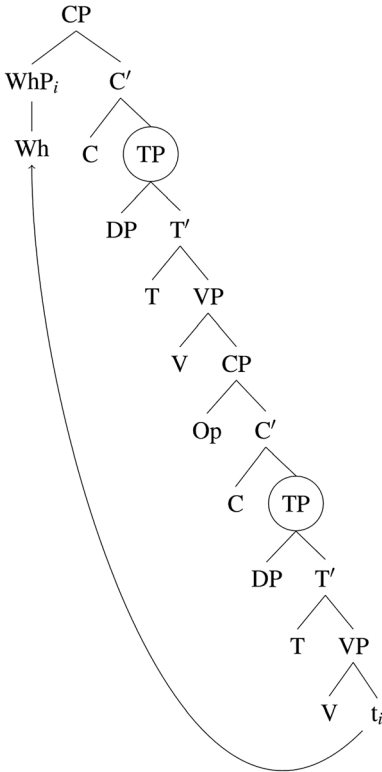
Figure 1: Structure of an English *wh*-question with extraction from EQs. Bounding nodes are circled.

periphery, which blocks successive-cyclic movement out (Figure 2). Specifically, Suñer argues that EQs under rogative verbs have an additional CP layer (labeled CP in Figure 2).

Evidence for this additional CP layer comes from the fact that Spanish questions embedded under rogative verbs can be preceded by the complementizer *que* 'that' in (7), whereas questions embedded under responsive verbs cannot as in (8).

(7)   *Me   pregunt-aron (que) a   quién invit-arás*
        me   ask-3PL.PST   that to who   invite-2SG.FUT
        tú   al   concierto.
        you   to.the   concert
        'They asked me who you will invite to the concert.'
        (Suñer 1991: 283)
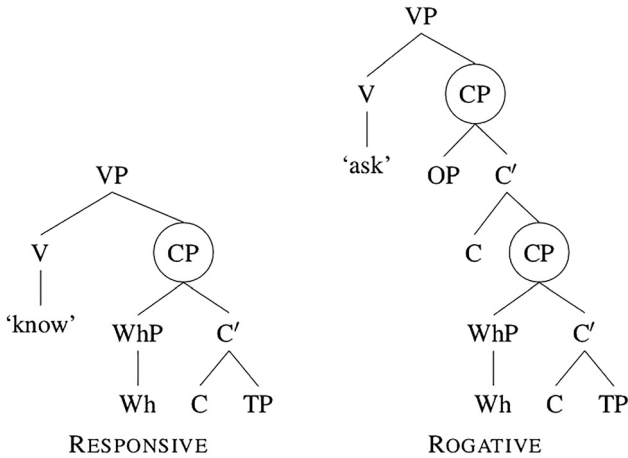
Figure 2: Structure of Spanish embedded questions under responsive and rogative verbs, according to Suñer (1991). The bounding nodes crossed by an extracted *wh*-word are circled.

(8) *Juana*   *no*   *sab-ía*   *(*que)*   *cuándo*   *visit-aría*
     Juana   NEG   know-3SG.PST   that   when   visit-3SG.COND
     a   *sus*   *abuelos.*
     to   POSS.3 grandparents
     'Juana didn't know when she would visit her grandparents.'
     (Suñer 1991: 284)

In EQs under rogative verbs, both spec,CP positions are assumed to be filled: the lower spec,CP contains the *wh*-phrase/operator introducing the EQ, and the higher spec,CP contains a null question operator, which Suñer assumes is required to mark the double CP structure as an indirect question. Moving out of a rogative EQ should therefore cross two CPs. Since CP is considered a bounding node in Spanish, extraction from these EQs violates Bounding constraints.

Thus, according to Suñer (1991), EQs under rogative verbs are *structural/ syntactic* islands, while EQs under responsive verbs are not. This predicts that extraction from EQs under rogative verbs should yield structural island effects, typically associated with stark unacceptability, whereas extraction from EQs under responsive verbs should yield no structural island effects and should obtain acceptable ratings. Her judgments suggest that this is the case in (4), (5), (6), but this contrast has never been tested experimentally.

1.2 Recent experimental work

Two recent experimental studies tested extraction from Spanish *si* 'whether' EQs embedded under the verb *preguntar(se)* 'to ask/wonder' (López-Sancio 2015;

Pañeda et al. 2020). Both studies compared the acceptability of sentences with FGDs in a 2×2 factorial design (discussed in greater detail below). An example set of test sentences (taken from López-Sancio 2015) is in (9).

(9) (a) *¿Quién*   __   *piens-a*   *que*   *Rocío*   *vio*
          who   __   think-3SG.PRS that   Rocío   see.3SG.PST
          el   *mensaje?*
          the   message
          'Who __ thinks that Rocío saw the message?'

     (b) *¿Quién*   __   *se*   *pregunt-a*   *si*   *Rocío*   *vio*
          who   __   REFL   ask-3SG.PRS   whether   Rocío   see.3SG.PST
          el   *mensaje?*
          the   message
          'Who __ wonders whether Rocío saw the message?'

     (c) *¿Qué*   *piens-as*   *que*   *vio*   __   *Rocío?*
          what   think-2SG.PRS   that   see.3SG.PST   __   Rocío
          'What do you think that Rocío saw __?'

     (d) *¿Qué*   te   *pregunt-as*   *si*   *Rocío*   *vio*   __?
          what   REFL   ask-2SG.PRS   whether   Rocío   see.3SG.PST   __
          'What do you wonder whether Rocío saw __?'
          (López-Sancio 2015: 10)

Both studies found *whether* EQ-island effects: Participants rejected (9d) at significantly higher rates than the other sentences in the paradigm – and at rates analogous to other less controversial syntactic island violations such as extractions from adjuncts.[2]

The results of López-Sancio (2015) and Pañeda et al. (2020) are compatible with the claim that EQs under rogative verbs are syntactic islands, though they do not establish that island effects are verb-dependent, since they did not compare responsive and rogative verbs.

Further, these studies do not address the source of the island effects observed. The results are compatible with a syntactic source, e.g., they could result from a syntactic Bounding constraint violation. They are alternatively compatible with extra-syntactic sources of unacceptability, such as semantic or pragmatic constraint violations or processing difficulty.

---

2 In fact, Spanish participants' judgments in these experiments look quite similar to the judgment patterns of English speakers in studies with similar designs: island sentences (9d) were generally less acceptable than the other conditions and they obtained unacceptable ratings (Almeida 2014; Michel 2014; Ortega-Santos et al. 2018; Sprouse et al. 2011, 2012, 2016).

1.3 The present study

In this paper we test for effects of verb choice, in order to determine whether extraction from Spanish EQs is influenced by the type of matrix verb (*ask* vs. *know*). We carried out two acceptability judgment experiments. Experiment 1 tests extraction from *si* ('whether') EQs, in order to follow up on previous experimental findings that extraction from *whether* EQs under rogative verbs is unacceptable (López-Sancio 2015; Pañeda et al. 2020). Experiment 2 examines whether the results extend to a different interrogative, *cuándo* ('when').

To assess whether effects of verb choice are likely to reflect a structural difference, we minimize the influence of non-structural factors on acceptability by presenting test sentences in a context designed to make them semantically and pragmatically motivated, and by reducing processing difficulty as much as possible. It has been proposed that sentences that are rejected due to non-structural constraints become more acceptable in context, whereas sentences which are rejected due to structural constraints do not (Erteschik-Shir 2006; Keller 2000; Sorace and Keller 2005).

With the same aim, we also examine judgment distributions. Kush et al. (2018, 2019) have recently suggested that judgment distributions can provide useful information about whether a domain is necessarily a structural island. The authors observed that small island effects can arise from variable/inconsistent judgments across trials. Kush and colleagues argue that if participants accept, on a significant proportion of trials, sentences with dependencies out of a supposed island domain, X, then one can minimally conclude that there exists a non-island structural analysis of domain X. Residual variability in judgments on such trials could either reflect that participants occasionally fail to adopt the non-island analysis of X or that extra-syntactic factors conspire to reduce acceptability. Applying this logic to EQs, we make the following predictions: If EQs under rogative verbs are necessarily analyzed as structural islands, extraction from EQs under rogative verbs should be consistently judged unacceptable and we expect to see large island effects. If, on the other hand, EQs under rogative verbs are not structural islands we either expect to see (i) no island effects at all, or (ii) smaller island effects characterized by inconsistent judgment distributions. In the last case we could conclude that there is a non-island analysis of EQs under rogative verbs (i.e., there is not a categorical syntactic ban on extraction).

As a control, we test extraction from adjuncts and relative clauses under the same conditions. Adjunct and relative clause island effects are generally attributed to a structural constraint violation (e.g., Chomsky 1973, 1977, 1986, 2001; Huang 1982; Nunes and Uriagereka 2002), so we expect them to yield unacceptability and

large island effects. EQs should show similar effects if they are necessarily analyzed as structural islands, but they should behave differently if they are not.

1.4 General experimental design: the factorial definition of islands

Our experiments adopt a version of the experimental paradigm introduced by Sprouse et al. (2011) and which has subsequently been used to collect formal acceptability judgments of island sentences in various languages (Sprouse et al. 2012, 2016; Aldosari 2015; Almeida 2014; Keshev and Meltzer-Asscher 2019; Kush et al. 2019, 2018; López-Sancio 2015; Michel 2014; Ortega-Santos et al. 2018; Pañeda et al. 2020; Stepanov et al. 2018; Tucker et al. 2019). This design isolates island effects from two factors that can affect acceptability, namely the *distance* between the filler and the gap (short/long) and the type of *structure* where the gap is located (non-island/island). An example from Sprouse et al. (2016) is shown in (10).

(10) (a) Non-island/short: *Who __ thinks that John bought a car?*
       (b) Island/short: *Who __ wonders whether John bought a car?*
       (c) Non-island/long: *What do you think that John bought __?*
       (d) Island/long: *What do you wonder whether John bought __?*
            Differences-in-differences (DD) score: ((10d) – (10b)) – ((10c) – (10a))

In the short conditions (10a), (10b), the gap is in the matrix sentence, whereas in the long conditions (10c), (10d), the gap is in the embedded clause. In the non-island conditions (10a), (10c), the embedded structure is a declarative clause, which does not give rise to island effects, whereas in the island conditions (10b), (10d), it is an EQ, expected to yield island effects. Sentences with extraction from an island are represented by condition *island/long* (10d), in which a long-distance dependency is established inside an EQ.

The design in (10) quantifies the independent effects of a long-distance dependency and an EQ, and it can thus be used to assess to what extent these two factors contribute to the (un)-acceptability of the island/long condition. If the island/long condition is as (un)-acceptable as predicted by the sum of the effects of the two factors, there is no evidence that EQs yield island effects, understood as specific effects caused by establishing a long-distance dependency inside an EQ. However, if the island/long condition is less acceptable than predicted by the addition of the effects of distance and the EQ structure (i.e., if there is "super-additivity"), this is taken as an indication that there are island effects, thus supporting the need to assume additional constraints (syntactic or otherwise) to explain them. An island effect is represented statistically by an interaction

between distance and structure. Island effect sizes can be computed with differences-in-differences (DD) scores as shown in (10).[3]

In our experiments, we deployed a 3×2 version of the design in (10) to address the impact of the type of matrix verb. Following most previous work on Spanish and English, our materials involved object extraction and a *wh*-question configuration.

