Schoenmakers, Gert-Jan & Stoica, Irina. 2024. An experimental investigation of *wh*-dependencies in four island types in Romanian. *Glossa: a journal of general linguistics* X(X). pp. 1–39. DOI: https://doi.org/10.16995/glossa.15193

Open Library of Humanities

An experimental investigation of *wh*-dependencies in four island types in Romanian

Gert-Jan Schoenmakers, Institute for Language Sciences, Utrecht University, Utrecht, NL, g.t.schoenmakers@uu.nl

Irina Stoica, Faculty of Foreign Languages and Literatures, University of Bucharest, Bucharest, RO, irina.stoica@lls.unibuc.ro

This paper gauges the acceptability of *wh*-extraction from four island types in Romanian, an understudied language in this respect. Linguistic theories predict *wh*-dependencies to be unacceptable when the gap is located inside adjuncts, complex NPs, embedded interrogative clauses, or subjects. After reviewing different theoretical approaches to island phenomena and experimental evidence from other Romance languages, we report on two acceptability judgment experiments, adapted from Sprouse et al.'s (2016) experiments for Italian. In order to test (conditional) adjunct, complex NP, and interrogative (*whether*) islands, the experiments crossed two factors: *distance* (between filler and gap, short vs. long) and *construction* (island vs. non-island). The investigation of subject islands required an alternative design for reasons specific to the structure of Romance languages. To test subject islands, we again followed the design used by Sprouse et al. (2016) and crossed the factors *gap site* (subject vs. object) and *complexity* (simple vs. complex). In line with predictions from most linguistic theories and the results from experimental work on Romance languages, we find significant island effects for each construction type. The subject island effect is weaker than the others, which may be an artifact of the alternative research design. Moreover, we discuss how our data from the adjunct and interrogative island item sets pattern in a somewhat unexpected manner. The present study provides initial experimental evidence for island sensitivity in Romanian bare *wh*-questions, which can serve as an empirical anchor for future investigations.

1 Introduction

Island constructions are known to constrain the formation of long-distance dependencies: *what* in (1a), for example, can freely move out of the embedded clause into the matrix clause,[1] whereas the embedded interrogative clause headed by *whether* in (1b) resists *wh*-extraction. The interrogative clause is thus said to constitute an *island*, a term coined by Ross (1967) to describe environments that block filler-gap dependencies – environments such as interrogative clauses (1b), adjuncts (2a), complex NPs (2b), and subjects (2c), among others (although not all types were originally reported in Ross 1967).

(1) a. **What** did you say [that Mary ate ___]?
b. *__What__ did you wonder [whether Mary ate ___]? (interrogative island)

(2) a. *__What__ did you get angry [because Mary ate __]? (adjunct island)
b. *__What__ did you hear [the rumor that Maria baked __]? (complex NP island)
c. *__What__ did you believe [the report about __] caused a fuzz? (subject island)

Island phenomena have long been a central topic in the theoretical literature, in part due to the multitude of factors which may impact the possibility of extraction from an island, such as the dependency type, the discourse status of the filler, or the construction and language under investigation (see the articles in Sprouse & Hornstein 2013).

Traditionally, the reason behind the unacceptability of sentences like (1b) and (2) has been grounded within syntax (see Boeckx 2012 or den Dikken & Lahne 2013 for a history), most prominently in the guise of Chomsky's (1973; 1977) Subjacency Condition. However, many different accounts have been developed since then, which do not take a 'pure syntax' stance, but instead look for the cause of the (un)acceptability patterns elsewhere, in some cases (but not all) ruling out grammatical constraints altogether. Such non-syntactic approaches claim that sentences such as (1b) and (2) violate discourse principles and/or overburden the processing system (see Boeckx 2012; Newmeyer 2016; Liu et al. 2022b). We discuss the different approaches to island phenomena in Section 2.1.

Further, the past decade has seen a strong proliferation of experimental investigations of island phenomena (see Section 2.2), which have provided a richer empirical basis to the theoretical discussion. These investigations also include various studies on the Romance languages. Romanian, however, remains an understudied language in this respect. This paper assesses whether Romanian exhibits island sensitivity, by adopting the same experimental

[1] Transformational accounts of syntactic phenomena hold that dependencies are formed via movement of the *wh*-element from the position in which it is interpreted (gap site) to the position in which it is pronounced (filler site). Although movement-free approaches exist, here we adopt the movement hypothesis.

design as has been used for Italian (Sprouse et al. 2016) and Spanish (López Sancio 2015). The consensus in the theoretical literature is that sentences such as (1b) and (2) are unacceptable in Romanian as well, but experimental support for this claim is scarce. As far as we can tell, it is limited to a single study (Stoica & Schoenmakers 2024) which investigates *wh*-dependencies with the gap located inside interrogative embedded clauses headed by *dacă* 'whether' and *când* 'when', with different filler types (bare *wh*, *what* N, and *which* N). Stoica and Schoenmakers find that long-distance *wh*-dependencies originating inside embedded *wh*-clauses indeed receive much lower acceptability ratings than those originating in embedded *că* 'that' clauses, in line with predictions from theoretical literature. We provide more details on the claims about island sensitivity in Romanian in Section 2.4, against the backdrop of earlier experimental studies on island phenomena in other Romance languages, presented in Section 2.3.

In Section 3, we report novel experimental data on the acceptability of *wh*-dependencies in four island types in Romanian (conditional adjuncts, complex NPs, embedded interrogative (*whether*) clauses, and subjects). Our experiments are similar in design to earlier experimental work investigating Spanish and Italian (López Sancio 2015; Sprouse et al. 2016) and cross two factors in a 2×2 design: the distance between the filler and the gap (i.e. whether the gap is located in the matrix or embedded clause) and the presence or absence of an alleged island structure. The subject island item set required an alternative design in Romanian, however, since the language does not allow preposition stranding (P-stranding, see Sprouse et al. 2016 and Abeillé et al. 2020 for discussion). This sub-experiment instead crossed the factors *gap site* (subject vs. object) and *complexity* (of the extraction domain; simple vs. complex). In most of the literature, island effects are said to emerge when there is a significant interaction effect between the two factors (Sprouse & Villata 2021).[2] The rationale is that the two factors may independently affect acceptability (e.g. because of the increased distance between filler and gap, or because of the increased complexity of the embedded clause), but a super-additive effect would require an additional explanation. As noted, this additional explanation is typically given either in the form of a syntactic constraint or in the form of other, non-syntactic conditions on the human cognitive system.