2 Experiment 1

2.1 Method

2.1.1 Participants

52 native speakers of European Spanish were recruited through social media and among students of the University of Oviedo (Spain). Four were excluded because their mean rating in the ungrammatical fillers was 3.5 (the midpoint of the rating scale) or higher. In addition, 1 participant was excluded due to failures in data recording. The remaining 47 participants had a mean age of 26.5 years (range: 18–57) and no self-reported language impairments. Thirty four were female. Most participants were born (33) and/or lived (38) in the northern Spanish region Asturias at the time of testing. Nine reported knowledge of other Romance languages spoken in Spain in addition to Spanish (Asturian: 2, Galician: 1, Catalan: 6). Two Amazon vouchers in value of 30 euros were raffled off among the participants as a reward. The experiment was conducted in accordance with the Declaration of Helsinki.

2.1.2 Materials

All test items were *wh*-questions, and they contained an embedded clause. We created 18 experimental items to test EQ islands following a 3×2 design (11). All experimental materials are available at the Open Science Framework repository (https://osf.io/enq4r/?view_only=ae50dc9a113f4e93a0e52dab44d25c1f).

---

3 Note that the design not only isolates the island effect from the independent contributions of distance and structure, but it also factors out the impact of specific lexical items or other influencing factors that are held constant across conditions. For instance, the island effect measure in a case like (10) is not influenced by the frequency of *bought*, because *bought* appears in all the conditions. Similarly, the effects of extracting an object as compared to a subject are factored out by subtracting the non-island conditions from the island conditions.

(11)   Sample *whether* EQ-island sentence set

    (a)   Non-island/short

       Context sentence:

       Funcionario   de   Correos: "*Cre-o   que   hab-éis*
       officer   of   post   believe-1SG.PRS   that   have-2PL.PRS
       recog-ido   el   paquete."
       pick.up-PTCP   the   package
       'Post officer: "I believe that you have picked up the package."'

       Experimental sentence:

       ¿Qué   funcionario __ *pens-aba   que   hab-íamos*
       what   officer   __   think-3SG.PST   that   have-1PL.PST
       recog-ido   el   paquete?
       pick.up-PTCP   the   package
       'Which officer __ thought that we had picked up the package?'

    (b)   *Know*-island/short

       Context sentence:

       Funcionario de   Correos: "No   teng-o   claro
       officer   of   post   NEG   have-1SG.PRS   clear
       si   hab-éis   recog-ido   el   paquete."
       whether   have-2PL.PRS   pick.up-PTCP   the   package
       'Post officer: "It is not clear to me whether you have picked up the package."'

       Experimental sentence:

       ¿Qué   funcionario __ *no   sab-ía   si*
       what   officer   __   NEG   know-3SG.PST   whether
       hab-íamos   recog-ido   el   paquete?
       have-1PL.PST pick.up-PTCP   the   package
       'Which officer __ didn't know whether we had picked up the package?'

    (c)   *Ask*-island/short

       Context sentence:

       Funcionario   de   Correos: "¿Hab-éis   recog-ido
       officer   of   post   have-2PL.PRS   pick.up-PTCP
       el   paquete?"
       the   package
       'Post officer: "Have you picked up the package?"'

       Experimental sentence:

       ¿Qué   funcionario __ *pregunt-ó   si   hab-íamos*
       what   officer   __   ask-3SG.PST   whether   have-1PL.PST

recog-ido   el   paquete?
pick.up-PTCP   the   package
'Which officer __ asked whether we had picked up the package?'

(d)   Non-island/long
     Context sentence:
     Funcionario: "Cre-o   que   hab-éis   recog-ido
     officer   believe-1SG.PRS that   have-2PL.PRS   pick.up-PTCP
     el   paquete   de   Amazon."
     the   package   of   Amazon
     'Post officer: "I believe that you have picked up the Amazon package."'
     Experimental sentence:
     ***¿Qué   paquete**   pens-aba   el   funcionario   que*
     what   package   think-3SG.PST the   officer   that
     hab-íamos   recog-ido   __?
     have-1PL.PST pick.up-PTCP   __
     'Which package did the officer think that we had picked up __?'

(e)   *Know*-island/long
     Context sentence:
     Funcionario: "No teng-o   claro   si
     officer   NEG have-1SG.PRS clear   whether
     hab-éis   recog-ido   el   paquete   de Amazon."
     have-2PL.PRS pick.up-PTCP   the   package   of Amazon
     'Post officer: "It is not clear to me whether you have picked up the
     Amazon package."'
     Experimental sentence:
     ***¿Qué   paquete**   no sab-ía   el   funcionario si*
     what   package   NEG know-3SG.PST the officer   whether
     hab-íamos   recog-ido   __?
     have-1PL.PST pick.up-PTCP   __
     'Which package didn't the officer know whether we had picked up __?'

(f)   *Ask*-island/long
     Context sentence:
     Funcionario: "¿Hab-éis   recog-ido   el   paquete
     officer   have-2PL.PRS   pick.up-PTCP   the   package
     de   Amazon?"
     of   Amazon
     'Post officer: "Have you picked up the Amazon package?"'
     Experimental sentence:
     ***¿Qué   paquete**   pregunt-ó   el   funcionario*
     what   package   ask-3SG.PST   the   officer

si	*hab-íamos*	*recog-ido*	*__?*
whether	have-1PL.PST	pick.up-PTCP	__

'Which package did the officer ask whether we had picked up __?'

This design was similar to (10) in that it manipulated the two-level factor Distance of a moved-*wh* word (short, long) and the Structure (the type of embedded clause). Our design differed from previous experiments in that Structure had three levels: the embedded clause was either (i) a non-island declarative complement embedded under the matrix verb *pensar* 'to think' and introduced by the complementizer *que* 'that', (ii) an EQ under the responsive matrix verb *saber* 'know', or (iii) an EQ under the rogative matrix verb *preguntar* 'ask'. Both island conditions embedded a *si* 'whether' EQ.

In addition, all *wh*-fillers were complex (e.g., *which package*). Complex fillers ameliorate extraction from EQs (Goodall 2015; Sprouse et al. 2016). This effect is predicted under syntactic (Pesetsky 1987; Rizzi 1990), semantic (Szabolcsi and Zwarts 1993), and processing accounts of islands (Hofmeister and Sag 2010), so we used complex fillers to minimize the burden across multiple different factors.

Furthermore, experimental items were presented in context to make the question pragmatically motivated. In all cases, the context consisted of a noun phrase referring to a speaker, followed by a sentence (presented as an utterance) that expressed a belief (in non-island conditions), a doubt (in *know*-EQs) or a question (in *ask*-EQs) about the event referred to in the experimental question. For instance, if the experimental question contained information about someone picking up a package, as in (11), this information was already present in the context. In non-island and *know*-EQ conditions, the context sentence had a similar structure as the experimental sentence, with a matrix and an embedded verb, but the matrix verb was never the same in the context and the experimental sentence. Another feature of the context was that the noun phrase corresponding to the question word in the test item was restricted (e.g., the *post officer* rather than the *officer* in the short conditions, the *Amazon package* rather than the *package* in the long conditions). This was meant to introduce the implicature that there were other members of the class, thereby motivating the question about a specific member of that class (e.g., *which officer*, *which package*).

We also created materials to test the acceptability of similar *wh*-movement dependencies from conditional adjuncts (12) and relative clauses (13) as controls. These followed the 2×2 design in (10). As in the EQ-island items, all sentences were questions with complex fillers, and they were presented in a similar context. Eight test items per island were created.

(12) Sample adjunct island sentence set

    (a) Non-island/short

        Context sentence:

Presidente	*de*	*la*	*comisión*	*presupuestaria:*	*"La*
president	of	the	commission	budgetary	the

ministra	*de Sanidad*	*desea-ría*	*que*
minister	of health	wish-3SG.COND	that

redujér-amos	*el*	*presupuesto."*
reduce-1PL.SBJV	the	budget

        'President of the budget commission: "The Health minister would like that we reduce the budget."'

        Experimental sentence:

¿Qué	***ministra***	__	*desea-ría*	*que*	*redujér-amos*
what	minister	__	wish-3SG.COND	that	reduce-1PL.SBJV

el	*presupuesto?*
the	budget

        'Which minister __ would like that we reduce the budget?'

    (b) Island/short

        Context sentence:

Presidente	*de*	*la*	*comisión*	*presupuestaria:*	*"La*
president	of	the	commission	budgetary	the

ministra	*de Sanidad*	*se*	*queja-ría*
minister	of health	REFL	complain-3SG.COND

si	*redujér-amos*	*el*	*presupuesto."*
if	reduce-1PL.SBJV	the	budget

        'President of the budget commission: "The Health minister would complain if we reduced the budget."'

        Experimental sentence:

¿Qué	***ministra***	__	*se*	*queja-ría*	*si*
what	minister	__	REFL	complain-3SG.COND	if

redujér-amos	*el*	*presupuesto?*
reduce-1PL.SBJV	the	budget

        'Which minister __ would complain if we reduced the budget?'

    (c) Non-island/long

        Context sentence:

Presidente	*de*	*la*	*comisión*	*presupuestaria:*	*"La*
president	of	the	commission	budgetary	the

ministra	*de*	*Sanidad*	*desea-ría*	*que*
minister	of	health	wish-3SG.COND	that

redujér-amos   *el*   *presupuesto*   *de*   *los*   *hospitales."*
reduce-1PL.SBJV   the   budget   of   the   hospitals
'President of the budget commission: "The Health minister would like that we reduce the hospitals budget."'
Experimental sentence:
¿Qué **presupuesto** *desea-ría*   *la*   *ministra*
what   budget   wish-3SG.COND   the   minister
que   *redujér-amos*   __?
that   reduce-1PL.SBJV   __
'Which budget would the minister like that we reduce __?'

(d)   Island/long
Context sentence:
Presidente de la   *comisión*   *presupuestaria: "La ministra*
president   of   the   commission   budgetary   the   minister
de Sanidad se   *queja-ría*   *si*   *redujér-amos*
of   health   REFL complain-3SG.COND   if   reduce-1PL.SBJV
el   *presupuesto*   *de*   *los*   *hospitales."*
the   budget   of   the   hospitals
'President of the budget commission: "The Health minister would complain if we reduced the hospitals budget."'
Experimental sentence:
¿Qué   **presupuesto**   *se*   *queja-ría*   *la*   *ministra*
what   budget   REFL   complain-3SG.COND   the   minister
si   *redujér-amos*   __?
if   reduce-1PL.SBJV   __
'Which budget would the minister complain if we reduced __?'