We find significant island effects in our data-sets, which indicates that Romanian is sensitive to all four island types, at least with (bare) *wh*-dependencies. This finding is in line with theoretical claims about Romanian and with earlier experimental work on island effects in other Romance

[2] Liu et al. (2022b) argue that the definition of an *island* as a particular structure that triggers unacceptability under given circumstances presumes a certain knowledge about the source of this unacceptability. However, acceptable examples of long-distance dependencies with the gap located inside a 'syntactic island' suggest that the source of the unacceptability may not in fact be syntactic. Liu et al. (2022b: 496) instead define an island as "an unacceptable filler-gap dependency, which has been traditionally interpreted as ungrammatical."

languages. It is worth noting here, however, that various factors which we did not test in these experiments may impact the island effect sizes (e.g. the dependency type or the mood of the embedded clauses). We discuss various linguistic theories that put forward different predictions about such alternative configurations. Therefore, Section 4 provides discussion of the limitations of our study and potential future directions. Section 5 draws our conclusions.

2 Theoretical background

2.1 Islands at the intersection between syntactic, discourse-based, and processing accounts

Long-distance dependencies in general, and islands in particular, have been at the core of many debates in the literature for decades. The contrast between sentences like (1a) and (1b) has been accounted for in different ways. Proponents of grammatical accounts have put forth a series of syntactic constraints that somehow constrain extraction from island structures. One of the most influential syntactic analyses revolves around the Subjacency Condition (Chomsky 1973; 1977) that blocks movement over two or more 'cyclic' (or 'bounding') nodes, defined at the time as S (now IP) and NP (now DP). Central to this theory is the idea of locality: movement transformations never directly cross clause boundaries. A fronted *wh*-element instead moves in successive steps, using the specifiers of CP as 'escape hatches', rather than moving in one fell swoop. In the formation of (1a), for example, the *wh*-dependency is local because it is confined to the embedded CP ***what*** *that Mary ate* __ in a first derivational step. Then, the distance between the overt filler and the gap formed in the specifier of the intermediate CP is structurally short as well: only one cyclic node is crossed in each movement step. Unlike other complementizers, however, *wh*-elements that introduce an interrogative clause are taken to occupy the specifier position of CP. This position is occupied by the embedded *wh*-phrase in (1b), and successive movement is hence blocked (because the position cannot be used as an escape hatch). The *wh*-filler is consequently forced to move to sentence-initial position in one fell swoop, and in so doing, it crosses two cyclic nodes: the two IPs corresponding to the matrix and embedded predicates. The sentence in (1b) thereby violates the Subjacency Condition and is ungrammatical, which in turn leads to unacceptability.

Rizzi (1982), however, shares examples from Italian in which a relative pronoun is extracted from an embedded interrogative, thereby crossing more than one cyclic node, see (3). Because *che* 'what' introducing the embedded interrogative clause prevents the specifier position from serving as an escape hatch, fell-swoop movement is forced in this case. The dependency consequently crosses two cyclic nodes (the IPs of *raccontare* 'tell' and *domandarsi* 'wonder', respectively), which violates the Subjacency condition. The sentence should therefore be ungrammatical, contrary to fact.

(3) **Tuo fratello**, **a cui** mi domando che storie abbiano raccontato __, era
your brother to whom me ask.1SG what story have.3PL told was.3SG
molto preoccupato.
very troubled
'Your brother, to whom I wonder which stories they told, was very troubled.'
(taken from Rizzi 1982: 50, his (6b))

Thus, Rizzi (1982) proposes that Subjacency might be subject to parametric variation: what qualifies as a cyclic node may differ from one language to the next, and in Italian (or the Romance languages more generally) it may not be IP (then S), but CP (then $\bar{S}$) that is a hurdle to movement. Not crossing the CP boundary of the matrix clause, the filler in (3) only moves over a single cyclic node on Rizzi's account.

Notice that the sentence in (3) demonstrates a relative clause (RC) dependency. *Wh*-dependencies across island structures, as in (4), are by contrast unacceptable in Italian, which Rizzi (1982) attributes to independent factors related to question formation in Italian: sentences with two or more *wh*-phrases are (presumably) rejected in Italian. Sprouse et al. (2016: 312) note, however, that this restriction is no longer problematic in many varieties of Italian spoken today, and moreover that island types other than interrogative islands do not have a second *wh*-element. Rizzi's (1982) theory (which was transposed to French by Sportiche 1981 and to Spanish by Torrego 1984) thus predicts no difference *ceteris paribus* between *wh*-dependencies and RC-dependencies (in fact, Torrego 1984 claims that *wh*-dependencies in Spanish essentially show the same pattern as RC-dependencies in Italian).

(4) ***Chi** ti domandi chi ha incontrato?
who you ask.2SG who has met
'Who do you wonder who met __?' (taken from Rizzi 1982: 51, his (7a))

In and of itself, the Subjacency Condition does not account for the adjunct and sentential subject island effects identified by Ross (1967).[3] Huang (1982) consequently postulated the *Condition on Extraction Domains* (CED) which blocks movement out of non-complements. In the Barriers framework (Chomsky 1986), adjuncts and subjects are similarly considered *barriers* that impede movement – a notion that in contrast to the formerly postulated cyclic nodes is structurally defined. Thus, if Romanian behaves like other Romance languages, the syntax-based Subjacency and Barriers accounts predict island effects to emerge when *wh*-dependencies are formed across adjuncts and subjects (as they are inherent barriers to movement), and complex NPs (i.e. NPs

[3] Subjacency can however account for other subject island effects, such as * ***Whom*** *did [your interest in __] seem to me rather strange?* (Chomsky 1977: 112, his (179)), where *whom* crosses a DP and an IP boundary.

modified by a relative clause; as per the Subjacency Condition), but not across interrogative clause boundaries (see the discussion on (3) above).

According to Newmeyer (2016: 191), the particular behavior of islands has only played "a fairly minor role in theoretical discussions among formal syntacticians" since the late 1980s, although we can assume that the hypothesis that islandhood may be related to formal constraints persists. Indeed, notable locality principles such as *(featural) Relativized Minimality* (Rizzi 1990; Starke 2001) and the *Minimal Link Condition* (Chomsky 1995) have since been proposed, which constrain movement in terms of the structural properties of the intervening material. However, space limitations prevent us from a detailed description of these accounts. In what follows, we instead discuss non-syntactic accounts that have been developed over the years, which for the most part appeal to certain conditions on discourse packaging and/or processing.

One influential account that proposes that island status is predicted by discourse factors comes from Goldberg (2006; 2013). She claims that the extent to which a construction *backgrounds* its content is indicative of island status,[4] formalized in the BCI Condition in (5). Long-distance dependency formation foregrounds the filler (inside its domain), and so if the gap is located inside a backgrounded construction (i.e. an island), this element is infelicitously both foregrounded and backgrounded at the same time. The 'discourse clash', Goldberg claims, is what leads to unacceptability. The BCI approach finds its roots in earlier works by e.g. Erteschik-Shir (1973) and Erteschik-Shir & Lapin (1979).

(5) **Backgrounded Constructions are Islands (BCI)**
Backgrounded constituents cannot be "extracted" in long-distance dependency constructions.

Another account that attributes island effects to a clash of discourse functions comes from Abeillé et al. (2020), in the form of the *Focus-Background Conflict* in (6). The main difference between these proposals is that the FBC predicts island effects in dependencies with focused fillers only (i.e. *wh*-dependencies). Long-distance dependencies with a non-focused filler (such as in RC-dependencies) by contrast comply with the criterion and are hence not ruled out. That the two dependency types may lead to distinct island effects is corroborated in experimental work on various languages. Abeillé et al. themselves tested PP extractions from subjects and objects in French and English. They found, for both languages, that *wh*-dependencies which originate in subjects yielded lower acceptability ratings than those which originate in objects, but they crucially did not find this same pattern for RC-dependencies. López Sancio (2015) and

[4] Backgrounded elements are non-topical elements which are also not part of the focus domain (Goldberg 2013); that is, they are not what the sentence is about, nor are they located in the informative domain. Note that subjects are not backgrounded, because they are usually the primary topic in a clause; elements within sentential subjects, however, are. Backgroundedness is intended as a gradual notion.

Sprouse et al. (2016) similarly observed subject island effects in Spanish and Italian, but only in items with *wh*-extraction, not in relative clauses. Sprouse et al. moreover report adjunct island effects for *wh*-dependencies, but not for RC-dependencies, in English.[5,6] Kobzeva et al. (2022), finally, report small island effects for *wh*-extraction but not relativization out of Norwegian conditional adjuncts and embedded interrogatives. If these clauses are more backgrounded than their declarative counterparts, *wh*-extraction from them should induce a penalty under the condition in (6).

(6) **Focus-Background Conflict (FBC)**
A focused element should not be part of a backgrounded constituent.

While the results of Kobzeva et al.'s (2022) experiment are in line with this prediction, the authors argue that their findings do not provide unequivocal support for the FBC condition: the interaction effect in the items with *wh*-extraction from embedded interrogatives was only small and possibly due to a ceiling effect in two of the conditions (the short-distance movement cases), while the acceptability patterns in the other two conditions (the long-distance movement cases) were nearly identical. Moreover, the predicted penalty for *wh*-extraction from (backgrounded) adjuncts was inconsistent and did not emerge in the majority of trials. Furthermore, Cuneo & Goldberg (2023) find strong correlations between the acceptability ratings of *wh*- and RC-dependencies in various (English) constructions, including different types of adjuncts.[7] The empirical facts so far thus do not unequivocally support the FBC condition in (6).

---

[5] Sprouse et al. (2016) and Kobzeva et al. (2022) report significant island effects for relativization out of English and Norwegian subject phrases, which is arguably in conflict with the FBC condition. The stimuli for these experiments contained stranded prepositions, however, unlike the experiments on the Romance languages. Abeillé et al. (2020) argue that the items in Sprouse et al. (2016), and by extension in Kobzeva et al. (2022), may have received lower ratings because of an independent constraint against P-stranding in subject phrases.

[6] López Sancio (2015) and Sprouse et al. (2016) do find various significant island effects in other relativization structures, and Kobzeva et al. (2022) find that relativization but not *wh*-extraction from existential RCs triggers an island effect in Norwegian. Since non-focus dependency types comply with the condition in (6), however, these structures are not technically ruled out by the FBC condition, and the source of their unacceptability could lie elsewhere. However, discourse-clash accounts are inherently challenged by parasitic gap constructions, such as *This is **the article** that John read __ [before filing __]*: the first gap is located in a foregrounded main clause while the second gap is located in a backgrounded domain (Cuneo & Goldberg 2023; Coopmans et al. 2024). This inevitably leads to a discourse conflict.

[7] Cuneo & Goldberg (2023) claim to have found evidence for the relation between backgroundedness and island status. Reanalyzing their data, however, Momma & Dillon (2023) show that the relation is correlational but there is no causal inference. They propose an alternative categorization of the data, claiming that "[a]ny linguistically relevant differences between constructions could in principle be a confounding or mediating variable, and hence need to be accounted for in causal analysis" (Momma & Dillon 2023: 7). Goldberg et al. (2024) in turn claim that independently collected gradient data are preferable over theoretically motivated categorical data as a predictor, suggesting that their data address the question *why* a long-distance dependency should be ill-formed.