(13)   Sample relative clause island sentence set
Context sentence (the same in all conditions):
Jefe   *de la*   *empresa*   *Tecnologías*   *LCA: "En*   *la*
chief   of   the   company   technologies   LCA   in   the
reunión   *de*   *hoy*   *algunos*   *de*   *mis*   *empleados*
meeting   of   today   some   of   my   employees
han   *present-ado*   *el*   *proyecto de-l*
have.3SG.PRS   present-PTCP   the   project   of-the
año   *pasado en*   *lugar*   *de-l*   *de*   *este*   *año."*
year   past   in   place   of-the   of   this   year
'Chief of the company Technologies LCA: "In today's meeting some of my employees have presented last year's project rather than this year's."'

(a) Non-island/short

*¿**Qué** **jefe*** __ *dij-o*   *que unos empleados*
what chief __ say-3SG.PST that some employees
hab-ían   *present-ado el proyecto de-l año pasado?*
have-3SG.PST present-PTCP the project of-the year past
'Which chief __ said that some employees had presented last year's project?'

(b) Island/short

*¿**Qué** **jefe*** __ *ten-ía*   *unos empleados que*
what chief __ have-3SG.PST some employees that
hab-ían   *present-ado el proyecto de-l año pasado?*
have-3SG.PST present-PTCP the project of-the year past
'Which chief __ had some employees that had presented last year's project?'

(c) Non-island/long

*¿**Qué** **proyecto** dij-o*   *el jefe que unos*
what project say-3SG.PST the chief that some
empleados hab-ían   *present-ado __?*
employees have-3SG.PST present-PTCP __
'Which project did the chief say that some employees had presented __?'

(d) Island/long

*¿**Qué** **proyecto** ten-ía*   *el jefe unos empleados*
what project have-3SG.PST the chief some employees
que hab-ían   *present-ado __?*
that have-3SG.PST present-PTCP __
'Which project did the chief have some employees that had presented __?'

2.1.3 Procedure and analysis

The experiment was an acceptability judgment task run on IbexFarm (Drummond 2013). Participants accessed the experiment through a link on their personal computers. After completing a demographic questionnaire they were presented with instructions and example sentences. Participants were instructed to judge the acceptability of example sentences according to their colloquial understanding,

and were discouraged from basing their judgments on school grammar, orthography or plausibility. Finally, they were given 5 practice items before the beginning of the task.

Participants rated sentences on a 7-point Likert scale. Sentences were presented at the center of the screen together with seven numbered buttons depicting the scale with the indications *totalmente inaceptable* 'totally unacceptable' and *totalmente aceptable* 'totally acceptable' next to numbers 1 and 7. Participants could either click on the numbers or press the corresponding keys in the keyboard. Context sentences were shown in italics above the experimental sentences. Participants were reminded both in the instructions and with a note beneath the scale to judge the second rather than the first sentence.

Experimental and control sentences were intermixed with 16 ungrammatical fillers and 2 grammatical fillers to have a 1:1 ratio of acceptable to unacceptable sentences (assuming that sentences with extraction from EQs, adjuncts and relative clauses are unacceptable). The order of presentation was pseudo-randomized by participant, such that the same condition was never shown in consecutive trials. Experimental items were distributed across six Latin square lists to ensure that each participant saw each experimental item only in one condition.

Acceptability judgments were *z*-score transformed by participant and analyzed using linear mixed-effects models implemented using the lme4 (Bates et al. 2015) and lmerTest packages (Kuznetsova et al. 2017) in R (R Core Team 2020). The control items were analyzed with a 2×2 model and the experimental items were analyzed with a 3×2 model. Both tested for main effects of Distance and the island Structure on acceptability, as well as their interaction, taking condition non-island/short as a baseline. In addition, both included random intercepts for participants and items. In the 3×2 model, the three-way factor Structure was coded into two contrasts: one for non-island vs. the aggregate of all islands and another one for *ask* vs. *know*.

2.2 Results

The grammatical fillers were rated in the acceptable range ($z = 0.80$, *SD*: 0.36) and the ungrammatical fillers were rated in the unacceptable range ($z = -1.08$, *SD*: 0.56), showing that participants understood the task and carried it out as expected.

Acceptability by condition and DD scores in the adjunct, relative clause and EQ-islands are shown in Figure 3. Statistical results are reported in Table 1.
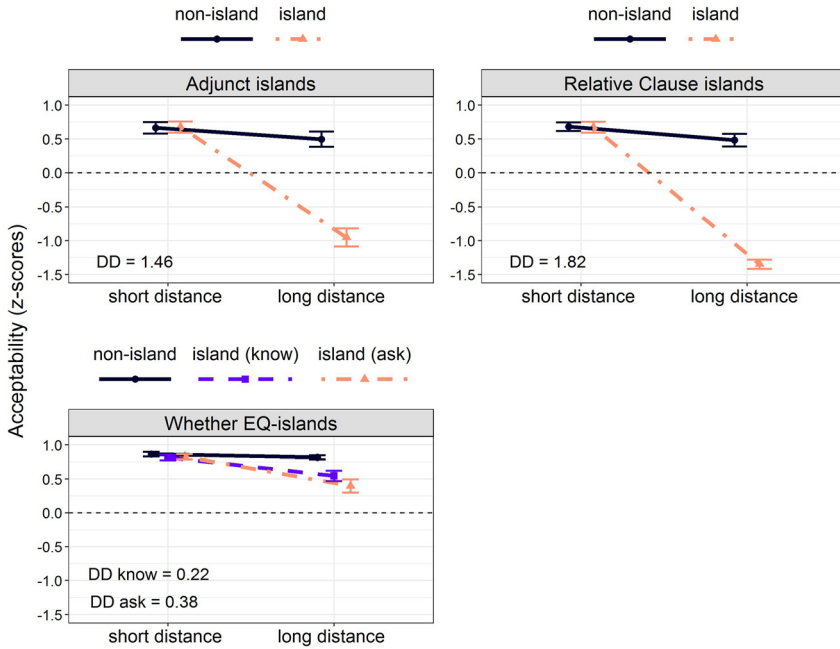
Figure 3: Acceptability by condition in the adjunct, relative clause and embedded question (EQ) islands of Experiment 1. Error bars show 95% confidence intervals. Non-parallel lines are indicative of an island effect.

Table 1: Results of linear mixed models of the island data in Experiment 1.

	β	SE	t	p	
Adjunct islands					
Distance	−0.165	0.074	−2.245	0.026	*
Structure	0.043	0.073	0.587	0.558	
Distance × Structure	−1.521	0.104	−14.616	< 0.001	***
Relative Clause islands					
Distance	−0.197	0.057	−3.471	< 0.001	***
Structure	−0.006	0.057	−0.106	0.916	
Distance × Structure	−1.827	0.080	−22.788	< 0.001	***
***Whether* EQ islands**					
Structure	−0.090	0.073	−1.229	0.219	
Ask-EQ vs. *know*-EQ	0.009	0.042	0.203	0.839	
Distance	−0.049	0.042	−1.172	0.241	
Distance × Structure	−0.599	0.103	−5.812	< 0.001	***
Distance × *ask*-EQ vs. Distance × *know*-EQ	−0.159	0.059	−2.665	0.008	**

We found significant super-additive Distance × Structure interactions both in adjunct (β = −1.521, SE = 0.104, p < 0.001) and relative clause islands (β = −1.827, SE = 0.080, p < 0.001), consistent with island effects. In both cases, island effects were large (adjunct islands: DD = 1.46; relative clause islands: DD = 1.82) and characterized by unacceptability in the island/long conditions (adjunct islands: −0.95; relative clause islands: −1.34).

EQs also yielded a super-additive Distance × Structure interaction, indicating that extraction from them was less acceptable on average than extraction from embedded declarative clauses (β = −0.599, SE = 0.103, p < 0.001). However, these island effects were smaller than adjunct and relative clause island effects (*know*-islands: DD = 0.22, *ask*-islands: DD = 0.38). The smaller interaction effects reflect the fact that ratings of the island/long conditions were well above z = 0, the midpoint of the scale, arguably in the acceptable range (*know*-island/long: 0.55, *ask*-island/long: 0.40). These ratings were lower than ratings to the other (uncontroversially grammatical) EQ-island conditions, but similar to ratings to the (also grammatical) adjunct and relative clause non-island/long conditions (adjunct: 0.50, relative clause: 0.48). Regarding the effects of verb choice, the interaction was significantly greater with *ask* than with *know* (β = −0.159, SE = 0.059, p = 0.008), indicating that island effects were greater in the former case. However, the verb did not cause a qualitative difference in acceptability.

The DD scores and condition means thus indicate that extraction from EQs is less acceptable *on average* than extraction from matched declarative clauses, but more acceptable than extraction from adjuncts and relative clauses. This could indicate that EQ-island effects are "subliminal" in Spanish, i.e., consistently mild and in the acceptable range (Almeida 2014). However, aggregate means alone are not enough to determine whether an effect is consistently mild, as they can either reflect a central tendency in judgments or obscure a pattern of variability that is incompatible with a "subliminal" interpretation (Kush et al. 2018, 2019). To see whether the acceptability means in the two EQ-island/long conditions reflect a central tendency in judgments, we plotted the distribution of z-scores in these conditions and compared it to the distribution of ratings in (i) the uncontroversially grammatical non-island/long conditions for all three islands and (ii) the adjunct and relative clause island/long conditions. Comparison plots are in Figure 4.