According to Chaves & Putnam (2020), the acceptability of sentences with a dependency crossing an island boundary is influenced by the structural frequency and plausibility of the sentence. Coherent and prototypical complex structures are easier to process, which translates into increased acceptability. That cognitive constraints on processing may explain the (un) acceptability of long-distance dependencies has also been claimed by Deane (1991), Kluender (1991; 1992; 1998; 2004), Pritchett (1991), Kluender & Kutas (1993a,b), and Hofmeister & Sag (2010) – studies that pioneered the 'experimental turn' in island research. These accounts are based on the observation that the acceptability of island-violating sentences is much higher when factors known to induce processing costs are controlled for. 'Island effects' are claimed to be due to an abundance of performance factors which all independently constrain the processing system: the observed superadditivity can be regarded as the cost of exceeding processing limitations. Thus, processing demands for island-violating sentences are too high, as numerous processing cost-inducing tasks must be tended to (see Hofmeister & Sag 2010 for a review of some factors). The success of sentence processing is then contingent on the sum and difficulty of demands imposed by (parts of) the material, given that cognitive resources spent on a given processing task cannot simultaneously be spent on other processing tasks as well. Note that these accounts naturally predict the often-observed gradience in acceptability ratings, and moreover place the focus on features of the whole sentence – not just the filler or gap site, but also the intervening material that may constrain working memory resources (such as "the frequency, specificity or 'semantic richness', verb class, and argument selection properties of verbs on the extraction path", Hofmeister & Sag 2010: 379). Crucially, some processing accounts maintain that no grammatical rules are necessary at all to explain island phenomena (i.e. *reductionist* accounts, see Sprouse et al. 2012a and Phillips 2013 for discussion). However, processing accounts like Hofmeister and Sag's (2010) are in principle compatible with the existence of syntactic rules (as well as discourse conditions). *Grounded* processing approaches, by contrast, consider island effects to be caused by formal syntactic constraints, but appeal to properties of the parser as well. In particular, these approaches consider the grammatical rules to have developed as a consequence of processing difficulty, in that the processing burdens have grammaticized into island constraints.

In summary, island effects and island insensitivity have been explained from different theoretical perspectives. One theoretical camp claims that specific syntactic constraints block particular types of (*wh*-)extraction. Other researchers propose that the unacceptability of (certain) long-distance dependencies is due to conflicts in the discourse representation, either because extraction happens from a backgrounded domain (Goldberg 2006; 2013), or because an extracted *wh*-element is both focused and non-focused at the same time (Abeillé et al. 2020). A third group of researchers submit that, when parsing long-distance dependencies, various interacting performance factors may cause the parser to 'break down', so to say. Regarding the *wh*-dependencies we tested in the experiment to be presented in Section 3, most accounts

predict island sensitivity in Romanian, with the exception of the Subjacency account (under the assumption that the cyclic nodes of Romanian are CP and DP, like in Italian, Spanish, and French); this account does not predict island effects to emerge when the gap site is located inside an interrogative clause. We discuss our linking hypotheses in more detail in Section 3.1.

2.2 Experimental work on island phenomena

In most experimental work on island phenomena, ratings of acceptability are collected for a larger set of experimental items from a larger sample of (non-linguist) participants. Many experiments cross two particular factors to isolate island effects: distance (i.e. dependency length) and the type of construction, as the usual definition of an 'island effect' involves a decrease in acceptability of sentences with a long-distance dependency with its gap located inside an alleged island construction (Sprouse & Villata 2021). Thus, a 2×2 design that manipulates i) the presence of a long-distance dependency, and ii) the presence of an alleged island, can be used to investigate island effects. It is worth noting here that the two factors may influence acceptability independently. The additional processing cost of keeping a long-distance dependency active in memory, compared to a short-distance dependency, may negatively affect the acceptability of a sentence, and the presence of an alleged island structure, typically a complex structure, may similarly impact sentence ratings. Therefore, we can only speak of an 'island effect' when the decrease in acceptability occurs *on top of the individual effects of the factors*. That is, statistically, we are interested in super-additivity, i.e. the interaction effect between the two factors.

Consider the sentences in (7) for illustration, in which the two factors *construction* (non-island vs. island) and *distance* (short vs. long distance) are crossed. Each factor might be associated with its own cost, but in the case of an island effect the interaction between the two factors is expected to give rise to a super-additive effect, which according to linguistic theory takes the shape of (7d) receiving lower acceptability ratings than predicted by the combined effects of construction and distance.

(7) a. **Who** _ thinks [that John bought a car]? (non-island/short)
b. **What** do you think [that John bought _]? (non-island/long)
c. **Who** _ wonders [whether John bought a car]? (island/short)
d. **What** do you wonder [whether John bought _]? (island/long)
(taken from Sprouse et al. 2016: 314, their (12))

The size of island effects is usually reported as the difference-in-differences (DD) score on standardized acceptability ratings. The DD-score is computed as follows (see Maxwell & Delaney 2003: 684–687 for more detailed description). A first difference score (D1) is calculated by subtracting the acceptability score of the island/long condition (7d) from the non-island/long condition (7b). D1 quantifies the effect of having an island structure in sentences with

a long-distance dependency. A second difference score (D2) is calculated by subtracting the mean acceptability rating of the island/short condition (7c) from the non-island/short condition (7a). D2 quantifies the effect of having an island structure in sentences with a short-distance dependency. Finally, the DD score is calculated by subtracting D2 from D1, and represents the difference between differences related to the two factors.

Since Sprouse (2007), the factorial design discussed above served as the basis for acceptability judgment experiments in investigations of island phenomena in various languages. Here, we focus on the Romance languages, and on Romanian in particular. We first discuss experimental studies on island effects across Romance languages in Section 2.3 and continue by providing novel data from Romanian.

2.3 Experimental studies on island effects in Romance languages

We summarize the main findings from several experiments on islands in the Romance languages in this section, but this is by no means an exhaustive review. Recall that theoretical linguists from the early 1980s claimed that Romance languages do not obey the same island constraints as English does (Rizzi 1982 for Italian, Sportiche 1981 for French, Torrego 1984 for Spanish). Thus, extraction from interrogative embedded clauses in particular is expected to be more acceptable in Romance languages, at least when a RC-dependency is involved, but the theory adopted in these studies can in principle be extended to other types of islands and to *wh*-dependencies as well (see also Sprouse et al. 2016). This section discusses experimental work investigating different types of island effects in the Romance languages.

Sprouse et al. (2016) conducted a series of judgment experiments to gauge the acceptability of extraction from four different island types (interrogative, complex NP, subject, and adjunct islands) in Italian and English. They tested both RC- and *wh*-dependencies, but in this paper we focus on the items with a *wh*-dependency. For each island type, the authors tested four conditions manipulated across the factors *distance* (short vs. long) and *construction* (island vs. non-island), with eight lexicalizations per experiment. Importantly, the design of the items with a subject island was different from the other island types, because Italian does not allow preposition stranding (see also Abeillé et al. 2020), and without a preposition following the object, there might be local ambiguity. Thus, Sprouse et al. (2016) tested the acceptability of extraction of simple subjects and objects, and from complex subjects and objects, see (8).

(8) a. **Chi** pensi che il quadro raffiguri __ ? (simple/object)
who think.2SG that the painting depict.SUBJ.3SG
'Who do you think that the painting portrays?'

b. **Chi** pensi che __ abbia dipinto il quadro? (simple/subject)
who think.2SG that have.SUBJ.SG painted the painting
'Who do you think has painted the painting?'

c. **Di chi** pensi che il quadro sul muro raffiguri (complex/object)
of who think.2SG that the painting on-the wall depicts
[la nascita __]?
the birth
'Who do you think the painting on the wall depicts the birth of?'

d. **Di chi** pensi che [il quadro __] raffiguri la nascita (complex/subject)
of who think.2SG that the painting depicts the birth
di Venere?
of Venus
'Who do you think the painting of depicts the birth of Venus?'
(adapted from Sprouse et al. 2016: 320, their (23))

Sprouse et al. (2016) tested 193 native speakers of Italian (spread across four experiments) and found super-additive effects for all four island types with a *wh*-dependency in Italian, with varying effect sizes (complex NP islands DD 0.89; adjunct islands DD 1.31; subject islands DD 1.37; interrogative islands DD 1.69).

A similar picture emerges in experimental work on Spanish. López Sancio (2015) replicated Sprouse et al.'s (2016) experiment for Spanish in his Master's thesis, and asked 217 native speakers of Spanish (distributed over four experiments) to rate stimulus items, testing two dependency types (*wh*- and RC-dependencies) and four island types (interrogative, complex NP, subject, and adjunct islands). The stimulus items in this experiment were constructed following Sprouse et al.'s (2016) items, with a slight adaptation to the subject island items: the subject and object DPs were made non-specific (e.g. *¿De quién crees que varios escándalos han suscitado rumores sobre el candidato?* 'Of whom do you think various scandals have sparked rumors about the candidate?', cf. (8)). Regarding the items with a *wh*-dependency, López Sancio reports significant island effects for all four island types (complex NP islands DD 1.41; adjunct islands DD 1.71; subject islands DD 0.80; interrogative islands DD 1.65).

Island effects in Spanish are also reported by Pañeda et al. (2020), who tested *wh*-dependencies across the same four constructions, but who used a speeded acceptability judgment experiment instead. Participants in Pañeda et al.'s experiment were asked to evaluate items as fast as possible after fast-paced word-by-word presentation on a computer screen, on a binary scale, instead of a 7-point scale. Regarding the subject island items, however, the experimental design used in Pañeda et al. was crucially different from that used in Sprouse et al. (2016) and López Sancio (2015). It instead crossed the familiar factors *distance* and *construction*, see (9): the filler either originated in the matrix clause (9a,b) or in the embedded clause (9c,d), and was only extracted from a subject in (9d). Notice that the conditions have different modifiers; this was done so that each condition has a comparable number of words in it.

(9) a. ¿**Quíen** cree que el discurso ofendió tanto a Julia ayer? (non-island/short)
who believe.3SG that the discourse offended so.much to Julia yesterday
'**Who** believes that the discourse offended Julia so much yesterday?'

b. ¿**Quíen** cree que el discurso del director ofendió a Julia? (island/short)
who believe.3SG that the discourse of.the director offended to Julia
'**Who** believes that the director's discourse offended Julia?'

c. ¿**Quíen** crees que __ ofendió tanto a Julia con el discurso? (non-island/long)
who believe.2SG that offended so.much to Julia with the discourse
'**Who** do you believe offended Julia so much with the discourse?'

d. ¿**De quíen** crees que [el discurso __] ofendió tanto a Julia? (island/long)
of who believe.2SG that the discourse offended so.much to Julia
'Of **who** do you believe that[the discourse offended Julia so much?'
(adapted from Pañeda et al. 2020: 11, their (10))

Pañeda et al. (2020) tested 80 native speakers of Spanish, who each saw all island types, and found super-additive effects for the subject, adjunct, and interrogative islands,[8] with subject island effects larger than the other two island types, but no clear evidence for island effects in the items with a complex NP island.[9] Notice that Pañeda et al.'s findings do not fully corroborate López Sancio's (2015) findings: the subject island effect that López Sancio reports is the weakest of all four investigated island types, whereas Pañeda et al. found the strongest interaction effect in the items with a subject island. Pañeda et al. argue that these differences might be due to the alternative design used to test subject islands, as the design in (9) might overestimate the island effect size because the interaction term may also contain independent effects of subextraction from complex phrases (as pointed out by Kush et al. 2018). The design in (8), by contrast, might underestimate the effect size, since the complexity of the added DPs in (8c) and (8d) may reduce the acceptability ratings to the extent that floor effects reduce the interaction effect (Kush et al. 2018).

[8] Further distinctions between different types of interrogative islands in Spanish are reported by Pañeda & Kush (2021) and Rodríguez & Goodall (2023): *whether*-islands yield smaller effects than *when*-islands. This distinction holds for Romanian as well (Stoica & Schoenmakers 2024).

[9] The speeded judgment experiment collected binary judgments of acceptability, so Pañeda et al. (2020) do not compute DD scores. Instead, they report the means of the posterior distributions: –6.04 log-odds for the subject island effect, –3.27 for the adjunct island effect, –3.65 for the interrogative island effect, and –1.36 for the complex NP island effect.

Abeillé et al. (2020) designed a series of experiments to test RC- and *wh*-extraction from subject islands in English and French. Abeillé et al. crossed the factors *grammatical function* (subject vs. object) and *extraction* (PP vs. no extraction),[10] resulting in yet another design to test subject islands, see (10). Still, Abeillé et al.'s design is rather close to that in (8), the main difference being that the non-extracted subjects and objects remain in situ in (10a,b) but are fronted in (8a,b), and the subjects and objects are swapped between conditions in (10). The situation described in (10) is the same across conditions.