Figure 4 shows that filler-gap dependencies into EQs were rated acceptable on most trials. The distribution of ratings in both EQ island/long conditions is similar
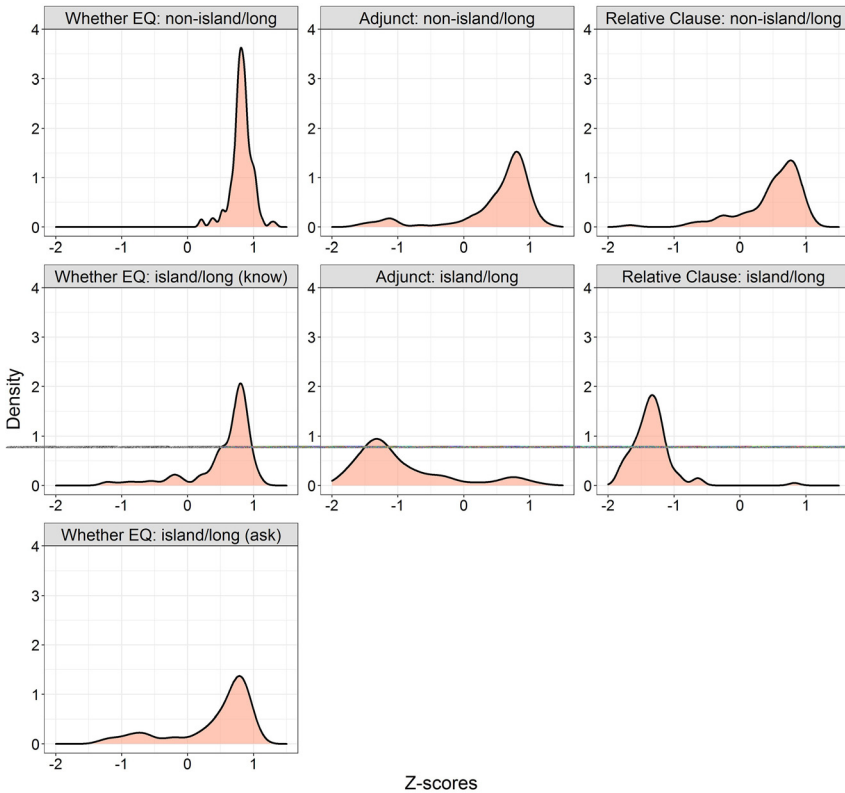
Figure 4: Distribution of *z*-scores in the non-island/long and island/long conditions of Experiment 1.

to the distribution of ratings in the EQ non-island/long condition: most judgments cluster around the highest scores (close to $z = 1$). The two EQ-island/long conditions differ from the non-island/long condition in that there was a higher probability of a lower rating or rejection. Thus, their slightly lower mean acceptability seems to reflect variability in judgments rather than a consistent preference for lower/intermediate ratings. Note also that the EQ island/long conditions patterned very similarly to the grammatical adjunct and relative clause non-island/long conditions, which obtained comparable acceptability means, and they differed strongly from the corresponding ungrammatical island/long conditions, in which judgments clustered around the lowest scores.

To determine whether the variability in the ratings of the EQ island/long conditions was associated to differences between or within participants, we

plotted each participant's highest z-score against her/his lowest z-score in the island/long and non-island/long conditions in Figure 5 (cf. Bondevik et al. 2020; Kush and Dahl 2020; Kush et al. 2018, 2019).
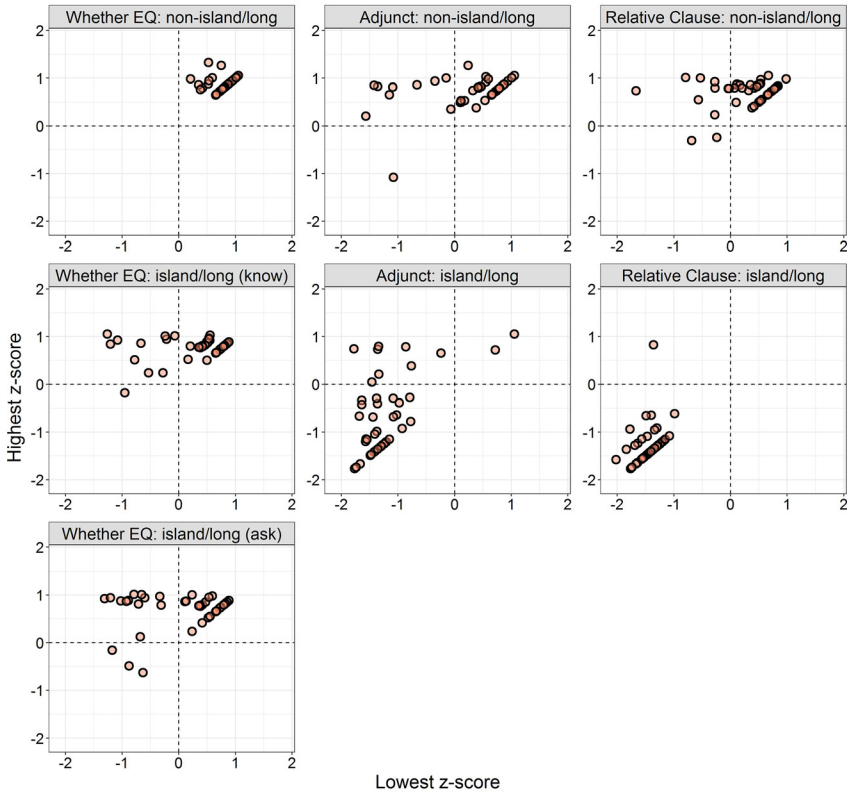
Figure 5: Highest against lowest z-score by participant in the island/long and non-island/long conditions of Experiment 1.

Note that there were three ratings by participant and condition, but for ease of visualization we only show the highest and the lowest rating, as these already comprise the maximum range of variability within individuals. Assuming that z-scores above 0 are acceptable and z-scores below 0 are unacceptable, a participant with both z-scores above 0 (upper right quadrant) is a consistent accepter, a participant with both z-scores below 0 (bottom left quadrant) is a consistent rejecter, and a participant with one z-score above and the other below 0 (upper left

quadrant) is an inconsistent rater. In the EQ island/long conditions, most individuals were consistent accepters, and almost all others were inconsistent. In this, EQ island/long differs from its non-island/long counterpart, consistently accepted by all participants, but it resembles the non-island/long conditions of the adjunct and relative clause subdesigns, which were also rated inconsistently by some individuals. Note also that in the EQ island/long conditions there were almost no consistent rejecters, in contrast to their adjunct and relative clause counterparts, in which consistent rejection was the norm.

Finally, to explore the significant difference in island effects with *ask* and *know*, in Figure 6 we plotted the relationship between each participant's mean DD score with *know* and their mean DD score with *ask*.
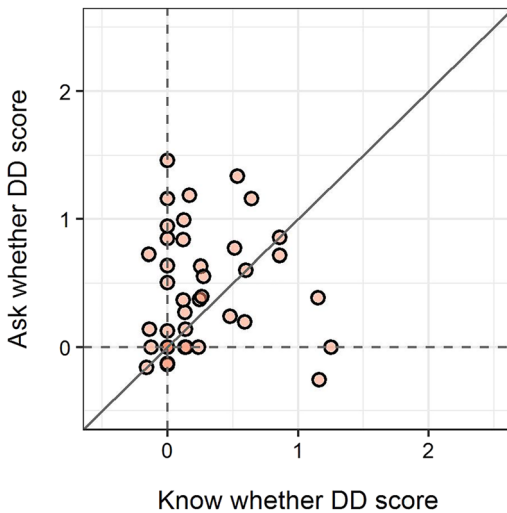
Figure 6: Individual differences-in-differences (DD) scores in *know whether* embedded question islands against individual DD scores in *ask whether* embedded question islands in Experiment 1. Each circle represents a participant. A DD score of 0 or below indicates no super-additive island effect. DD scores greater than 0 are indicative of super-additive island effects.

Participants on the diagonal line had the same DD score with both types of EQs. Participants on the left side of the diagonal line had a greater DD score with *ask*, and participants on its right side had a greater DD score with *know*. The plot shows that many individuals had greater island effects with *ask,* although some displayed a different pattern, with greater *know* island effects or the same effect in both constructions.

2.3 Discussion

In Experiment 1, extraction from *si* 'whether' EQs yielded small island effects. Sentences obtained mean ratings in the acceptable range and mostly high scores, like grammatical control conditions. In contrast, extraction from adjunct and relative clause islands – which we take to be ungrammatical – yielded larger island effects, general unacceptability and mostly low ratings. Because EQs patterned more like grammatical conditions than like ungrammatical controls, we suggest that *whether*-EQs, on the whole, are not syntactic islands in Spanish.

The small *whether* EQ-island effect seems inconsistent with the notion of "subliminal" island effects (Almeida 2014), since the effect was not consistently mild across trials and speakers. Rather, it was caused by a combination of high and a few low ratings. Most participants accepted extraction from *whether* EQs consistently, and thus it appears that for many speakers, extraction from *whether* EQs is (almost) as acceptable as extraction from declarative clauses. Low ratings came by and large from inconsistent participants. We think this suggests that some individuals' judgments were particularly affected by factors unrelated to a syntactic ban on extraction, as there was a similar rate of inconsistent participants in the adjunct non-island/long condition, in which extraction is syntactically allowed.

Island effects were numerically larger with the matrix verb *ask* than with *know*, indicating that *know* sentences had a higher acceptance rate. This effect was observable in most participants. Even so, the matrix verb did not cause any qualitative differences, as extraction was acceptable in both cases and ratings were distributed in a very similar way with both verbs.

The preponderance of acceptable ratings in both EQ-island conditions suggests that neither EQ is a syntactic island in Spanish, contrary to the analysis in Suñer (1991), where it was proposed that *ask*-EQs were syntactic islands and *know*-EQs were not. We note, however, that the effect goes in the direction predicted by Suñer, suggesting that *ask* may introduce some additional factor that increases the probability of mild degradation.

Our *whether*-island results differ strongly from those of López-Sancio (2015) and Pañeda et al. (2020), as these studies found large island effects and unacceptability in sentences with extraction from these clauses. They also differ from findings recently reported by Rodríguez and Goodall (2020), who similarly obtained large *whether*-island effects in Spanish. We attribute this difference to the use of context and complex fillers in our materials, which we included to motivate the sentences pragmatically and reduce the burden of several factors on acceptability. Note that the control adjunct and relative clause island sentences were

strongly unacceptable even though they were presented under the same conditions, indicating that context and complex fillers do not make sentences acceptable when they are – by assumption – ungrammatical (cf. Erteschik-Shir 2006; Keller 2000; Sorace and Keller 2005). In our view, the acceptability contrast between the experimental sentences and the control sentences indicates that there is no syntactic constraint on extraction from *whether*-EQs in Spanish. Thus, the large island effects in previous studies may have reflected semantic, pragmatic, and/or processing factors.

As an anonymous reviewer notes, our results are compatible with the classic observation that embedded questions are weak/selective islands, i.e., islands that allow extraction of D(iscourse)-linked arguments, but not of non-D-linked arguments or adjuncts (Cinque 1990; Rizzi 1990). This is particularly so considering the large island effects and unacceptability that previous studies obtained with bare fillers, which were presumably non-D-linked. We can only agree on this observation. However, we would like to add that, in our view, the fact that these sentences are more acceptable when the *wh*-filler is D-linked is in itself an indication that they are not syntactic violations, as D-linking can be considered a sort of contextualization and should thus have similar effects as context. We are aware that the acceptability contrast between island sentences with and without D-linking has been given syntactic explanations: for instance, Cinque (1990) and Rizzi (1990) have argued that filler-gap dependencies inside weak islands are syntactically licit with D-linked/referential fillers, but syntactically ill-formed with non-D-linked/non-referential fillers (Cinque 1990; Rizzi 1990). However, such explanations seem ad hoc, as there are no obvious syntactic differences between island sentences with and without D-linking. Thus, we believe that, unless evidence for a syntactic difference is provided, it is more parsimonious to assume that all weak island sentences have the same syntactic status and that acceptability differences related to D-linking have extra-syntactic causes.

3 Experiment 2

Our results are in line with the theoretical claim that neither responsive nor rogative EQs are syntactic islands in Spanish, but we only tested embedded polar questions. We know very little about the acceptability profile of *wh*-extraction from other types of EQs. According to Torrego (1984), extraction from EQs introduced by non-argumental *wh*-words is allowed. However, two studies suggest that this might not be the case: first, Rodríguez and Goodall (2020) found that EQs introduced by *dónde* 'where' and *cuándo* 'when' yielded large island effects and unacceptability. Following the logic above, the reason for this could be that they

presented sentences without context. However, Ortega-Santos et al. (2018) found similar results with *por qué* 'why' EQs, even though sentences were presented in context. This suggests that non-polar EQs may disallow extraction even under favorable pragmatic conditions. To assess whether this is true, in Experiment 2, we tested extraction from questions introduced by *cuándo* 'when' with context and complex fillers.

3.1 Method

3.1.1 Participants

51 native speakers of European Spanish were recruited through social media and among students from the University of Oviedo (Spain). Two were excluded because their mean rating in the ungrammatical fillers was higher than 3.5. The remaining 49 participants had a mean age of 40.1 years (range: 18–80) and no self-reported language impairments. Twenty nine were female. Most were born (38) or lived (30) in the region Asturias at the time of testing. Six reported knowledge of other Romance languages spoken in Spain in addition to Spanish (Asturian: 2, Galician: 2, Catalan: 2). Two Amazon vouchers in value of 30 euros were raffled off among the participants as a reward. The experiment was conducted in accordance with the Declaration of Helsinki.

3.1.2 Materials

We used the same experimental, control and filler items as in Experiment 1, but we modified the interrogative experimental items such that they contained *cuándo* (interrogative 'when') rather than *si* 'whether' questions.[4] The context sentences were changed accordingly: *yes/no* context questions were replaced by *when* questions. The non-island conditions remained unchanged. (14) shows a sample *when*-EQ island sentence set.

(14)   Sample *when* EQ-island sentence set
    (a)   Non-island/short
        Context sentence:
        Funcionario   de   Correos: "Cre-o   que hab-éis
        officer   of   post   believe-1SG.PRS that have-2PL.PRS

---

4 In Spanish orthography, interrogative *when* carries an accent (*cuándo*), whereas relative *when* does not (*cuando*).

recog-ido   el   paquete."
pick.up-PTCP the   package
'Post officer: "I believe that you have picked up the package."'
Experimental sentence:
¿Qué   funcionario   __   pens-aba   que hab-íamos
what   officer   __   think-3SG.PST that have-1PL.PST
recog-ido   el   paquete?
pick.up-PTCP the   package
'Which officer __ thought that we had picked up the package?'

(b) *Know*-island/short
Context sentence:
Funcionario de   Correos: "No teng-o   claro
officer   of   post   NEG have-1SG.PRS clear
cuándo   hab-éis   recog-ido   el   paquete."
when   have-2PL.PRS pick.up-PTCP the   package
'Post officer: "It is not clear to me when you have picked up the package."'
Experimental sentence:
¿Qué   funcionario   __   no   sab-ía   cuándo
what   officer   __   NEG   know-3SG.PST when
hab-íamos recog-ido   el   paquete?
have-1PL.PST pick.up-PTCP the package
'Which officer __ didn't know when we had picked up the package?'

(c) *Ask*-island/short
Context sentence:
Funcionario de   Correos: "¿Cuándo hab-éis
officer   of   post   when   have-2PL.PRS
recog-ido   el   paquete?"
pick.up-PTCP the   package
'Post officer: "When have you picked up the package?"'
Experimental sentence:
¿Qué   funcionario __   pregunt-ó cuándo hab-íamos
what   officer   __   ask-3SG.PST when   have-1PL.PST
recog-ido   el   paquete?
pick.up-PTCP the   package
'Which officer __ asked when we had picked up the package?'

(d) Non-island/long

Context sentence:

Funcionario: "Cre-o   que   hab-éis   recog-ido
officer   believe-1SG.PRS that   have-2PL.PRS   pick.up-PTCP

el   paquete   de   Amazon."
the   package   of   Amazon

'Post officer: "I believe that you have picked up the Amazon package."'

Experimental sentence:

¿Qué   paquete   *pens-aba   el   funcionario   que*
what   package   think-3SG.PST the   officer   that

hab-íamos   recog-ido   __?
have-1PL.PST   pick.up-PTCP   __

'Which package did the officer think that we had picked up __?'

(e) *Know*-island/long

Context sentence:

Funcionario: "No   teng-o   claro   cuándo
officer   NEG   have-1SG.PRS clear   when

hab-éis   recog-ido   el   paquete   de Amazon."
have-2PL.PRS pick.up-PTCP the   package   of Amazon

'Post officer: "It is not clear to me when you have picked up the Amazon package."'

Experimental sentence:

¿Qué paquete   *no   sab-ía   el funcionario cuándo*
what   package   NEG   know-3SG.PST the officer   when

hab-íamos recog-ido   __?
have-1PL.PST pick.up-PTCP __

'Which package didn't the officer know when we had picked up __?'

(f) *Ask*-island/long

Context sentence:

Funcionario: "¿Cuándo   hab-éis   recog-ido   el
officer   when   have-2PL.PRS pick.up-PTCP the

paquete   de   Amazon?"
package   of   Amazon

'Post officer: "When have you picked up the Amazon package?"'

Experimental sentence:

¿Qué   paquete   *pregunt-ó   el   funcionario*
what   package   ask-3SG.PST the   officer

cuándo   hab-íamos   recog-ido   __?
when   have-1PL.PST pick.up-PTCP __

'Which package did the officer ask when we had picked up __?'

3.1.3 Procedure and analysis

The experimental procedure and data analysis were identical to Experiment 1.

3.2 Results

The grammatical fillers were rated in the acceptable range (0.95, *SD*: 0.29) and the ungrammatical fillers were rated in the unacceptable range (−0.94, *SD*: 0.54), showing that participants understood the task and carried it out as expected.

   Acceptability by condition in the adjunct, relative clause and *when*-islands is shown in Figure 7. Statistical results are reported in Table 2.
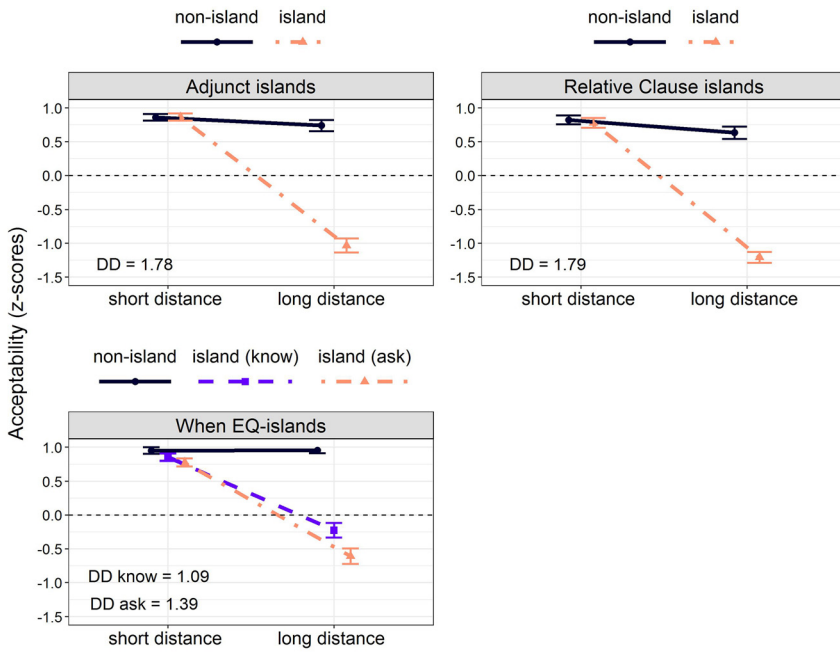
Figure 7: Acceptability by condition in the adjunct, relative clause and embedded question (EQ) islands of Experiment 2. Error bars show 95% confidence intervals. Non-parallel lines are indicative of an island effect.

Table 2: Results of linear mixed models of the island data in Experiment 2.

	β	SE	t	p
Adjunct islands				
Distance	−0.112	0.053	−2.102	0.036 *
Structure	0.004	0.053	0.068	0.946
Distance × Structure	−1.775	0.076	−23.462	<0.001 ***
Relative Clause islands				
Distance	−0.191	0.058	−3.292	0.001 **
Structure	−0.045	0.058	−0.778	0.437
Distance × Structure	−1.795	0.082	−21.898	<0.001 ***
***When* EQ-islands**				
Distance	0.005	0.056	0.095	0.924
Structure	−0.270	0.097	−2.797	0.005 **
Ask-EQ vs. *know*-EQ	−0.078	0.056	−1.402	0.161
Distance × Structure	−2.477	0.137	−18.134	<0.001 ***
Distance × *ask*-EQ vs. Distance × *know*-EQ	−0.306	0.079	−3.876	<0.001 ***

Similar to Experiment 1, we found large, significant super-additive Distance × Structure interactions in the control conditions: the adjunct islands (β = −1.775, SE = 0.076, p < 0.001, DD: 1.78, z-score: −1.03) and relative clause islands (β = −1.795, SE = 0.082, p < 0.001, DD: 1.79, z-score: −1.21).

In the *when*-EQs, we also obtained a super-additive Distance × Structure interaction (β = −2.477, SE = 0.137, p < 0.001), but island effects were smaller than in adjunct and relative clause constructions (DDs: 1.09 with *know* and 1.39 with *ask*). The *when*-EQ island/long conditions obtained higher ratings than their adjunct and relative clause counterparts. However, in contrast to Experiment 1, these were below 0 and thus arguably unacceptable (*know*-island/long: −0.23, *ask*-island/long: −0.61). As in Experiment 1, the interaction was significantly greater with *ask* than with *know* (β = −0.306, SE = 0.079, p < 0.001), indicating that island effects were greater under *ask*.

In order to see the pattern of judgments that underlies the aggregate means, in Figure 8 we plotted the distribution of z-scores in the non-island/long and island/long conditions.