(10) a. Est-ce que le footballeur adore la couleur de la (object/no extraction)
is-it that the football.player loves the color of the
décapotable à cause de sa luminosité?
convertible at cause of its luminance
'Does the football player love the color of the convertible because of its luminance?'

b. Est-ce que la couleur de la décapotable enchante le (subject/no extraction)
is-it that the color of the convertible delights the
footballeur à cause de sa luminosité?
football.player at cause of its luminance
'Which convertible does the color delight the football player because of its luminance?'

c. **De quelle décapotable** est-ce que le footballeur (object/PP extraction)
of which convertible is-it that the football.player
adore [la couleur à cause de sa luminosité?
loves the color at cause of its luminance
'Of which convertible does the football player love the color because of its luminance?'

d. **De quelle décapotable** est-ce que [la couleur _] (subject/PP extraction)
of which convertible is-it that the color
enchante le footballeur à cause de sa luminosité?
delights the football.player at cause of its luminance
'Of which convertible does the color delight the football player because of its luminance?'

(adapted from Abeillé et al. 2020: 16–17, their (32))

Abeillé et al. (2020) tested 47 native speakers of French, and found significant subject island effects in the *wh*-dependency experiment. As Abeillé et al. posted their research data online, we could calculate the corresponding DD-score for this experiment, which was 0.68, but note that this DD score is not fully comparable to those reported for the other experiments because of the

[10] Abeillé et al. (2020) also included an ungrammatical condition as a baseline. They find that sentences with *wh*-extraction from French subjects were not rejected to the same extent as their ungrammatical counterparts, and so they conclude that these constructions are infelicitous rather than ungrammatical.

alternative experimental design. Crucial for the authors' purposes was that they did not find significant island effects in the items with a RC-dependency, and that the patterns for French and English are highly similar. We, however, focus on the fact that their results corroborate the findings for Italian (Sprouse et al. 2016) and Spanish (López Sancio 2015; Pañeda et al. 2020), in that *wh*-extraction from subjects received worse ratings than would be expected, at least to the extent that the experimental designs can be compared.

The experimental studies discussed above show that island effects are consistently reported for *wh*-dependencies in various island constructions in the Romance languages that have been investigated so far. The reported island effect sizes are different, however, which in many cases is (presumably) due to differences in the design of the stimulus materials. Our experiments (reported in Section 3) followed the design used in Sprouse et al. (2016) and López Sancio (2015) to a large extent; the main difference is that we tested two island types with a *wh*-dependency per experiment, while Sprouse et al. and López Sancio tested one island type with a *wh*-dependency and another with a RC-dependency per experiment. **Table 1** reproduces the island effect sizes that Sprouse et al. and López Sancio found in their items with a *wh*-dependency, per construction type; the findings from Pañeda et al. (2020) and Abeillé et al. (2020) are not represented here because of their alternative research designs. Our investigation contributes data from *wh*-extraction from four different island types in Romanian to the archipelago (in contrast with Stoica & Schoenmakers 2024, who only tested interrogative islands). We report on our experiments in Section 3, but before that, we turn to the literature on islands in Romanian.

	Language	Adjuncts	Complex NPs	Interrogatives	Subjects
Sprouse et al. (2016)	Italian	1.31	0.89	1.69	1.37
López Sancio (2015)	Spanish	1.71	1.41	1.65	0.80

Table 1: DD-scores (based on *z*-scores) per island type reported in previous experimental work on *wh*-extraction in Romance languages with a comparable experimental design.

2.4 Islands in Romanian

As noted, island phenomena are relatively understudied in Romanian. The theoretical literature largely acknowledges the presence of island effects with respect to adjuncts, interrogative clauses, and complex NPs (Pană Dindelegan 2013), illustrated with *wh*-dependencies in (11).[11]

[11] Romanian is a pro-drop language, and is commonly claimed to have the basic word order VSO (e.g. Di Sciullo et al. 2003). *Ion* in (11b) is a post-verbal subject followed by the gap associated with the indirect object filler.

The example in (12), provided by the native-speaking author of this article, is added to illustrate that extraction from subject islands is presumably also unacceptable in Romanian.

(11) a. *__Ce__ a întârziat Ion [pentru că a pierdut __]? (adjunct island)
what has delayed Ion because has lost
'What was Ion late because he lost?'

b. *__Cui__ nu-ți place [ce i-a (interrogative island)
who.DAT not-CL.DAT.2SG likes what CL.DAT.3SG-has
spus Ion __]?
said Ion
'To whom don't you like what Ion said?'

c. *__Cine__ mi-a arătat Maria [cartea pe care a (complex NP island)
who CL.DAT.1SG=has shown Maria book.the in which has
scris-o __]?
written=CL.ACC.F.3SG
'Who did Mary show me the book in which __ has written?'
(taken from Pană-Dindelegan 2013: 544, her (22))

(12) *__Despre cine__ crezi că a devenit best-seller (subject island)
about whom think.2SG that has become best-seller
[o carte __] ?
a book
'About whom do you think the book became a best-seller?'

Certain factors have been argued to alleviate island effects in Romanian (and the Romance languages more generally). Based on the fact that Romanian allows multiple fronted *wh*-elements (or, it allows a multiply filled COMP), Comorovski (1986: 177) originally predicted that the language "will not obey any form of the *Wh*-island Constraint"; Rudin (1988) similarly claimed that Romanian allows extraction from interrogative clauses, illustrated in (13). Later, however, Comorovski (1989) and Alboiu (2002) argued that this claim only holds for D(iscourse)-linked elements (Pesetsky 1987), i.e. elements whose reference can be reconstructed from the existing discourse context. Fillers that contain the phrase *which*, for instance, are contextually more salient than fillers that contain the phrase *what*, and their extraction from island constructions is consequently more acceptable, as in (14). Sevcenco (2006) suggests that D-linking ameliorates extraction from complex NPs as well, see (15). Notice that the extracted object in (15a) is clitic-doubled because of the form of the filler (Dobrovie-Sorin 1990); the sentence in (15b) is ungrammatical regardless of the presence of this clitic.