The two *when* EQ-island/long conditions pattern differently from all the non-island/long conditions: judgments do not cluster at the high end of the scale. Judgment distributions for both EQ-island conditions exhibit bimodality, with judgments falling on both ends of the scale, though in differing proportions. With the verb *ask*, most judgments cluster around the low end of the scale, although there are a small number of judgments above 0. With the verb *know*, ratings of *when* EQ-island/long sentences in the two modes are comparable: reflecting an equal proportion of low and high ratings.
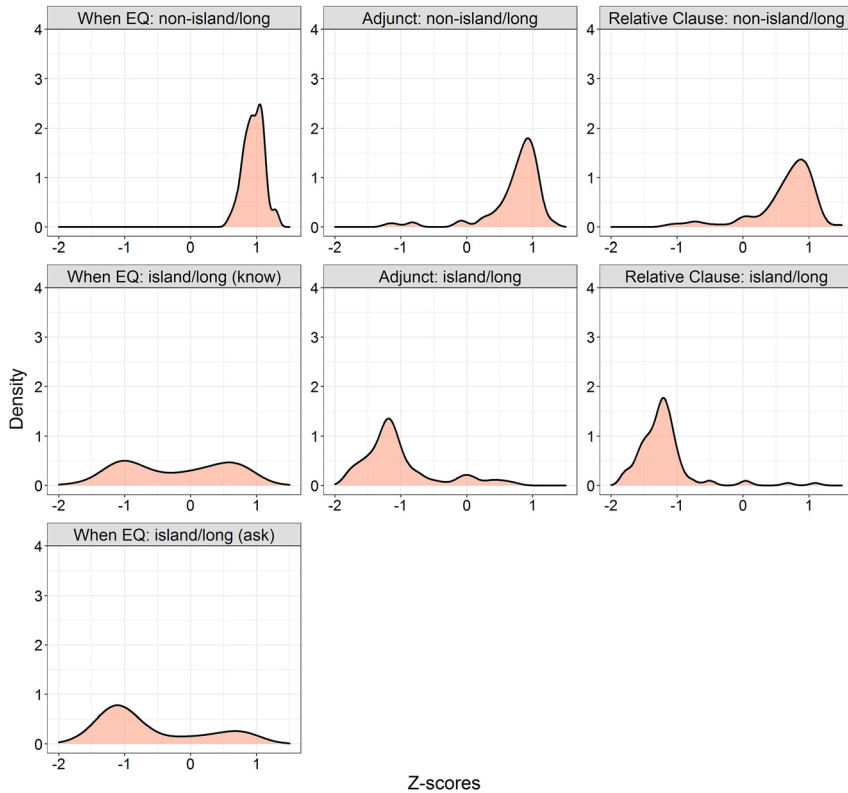
Figure 8: Distribution of *z*-scores in the non-island/long and island/long conditions of Experiment 2.

Figure 9 shows each participant's highest *z*-score against her/his lowest *z*-score in the same two conditions.

Here, the *ask* and *know when* island/long conditions were similar: most participants were either consistent rejecters or inconsistent raters, though there were also a few consistent accepters. The *when*-island/long conditions differ from their non-island counterpart, which was consistently accepted by all participants. They also contrast with the other non-island/long conditions, which were mostly consistently rated. The *when*-island/long conditions further differ from the adjunct and relative clause island/long conditions, which were consistently rejected by almost all speakers.

Finally, Figure 10 shows the relationship between individual DD scores with *know* and *ask* islands. As in Experiment 1, most participants are on the left of the diagonal line, indicating that they exhibited greater island effects with *ask*, but the effect does not hold consistently for all individuals.
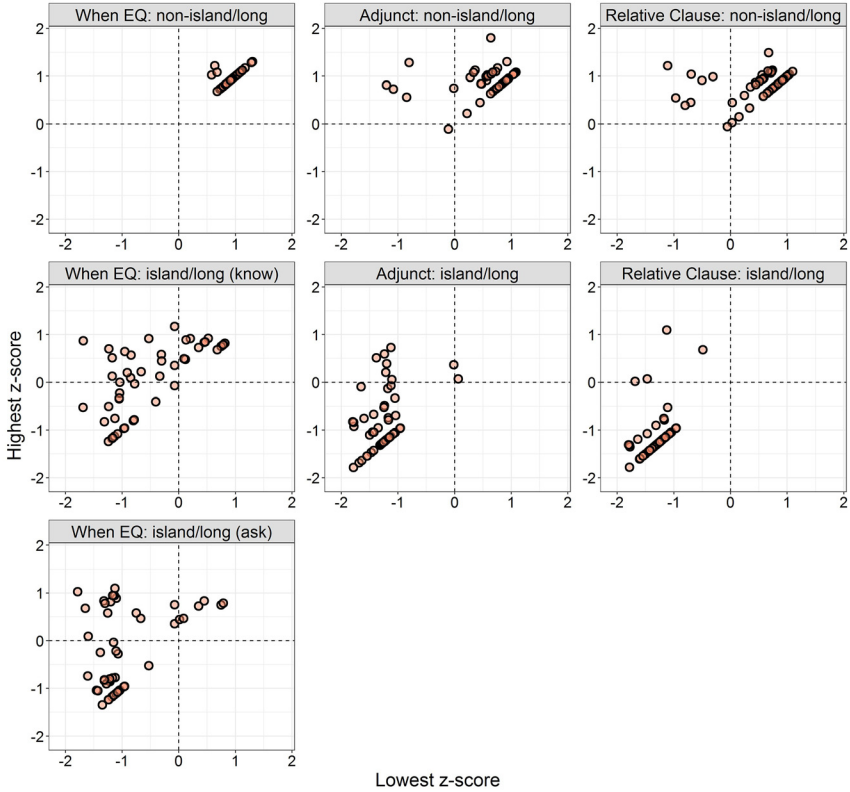
Figure 9: Highest against lowest *z*-score by participant in the non-island/long and island/long conditions of Experiment 2.
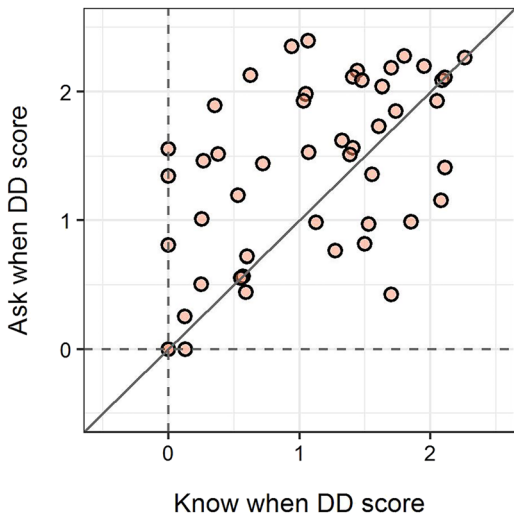
Figure 10: Individual differences-in-differences (DD) scores in *know when* embedded question islands against individual DD scores in *ask when* embedded question islands in Experiment 2. Each circle represents a participant. A DD score of 0 indicates no island effect. DD scores greater than 0 are indicative of island effects.

3.3 Discussion

Replicating the results of Experiment 1, in Experiment 2 island effects were greater and caused more degradation with *ask* than with *know*, but the verb did not yield a qualitative difference in acceptability. The contrast between *ask* and *know* was caused by a higher number of rejections with *ask*, and could be seen in most participants, suggesting again that *ask* introduces an additional constraint that increases degradation.

Extraction from *cuándo* (interrogative 'when') EQs yielded island effects and mean unacceptable ratings. *When*-EQs differed from adjunct and relative clause islands in that they yielded smaller island effects. Underlying the smaller EQ-island effects was a rather variable rating distribution. Judgments of EQ-islands exhibited bimodality, suggesting globally inconsistent judgments of either acceptance or rejection. There were low scores and consistent rejecters, as in ungrammatical conditions, some high scores and consistent accepters, as in grammatical conditions, and many inconsistent participants. The fact that there was a significant number of high scores suggests that there is at least one structural analysis under which *when* EQs are not islands. The cases of rejection may reflect (i) the impact of non-structural factors on acceptability, like semantic-pragmatic infelicity or processing difficulty, or (ii) an alternative analysis under which *when*-EQs are structural islands.

For instance, Torrego (1984) claims that EQs with subject-verb inversion are islands even if the extraction observes bounding constraints. If Torrego's observation is correct, participants would be expected to have rejected extraction from EQs they analyzed as having undergone inversion. Our EQs did not contain any overt subjects, so our stimuli were, in principle, compatible with both an inversion and a non-inversion analysis. It is possible that participants rejected extraction from *when*-EQs that they analyzed as having undergone inversion. Torrego suggests that *when* forces inversion when the predicate is not "heavy", but in that case, we would have expected participants to uniformly reject extraction from *when*-EQs. The fact that we saw variation in acceptance could mean that individual participants applied inversion inconsistently across trials, or that there is inter-speaker variation in whether *when* triggers inversion. This is an empirical question yet to be investigated.

4 General discussion

In two 7-point acceptability judgment experiments on Spanish, we examined whether matrix verb choice (rogative *ask* vs. responsive *know*) affected extraction

from *si* 'whether' and *cuándo* 'when' embedded questions. Island effects were tested in *wh*-question configurations with complex fillers in object position, and we presented test sentences in context to make them semantically and pragmatically motivated. Adjunct and relative clause islands were tested under the same conditions as a control. We found island effects in all cases, but these were greater in adjunct and relative clause islands than in EQs.

In Experiment 1, island effects for extraction from *whether*-EQs were relatively small under both responsive *know* and rogative *ask*. On most trials participants accepted *wh*-extraction from EQs under both types of verbs. Island effects were slightly larger with *ask* than with *know*, which reflected a slightly higher probability of rejection, but not a qualitative difference in overall acceptability.

In Experiment 2, sentences with extraction from *when*-EQs exhibited larger island effects, with average judgments that fell under standard thresholds for acceptability. Ratings of *wh*-extraction from *when*-EQs were bimodally distributed. Judgments of *wh*-extraction from *when*-EQs under the responsive *know* were equally likely to be high or low, whereas judgments of extraction from *when*-EQs under rogative *ask* were more often low. In what follows, we first address the effects of matrix verb choice and then discuss the contrast between the *whether* and *when* results.

4.1 Matrix verb effects

Our findings from Experiment 1 (on *whether* islands) argue against Suñer's (1991) syntactic analysis of selective verb-driven island effects. If extraction from an EQ under a rogative verb resulted in a syntactic constraint violation, as claimed by Suñer, participants should have rejected *wh*-extraction from *ask*-EQs and we should have observed larger island effects. Instead, extraction from *ask*-EQs obtained average ratings in the acceptable range and yielded small effects. Seventy seven percent of trials obtained ratings of 5 or above on a 7 pt scale (compared to 14% in adjuncts and 1% in relative clauses), and 64% obtained *z*-scores of 0.5 or above (compared to 10% in adjuncts and 1% in relative clauses). Furthermore, most participants accepted these sentences consistently, and there were very few consistent rejecters, at rates comparable to those of consistent rejecters in uncontroversially grammatical sentences. This suggests that extraction from *ask*-EQs is syntactically allowed, like extraction from *know*-EQs, which was rated in a similar way.

In Experiment 2, extraction from *ask*-EQs had a lower acceptance rate, which fell below the midpoint of the scale. Still, acceptability was higher than in the control islands (27% of trials obtained ratings of 5 or above, compared to 10% in

adjuncts and 4% in relative clauses, and 16% had *z*-scores of 0.5 or above, compared to 3% in adjuncts and 2% in relative clauses). Furthermore, just like in Experiment 1, *ask*-EQ islands patterned more with *know*-EQ islands than with the control islands, both in terms of average acceptability and in the rating distribution, which was variable. Because of all this, we believe that our results provide strong evidence against the hypothesis that the *structural* islandhood of Spanish EQs depends on the type of matrix verb.[5]

Even so, island effects were quantitatively greater with *ask* than with *know* in both experiments. Thus, it seems that the rogative/responsive distinction does have a (non-categorical) effect on acceptability. It remains to be determined what to attribute this distinction to. One possibility is that the difference in the acceptability of extraction from *ask* and *know* questions has a semantic cause. Indeed, questions embedded under each type of verb have different readings: they are interpreted as "real" indirect questions under *ask*, but as propositions under *know*. In relation to this, *ask*-EQs may contain a question feature that their *know* counterparts lack. This feature may cause mild intervention effects or interfere with the retrieval of the filler at the gap position (see de Cuba and MacDonald 2013 for a related proposal).

Another possibility within the domain of semantic approaches to island effects comes from Abrusán (2014). Abrusán notes that different embedding predicates induce stronger *wh*-island effects depending on how difficult the lexical semantics of the predicate makes satisfying a *maximal informativity requirement* on *wh*-questions. It is possible, under certain conditions, to meet the requirement with verbs such as *know* (Abrusán 2014: 139–143), but the conditions under which questions with *ask* satisfy the requirement may be more narrow or exacting. We leave working this out to future research.

Note also that, unlike *know*-EQs, *ask*-EQs may be interpreted as uttered questions or "question acts" (Krifka 2001; Lahiri 2002). This possibility was favored in our study by the context, which presented the questions as reported speech. Question acts have been argued to have a simplified Boolean algebra that only allows some types of quantifiers to scope out of them (Krifka 2001). This may interact with island effects, as these have also been considered a scope phenomenon related to a simplified Boolean algebra (Szabolcsi and Zwarts 1993). Future research should address whether extraction from embedded question acts is

---

5 Note that variation in judgments to *ask when*-EQ islands cannot be taken to indicate that extraction from *ask*-EQs is grammatical for some participants and ungrammatical for others, as there was little variation in *ask whether*-EQs and even more variation in *know when*-EQs. Variation seems to be related to the *whether* vs. *when* contrast rather than to the rogative vs. responsive contrast.

expected to be less acceptable than extraction from other types of EQs on account of their semantic differences.[6]

Our findings are further relevant for the interpretation of the results from previous studies on Spanish islands. Both López-Sancio (2015) and Pañeda et al. (2020) found large island effects in *whether* questions embedded under the verb *preguntar(se)* 'to ask/to wonder'. Since the effects of this verb on extraction had not been directly assessed, it was unclear to what extent it could have influenced the results. Our results suggest that even if it may have indeed increased the size of the effects, the fundamental cause for these large island effects must lie elsewhere.

4.2 *Whether* vs. *when* embedded questions

While the presence of EQ-island effects in Spanish does not seem to be strongly determined by the type of embedding verb, the type of interrogative word introducing the EQ seems to have a larger impact: *si* 'whether' island effects were small, obtained mean ratings in the acceptable range and yielded relatively little variation, whereas *cuándo* 'when' island effects were greater, yielded mean unacceptable ratings and much more variation. This difference suggests that the picture of extraction from Spanish EQs is more complex than commonly thought.

Note, though, that under classical bounding approaches (Chomsky 1973, 1977), this contrast does not reflect a bounding difference. From this perspective, EQ island effects arise because *wh*-movement is forced to cross more than one (TP) bounding node in a single movement. This occurs because the intervening

---

6 One potential concern is that we are underestimating the size of the difference between *know* and *ask*, because an additional factor is increasing the size of the island effect in *know* conditions: our *know* conditions were preceded by negation, while the corresponding *ask* conditions were not. Extraction out of a negated complement can result in (weak) island effects (Rizzi 1990; Ross 1984; Szabolcsi and Zwarts 1993). If negation contributes to additional unacceptability, then a greater contrast between the two verbs could have been found if *know* had not been negated. While we cannot reject the possibility, we consider it unlikely that negation affected our results strongly because *whether* questions embedded under *know* obtained acceptable ratings (z-score: 0.55), which were similar (and descriptively higher) in mean than those of the grammatical control adjunct and relative clause non-island/long sentences (0.50 and 0.49, respectively). Thus, even if negation reduces acceptability, it does not seem to do so more than any extra-grammatical factors that might affect the ratings of the control non-island/long sentences. We also note that *know* questions without negation (e.g., *Which package did the officer know whether we had picked up?*) seem less natural. Thus, a comparison of *ask* and *know* without negation could also introduce an additional source of unacceptability. Therefore, we believe that such a comparison would not be more likely than ours to find a strong difference in the acceptability of extraction from questions embedded under the two verbs.

specifier,CP "escape hatch" is occupied by the embedded wh-phrase. If Spanish chooses CP as its left-peripheral bounding node (Rizzi 1982; Torrego 1984) how-ever, direct movement across the filled intermediate specifier,CP should not result in a Subjacency violation. If indeed CP is a bounding node in Spanish, then extraction from all EQs should be syntactically licit.

An open question is whether a bounding explanation for the difference be-tween *whether* and *when*-EQs is possible under more modern assumptions, for instance under the framework of Phase Theory (e.g., Chomsky 2000, 2001, 2008; see also Citko 2013). Phase Theory proposes that sentences are derived by phases and that once a phase is completed, part of it undergoes Spell-Out, becoming inaccessible for operations such as movement. Which categories count as phases is a matter of discussion. It is often argued that CPs are phases, but CPs are also assumed to have a rich internal structure, with multiple phrasal projections (Rizzi 1997, 2001). This opens up the possibility that only some CP projections are phases. In this context, the contrast between *whether* and *when*-EQs could be explained provided *when* occupies a phasal projection and *whether* does not. It has been suggested that *si* 'whether' occupies an Int(errogative) projection within the CP, whereas *wh*-words such as *when* occupy a Foc(us) projection (Rizzi 2001; see also Hernanz Carbó 2012). Given this, it could be argued that Foc is a phase head, while Int is not (see Yoshimoto 2012 for a related proposal). However, we are unaware of empirical and conceptual arguments that support this assumption. Assuming that the outermost specifier of a domain is the phase head (Bošković 2016), Int seems more likely to be a phase than Foc, as it is in a higher projection according to Rizzi (2001). This seems to predict that extraction from *whether* EQs should be more degraded than extraction from *when* EQs, against our findings.

We have argued that a syntactic explanation in terms of bounding is not possible, at least not under the common understanding of this notion. It remains to be determined whether the contrast between *whether* and *when*-EQs could be accounted for within syntax by other proposals. For instance, if, as argued by Torrego (1984), EQs with subject-verb inversion disallow extraction (see Section 3.3), the contrast between *whether* and *when* EQs might reflect differences in the extent to which each interrogative forces inversion.[7] It remains to be established why inversion should interact with islandhood, as Torrego originally accounted for it by means of the Empty Category Principle (Chomsky 1981), which is not part of the theoretical apparatus anymore.

Another syntactic proposal that could account for the differences is featural Relativized Minimality (fRM; Belletti et al. 2012; Friedmann et al. 2009; Rizzi 1990,

---

7 According to Torrego (1984), inversion may be required in *when* EQs with non-"heavy" predi-cates, but it is fully optional in *whether* EQs.

2013; Villata et al. 2016). This theory explains EQ-island effects as intervention effects caused by the similarity between the *wh*-filler and the interrogative intro- ducing the EQ. An explanation of the *whether-when* differences in terms of fRM would have to posit that Spanish *cuándo* 'when' is more similar to the *wh*-filler than *si* 'whether', and it is therefore a stronger intervener.[8] However, we think that an fRM account will have trouble explaining inter-trial inconsistency in judgments.

It might also be argued that the differences between *whether* and *when*-EQs are not caused by any syntactic constraint and that they arise instead for grammatical, but extra-syntactic reasons, like a semantic factor. One potential avenue for explaining the differences has to do with presuppositions. *When we had picked up __* in (14e), (14f) presupposes that something was picked up, whereas *whether we had picked up __* in (11e), (11f) does not. Clauses that express presupposed content (such as complements of factive verbs) are known to yield weak island effects (Abrusán 2011, 2014; Cinque 1990; Szabolcsi 2006), so the larger island effects in *when*-EQs could be related to their presupposing the event in the embedded clause. We leave the exploration of semantic accounts for future research.

Finally, the contrast in the acceptability of extraction from *whether* and *when*-EQs could (at least partially) reflect that the former are easier to process than the latter. First, *when*-island sentences contain two filler-gap dependencies, while their *whether* counterparts contain only one, and an additional dependency can increase processing burden. Second, if *when* shares more features with the *wh*-filler than *whether*, as discussed above, this can cause similarity-based interference, hindering the retrieval of the *wh*-filler at the gap position (Atkinson et al. 2016; Keshev and Meltzer-Asscher 2019; Villata et al. 2016). Similarity-based interference may also explain why extraction from *whether* clauses was rejected more often than extraction from declarative clauses, because non-syntactic features also count for the computation of similarity under a processing account.[9]

Finally, moving on to other EQ-islands: most of the discussed accounts for the difference between *whether* and *when*-island effects predict that other *wh*-in- terrogatives should pattern like *when*. Indeed, Ortega-Santos et al. (2018) provide

---

8 This could be possible under an analysis in which *si*, as a complementizer (Brucart 1993; Hernanz Carbó 2012), does not interfere in the same way as a movable phrase (e.g., *when*) in a specifier position does. Alternatively, if there is a null operator in spec,CP of a *si*-headed CP, a fRM analysis might instead distinguish between overt and covert question operators for the purposes of intervention.

9 In contrast, the syntactic fRM account only considers syntactic features for the computation of similarity, and thus the contrast between extraction from *whether*-EQs and extraction from declarative clauses remains unexplained. In any case, note the similarity between the two ac- counts, which has led Ortega-Santos (2011) to argue that Relativized Minimality is a grammatical convention based on the workings of memory.

evidence that *por qué* 'why' patterns like *when*, as they found that *por qué* yielded large island effects even though test sentences were presented in a context that arguably reduced semantic/pragmatic infelicity. Thus, it is possible that extraction from Spanish EQs is considerably degraded with all interrogatives except for *whether*, against previous theoretical claims, even if this degradation is not caused by the violation of a bounding constraint.[10]

5 Conclusion

In two acceptability experiments, we investigated extraction from Spanish *whether* and *when* questions embedded under the responsive verb *know* and the rogative verb *ask*, in order to test the theoretical claim that questions embedded under responsive verbs do not violate Bounding constraints, but questions embedded under rogative verbs do (Suñer 1991). Because we found no qualitative differences in the acceptability of extraction from questions under the two verbs, we argue that there is no such Bounding contrast. Even so, rogative verbs might pose additional non-structural constraints on extraction, as island effects were numerically greater under *ask*.

Furthermore, because extraction from *whether* questions was generally acceptable, our results provide empirical evidence that extraction from Spanish embedded questions, on the whole, does not violate bounding constraints (Torrego 1984). While extraction from *when* embedded questions was more degraded, the contrast between *whether* and *when* cannot reflect a Bounding difference under the common understanding of Bounding. Instead, it may be related to the featural composition of the two interrogatives, the presuppositions they introduce and/or processing factors.