(13) **Pentru care clauză** vrei să afli cine nu a decis încă [ce
for which paragraph want.2SG to learn who not has decided yet what
va vota __]?
will.3SG vote
'For which paragraph do you want to learn who has not yet decided what they will vote?'

(taken from Rudin 1988: 458, her (22a))

(14) Despre { **care** / * **ce** } știi [cine i-a povestit __]?
about which what know.2SG who DAT.has told
'Which one/what do you know who told him about?'

(taken from Comorovski 1989: 82, her (6))

(15) a. **Pe care polițist** au lansat [zvonul că l-a mituit
on which policeman have launched rumor-the SJV him.CL.ACC = has bribed
primul ministru __]?
prime.the minister
'Which policeman did they launch the rumor that the prime minister bribed?'

(taken from Sevcenco 2006: 167, her (63))

b. *__Pe cine__ au lansat [zvonul că (l)-a mituit primul
on whom have launched rumor-the SJV (him.CL.ACC =)has bribed prime.the
ministru __]?
minister
'Who did they launch the rumor that the prime-minister bribed?'

Stoica & Schoenmakers (2024) experimentally investigated the claim that D-linking alleviates island effects that are due to extraction from Romanian interrogative clauses headed by *dacă* 'whether' and *când* 'when'. In their experiment, they manipulated the filler type between bare *wh*, *what* N, and *which* N fillers, predicting a gradual increase in acceptability, the more D-linked the filler is. Although Stoica and Schoenmakers found significant island effects in each sub-experiment, their predictions regarding the discourse embedding of the fillers were not borne out entirely. That is, in line with their predictions, they found that *which* N fillers led to weaker island effects than bare *wh* fillers; however, *what* N fillers led to unexpectedly strong island effects – stronger, in fact, than *which* N fillers. The authors conclude that D-linking accounts cannot straightforwardly explain their results. However, they also note that long-distance *wh*-extraction from non-island constructions (16) yielded unexpectedly high ratings compared to the other conditions, especially in the items with a *what* N filler, which seems to have been the driving factor behind the differences in the island effect sizes, even though no theory would a priori predict this pattern. We return to this observation in Section 3.6.2. The present paper only

investigated bare *ce* 'what' fillers, in line with what López Sancio (2015) and Sprouse et al. (2016) did for Spanish and Italian,[12] but unlike Stoica & Schoenmakers (2024) we tested four different island types.

(16) **Ce carte** crede profesoara [că a citit Matei __]?
'Which book does the teacher think that Matei read?'

Finally, Sevcenco (2006) claims that Romanian island effects are also alleviated when a subjunctive is used in a complex NP, see (17), and this appears to hold for other island types as well: the sentences in (18) with an interrogative and a subject island become unacceptable when the subjunctive is changed to an indicative (see also Baunaz & Puskás 2014). As far as we know, this difference has not yet been investigated experimentally. The alleged islands in our experiment were in the indicative.

(17) **Cine** ai făcut [sugestia ca mâine să se prezinte __ de
who have.2SG made suggestion-the that tomorrow SUBJ REFL.CL present of
urgență la șef]?
emergency at boss
'Who did you make the suggestion that should show up by all means at the boss?'
(taken from Sevcenco 2006: 167, her (64))

(18) a. **Ce** te întrebi [dacă trebuie să cumperi __]?
what 2.SG.CL wonder whether must SUBJ buy
'What do you wonder whether you must buy?'
(taken from Dobrovie-Sorin 1990: 354, footnote 8, her (ii))

b. **Ce** ar fi bine [să manânce copiii __ deseară]?
what would be good SUBJ eat children.the evening.this
'What would it be good that the children eat this evening?'
(taken from Di Sciullo et al. 2003: 285, their (17b))

In sum, it has been noted that Romanian *wh*-questions may be sensitive to various types of islands, with factors such as D-linking and verbal mood playing a role. With the exception of Stoica & Schoenmakers (2024), to our knowledge there have been no experimental investigations of island effects in Romanian. The next section reports on two experiments that followed the same basic design as in López Sancio (2015) and Sprouse et al. (2016), measuring the acceptability of *wh*-dependencies across four different types of island structures.

[12] Sprouse et al. (2016) also test complex *wh*-phrases that contain a noun (e.g. *which car*), but only for English.

3 Experiment

We conducted two acceptability judgment experiments to investigate island sensitivity in Romanian, using *wh*-dependencies and four island types: conditional adjunct, complex NP, interrogative (*whether*), and subject islands. Our experiments adopted the same factorial design as in Sprouse et al. (2016) and López Sancio (2015). Earlier experimental studies on languages closely related to Romanian (Section 2.3) and claims about Romanian (Section 2.4) lead us to believe that bare *wh*-questions in which the four types of island boundaries are crossed (with the embedded clause in the indicative) are unacceptable in Romanian. This outcome is predicted by (most of) the linguistic theories discussed in Section 2.1, on which we elaborate in the next subsection.

3.1 Linking hypotheses

Under the assumption that the cyclic nodes in Romance languages are CP and DP (extrapolating from Sportiche 1981, Rizzi 1982, and Torrego 1984), Chomsky's (1973; 1977) Subjacency Condition predicts that interrogative islands are unproblematic in Romanian, whereas complex NP islands resist extraction. Specifically, while initial movement to the edge of the embedded CP (within the complex noun phrase) is allowed, subsequent movement into sentence-initial position would cross two cyclic nodes (*viz.* the CP and DP boundaries). This would consequently be disallowed under Subjacency. With interrogative clauses, by contrast, only one cyclic node is crossed at a time (see the discussion on (3)). Experimental data from Spanish and Italian (López Sancio 2015; Sprouse et al. 2016), however, suggest that these languages do in fact resist *wh*-extraction over interrogative clause boundaries. Further, extraction from adjuncts and subjects is expected to be disallowed according to the CED (Huang 1982) and the Barriers framework (Chomsky 1986), as adjuncts and subjects are both non-complement constituents.

According to Goldberg's (2006; 2013) BCI condition (see (5)), all island constructions that we investigate are backgrounded constituents and resist extraction, as long-distance dependency formation foregrounds the filler while the gap is located inside a backgrounded domain. Similarly, Abeillé et al.'s (2020) *Focus-Background Conflict* predicts that extraction from the island constructions investigated here yields unacceptability, given that *wh*-dependency formation involves focus movement. The filler-gap dependency would therefore simultaneously be focused (because the filler is a *wh*-phrase) and non-focused (because the gap site is located inside an island).