---

10 Note, though, that *por qué* 'why' is usually assumed to be base-generated (e.g., Rizzi 2001). If this is true, *por qué* may not force subject-verb inversion (since inversion is related to *wh*-movement; Torrego 1984) and sentences with extraction from *por qué*-EQs should only contain one filler-gap dependency. From this perspective, *por qué* island effects can neither be attributed to inversion nor to the difficulty of processing two dependencies. Assuming *por qué* and *cuándo* island effects have the same cause, this could help tease apart the competing accounts presented in this section.

Author contributions: CP ran the experiments, analyzed the data and wrote the first draft. Both authors conceived and designed the experiments and contributed to the writing of the article.

References

Abrusán, Márta. 2014. *Weak island semantics*. New York: Oxford University Press.

Abrusán, Márta. 2011. Presuppositional and negative islands: A semantic account. *Natural Language Semantics* 19. 257–321.

Aldosari, Saad Mohammed. 2015. *The role of individual differences in the acceptability of island violations in native and non-native speakers*. Lawrence, KS: University of Kansas dissertation.

Almeida, Diogo. 2014. Subliminal wh-islands in Brazilian Portuguese and the consequences for syntactic theory. *Aballin* 13(2). 55–93.

Atkinson, Emily, Aaron Apple, Kyle Rawlins & Akira Omaki. 2016. Similarity of wh-phrases and acceptability variation in wh-islands. *Frontiers in Psychology* 6. https://www.frontiersin.org/articles/10.3389/fpsyg.2015.02048/full (accessed 29 May 2020).

Bates, Douglas, Martin Mächler, Bolker Ben & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). https://www.jstatsoft.org/article/view/v067i01 (accessed 29 May 2020).

Belletti, Adriana, Naama Friedmann, Dominique Brunato & Luigi Rizzi. 2012. Does gender make a difference? Comparing the effect of gender on children's comprehension of relative clauses in Hebrew and Italian. *Lingua* 122(10). 1053–1069.

Bondevik, Ingrid, Dave Kush & Terje Lohndal. 2020. Variation in adjunct islands: The case of Norwegian. *Nordic Journal of Linguistics*. 1–32. https://www.cambridge.org/core/journals/nordic-journal-of-linguistics/article/variation-in-adjunct-islands-the-case-of-norwegian/2425F6A18A0DF20AD9AB8B69368658CE (accessed 02 January 2021).

Bošković, Željko. 2016. Getting really edgy: On the edge of the edge. *Linguistic Inquiry* 47(1). 1–33.

Brucart, José María. 1993. Sobre la estructura de SCOMP en español. In Amadeu Viana (ed.), *Sintaxi: Teoria i perspectives*, 59–102. Lleida: Pagès.

Chomsky, Noam. 1964. *Current issues in linguistic theory*. The Hague: Mouton.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen Anderson & Paul Kiparsky (eds.), *A festschrift for Morris Halle*, 232–286. Holt: Rinehart and Winston.

Chomsky, Noam. 1977. On Wh-Movement. In Peter Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal syntax*, 71–132. New York: Academic Press.

Chomsky, Noam. 1981. *Lectures on government and binding: The Pisa lectures*. Berlin & New York: Mouton De Gruyter.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels, Juan Uriagereka & Samuel Jay Keyser (eds.), *Step by step*, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–50. Cambridge, MA: MIT Press.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & María Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 1990. *Types of Ā-dependencies*. Cambridge, MA: MIT Press.

Citko, Barbara. 2013. *Phase theory: An introduction*. Cambridge: Cambridge University Press.

de Cuba, Carlos & Jonathan E. MacDonald. 2013. Referentiality in Spanish CPs. In Victoria Camacho-Taboada, Ángel L. Jiménez Fernández, Javier Martín-González & Mariano Reyes-Tejedor (eds.), *Information structure and agreement*, 117–140. Amsterdam & Philadelphia: John Benjamins.

Drummond, Alex. 2013. Ibex farm. http://spellout.net/ibexfarm/ (accessed 10 April 2020).

Erteschik-Shir, Nomi. 2006. What's what? In Gisbert Fanselow, Caroline Fery, Matthias Schlesewsky & Ralf Vogel (eds.), *Gradience in grammar: Generative perspectives*, 317–335. Oxford: Oxford University Press.

Friedmann, Naama, Adriana Belletti & Luigi Rizzi. 2009. Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua* 119(1). 67–88.

Goodall, Grant. 2015. The D-linking effect on extraction from islands and non-islands. *Frontiers in Psychology* 5. 1–11.

Hernanz, Carbó & María Lluïsa. 2012. Sobre la periferia izquierda y el movimiento: El complementante "si" en español. In José María Brucart & Ángel J. Gallego (eds.), *El movimiento de constituyentes*, 151–171. Madrid: Visor.

Hofmeister, Philip & Ivan A. Sag. 2010. Cognitive constraints and island effects. *Language* 86(2). 366–415.

Huang, Cheng-Teh James. 1982. *Logical relations in Chinese and the theory of grammar*. Cambridge, MA: MIT dissertation.

Keller, Frank. 2000. *Gradience in grammar: Experimental and computational aspects of degrees of grammaticality*. Edinburgh: University of Edinburgh dissertation.

Keshev, Maayan & Aya Meltzer-Asscher. 2019. A processing-based account of subliminal wh-island effects. *Natural Language & Linguistic Theory* 37(2). 621–657.

Kluender, Robert & Marta Kutas. 1993. Subjacency as a processing phenomenon. *Language & Cognitive Processes* 8(4). 573–633.

Krifka, Manfred. 2001. Quantifying into question acts. *Natural Language Semantics* 9(1). 1–40.

Kush, Dave & Anne Dahl. 2020. L2-transfer of L1 island-insensitivity: The case of Norwegian. *Second Language Research*. 1–32. https://journals.sagepub.com/doi/full/10.1177/0267658320956704 (accessed 05 October 2020).

Kush, Dave, Terje Lohndal & Jon Sprouse. 2018. Investigating variation in island effects: A case study of Norwegian wh-extraction. *Natural Language & Linguistic Theory* 36(3). 743–779.

Kush, Dave, Terje Lohndal & Jon Sprouse. 2019. On the island sensitivity of topicalization in Norwegian: An experimental investigation. *Language* 95(3). 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(13). 1–26.

Lahiri, Utpal. 2002. *Questions and answers in embedded contexts*. Oxford: Oxford University Press.

López-Sancio, Sergio. 2015. *Testing syntactic islands in Spanish*. Vitoria-Gasteiz: University of the Basque Country MA thesis.

Michel, Dan. 2014. *Individual cognitive measures and working memory accounts of syntactic island phenomena*. San Diego, CA: University of California San Diego dissertation.

Nunes, Jairo & Juan Uriagereka. 2002. Cyclicity and extraction domains. *Syntax* 3(1). 20–43.

Ortega-Santos, Iván. 2011. On Relativized Minimality, memory and cue-based parsing. *Iberia: An International Journal of Theoretical Linguistics* 3(1). 35–64.

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, Sol Lago, Elena Vares, João Veríssimo & Claudia Felser. 2020. Island effects in Spanish comprehension. *Glossa: A Journal of General Linguistics* 5(1). 1–30.

Pesetsky, David. 1987. Wh-in-situ: Movement and unselective binding. In Eric J. Reuland & Alice G. B. ter Meulen (eds.), *The representation of (in)definiteness*, 98–129. Cambridge, MA: MIT Press.

Phillips, Colin. 2006. The real-time status of island phenomena. *Language* 82(4). 795–823.

R Core Team. 2020. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. https://www.r-project.org/ (accessed 10 April 2020).

Rizzi, Luigi. 1982. Violations of the wh- island constraint and the subjacency condition. In Luigi Rizzi (ed.), *Issues in Italian syntax*, 49–76. Dordrecht: Foris.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar. Handbook in generative syntax*, 281–337. Dordrecht: Springer.

Rizzi, Luigi. 2001. On the position "Int (errogative)" in the left periphery of the clause. In Guglielmo Cinque & Giampaolo Salvi (eds.), *Current studies in Italian syntax: Essays offered to Lorenzo Renzi* (North-Holland linguistic series 59), 287–296. New York: Elsevier.

Rizzi, Luigi. 2013. Locality. *Lingua* 130. 169–186.

Rodríguez, Alejandro & Grant Goodall. 2020. On the universality of wh-islands: Experimental evidence from Spanish. In *Poster presented at the 50th Linguistic Symposium on Romance Languages*. Austin: University of Texas, 1–8 July.

Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, MA: MIT dissertation.

Ross, John Robert. 1984. Inner islands. In Claudia Brugman & Monica Macaulay (eds.), *Proceedings of the 10th Annual Meeting of the Berkeley Linguistics Society* (BLS), 258–265. Berkeley: Berkeley Linguistics Society.

Sorace, Antonella & Frank Keller. 2005. Gradience in linguistic data. *Lingua* 115. 1497–1524.

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(1). 307–344.

Sprouse, Jon, Shin Fukuda, Hajime Ono & Robert Kluender. 2011. Reverse island effects and the backward search for a licensor in multiple wh-questions. *Syntax* 14(2). 179–203.

Sprouse, Jon, Matt Wagers & Colin Phillips. 2012. A test of the relation between working-memory capacity and syntactic island effects. *Language* 88(1). 82–123.

Stepanov, Arthur, Manca Mušič & Penka Stateva. 2018. Two (non-) islands in Slovenian: A study in experimental syntax. *Linguistics* 56(3). 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 Martin Everaert & Henk C. van Riemsdijk (eds.), *The Blackwell companion to syntax*, vol. 4, 479–531. Malden, MA: Blackwell.

Szabolcsi, Anna & Frans Zwarts. 1993. Weak islands and an algebraic semantics for scope taking. *Natural Language Semantics* 1(3). 235–284.

Torrego, Esther. 1984. On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15(1). 103–129.

Tucker, Matthew A., Idrissi Ali, Jon Sprouse & Diogo Almeida. 2019. Resumption ameliorates different islands differentially: Acceptability data from Modern Standard Arabic. In Amel Khalfaoui & Matthew Tucker(eds.), *Perspectives on Arabic linguistics*, vol. 30, 1–52. Amsterdam & Philadelphia: John Benjamins.

Villa-García, Julio. 2015. *The syntax of multiple-que sentences in Spanish*. Amsterdam & Philadelphia: John Benjamins.

Villata, Sandra, Luigi Rizzi & Julie Franck. 2016. Intervention effects and Relativized Minimality: New experimental evidence from graded judgments. *Lingua* 179. 76–96.

Yoshida, Masaya, Nina Kazanina, Leticia Pablos & Patrick Sturt. 2014. On the origin of islands. *Language & Cognitive Processes* 29(7). 761–770.

Yoshimoto, Keisuke. 2012. The left periphery of CP phases in Japanese. *Acta Linguistica Hungarica* 59(3). 339–384.