Processing accounts argue that the combination of multiple factors that induce a processing cost impede parsing in structures with a *wh*-dependency across an island boundary. The super-additive effect observed in experimental research is then viewed as the cost of exceeding cognitive resources. All else being equal, processing accounts also predict island effects in our

experiments, but these island effects are expected to correlate with individual participants' working memory capacity (Sprouse et al. 2012a). Since we did not include a working memory task in our experiment, which would help separate the predictions of processing accounts from those from the syntax-based accounts, we instead refer the reader to the discussion between Hofmeister et al. (2012a,b; 2014) and Sprouse et al. (2012a,b) (see also Michel 2014 and Pham et al. 2020).

To recap, the Subjacency account that adopts parametrization predicts sensitivity in the items with adjunct, subject, or complex NP islands, but not in the items with interrogative islands. Accounts that are based on discourse clashes or resource overload predict island sensitivity in all construction types. Earlier experimental work on Romance languages found island effects in each of the item sets that we tested (López Sancio 2015; Sprouse et al. 2016). Given that our experiments employ the same design and also test a Romance language, we expect to find similar results. That is, we expect to find island effects as defined in Sprouse & Villata (2021) as an interaction effect between the factors *distance* (long vs. short) and *construction* (island vs. non-island) in the adjunct, embedded interrogative, and complex NP item sets, and as an interaction effect between *gap site* (subject vs. object) and *complexity* (simple vs. complex) for subject islands (Sprouse et al. 2016).

3.2 Participants

211 native speakers of Romanian without language deficits completed one of two versions of the questionnaire, either with adjunct and complex NP islands or with subject and interrogative islands, and received course credit upon completion. The participants were unfamiliar with syntactic theorizing or island phenomena. Data from participants were removed if their average score for the grammatical filler items was lower than 4 (on a 7-point scale), if their average score for the ungrammatical filler items was higher than 4, or their standard deviation was lower than 1. This was the case for seven participants. We entered data from 103 participants in the adjunct/complex NP-island task (M_{age}: 20.6, SD: 3.1, range 18–46) and 101 participants in the subject/whether-island task (M_{age}: 21.7, SD: 5.1, range 18–47) into statistical analysis.

3.3 Materials

We conducted two acceptability judgment tasks, one testing adjunct islands and complex NP islands, the other testing interrogative islands and subject islands. Both tasks followed the factorial design used in Sprouse et al. (2016), which isolates the factors *distance* (short vs. long) and *construction* (island vs. non-island). Sample items for each island type are given in (19) – (22); here we added the boldface font, brackets, and underscores as a visual aid. The stimuli followed the same pattern as in Sprouse et al.'s experiment, but were adapted in order to control

for definiteness and tense (the full stimulus list can be found in our OSF repository). All islands were in the indicative. All adjunct island items contained conditional *dacă* 'if' clauses, and all interrogative island items contained *dacă* 'whether' clauses. As discussed in Section 2.1, we also follow Sprouse et al.'s alternative design for the subject island items, which instead crossed the factors *gap site* (subject vs. object) and *complexity* (simple vs. complex). The reason to use this design is that the Romance languages do not allow preposition stranding, and the alternative design minimizes the possibility that participants interpret the sentences on the wrong reading (Sprouse et al. 2016).[13] Notice that the subjects and objects in the complex conditions in (22) are both modified by a PP. The alternative design moreover controls for subextraction (unlike the one adopted in e.g. Kush et al. 2018; Kobzeva et al. 2022): by manipulating the extraction domain in the complex conditions, it permits a comparison of subextraction from subjects vs. subextraction from constituents in different positions. The embedded clause in the condition with subextraction from a complex object (20b) was VOS for processing reasons. If the sentence were *Despre ce crezi că a iscat raportul despre dictator o discuție __ ?* 'What do you believe the report about the dictator sparked a conversation about?', the second *despre*-PP before the gap site might induce higher processing costs, because it introduces another DP before the filler-gap dependency is resolved. Moreover, this maximizes similarity with the condition subextracting from a complex subject (22d) in terms of dependency length. The embedded clauses in all other conditions were VSO sentences. An anonymous reviewer points out that the subjects in the subject island item set were always definite, while the objects were always indefinite. Since definite DPs are less transparent than indefinite DPs ('definite islands' in Ross 1967; see Neal & Dillon 2021 and Shen & Lim 2024 for experimental investigations), a putative island effect may include a definiteness effect. We return to this confound in Section 3.6.1, noting here that the effect sizes in the cited experimental investigations were only small.

(19) ADJUNCT ISLAND

a. **Cine** __ crede [că a fumat actorul o țigară pe scenă]?
who thinks that has smoked actor.the a cigarette on stage
'Who thinks the actor smoked a cigarette on stage?' (non-island/short)

[13] An anonymous reviewer asks if there could be some remaining ambiguity, such that participants integrate *despre ce/cine* 'about what/whom' with the matrix verb (e.g. 'about what/whom do you believe that …' in (20b) and (20d)). This is technically possible, but only temporarily, because the embedded clauses contain two DPs (*rapportul* 'the report' and *o discuție* 'a conversation') which need to be allocated according to the argument structure of the embedded verb. Under this interpretation, the structures in (20b) and (20d) are ungrammatical, because the theta-criterion is violated (Chomsky 1981). This should lead to a strong main effect of the factor *complexity*, but we deem it unlikely that participants analyze the sentences in this way.

b. **Cine** __ se îngrijorează [dacă a fumat actorul o ţigară pe scenă] (island/short)
who REFL.3SG worries if has smoked actor.the a cigarette on stage
'Who is worried if the actor smoked a cigarette on stage?'

c. **Ce** crezi [că a fumat actorul __ pe scenă]? (non-island/long)
what think that has smoked actor.the on stage
'What do you think the actor smoked on stage?'

d. **Ce** te îngrijorezi [dacă a fumat actorul __ pe scenă]? (island/long)
what REFL.2SG worry if has smoked actor.the on stage
'What are you worried if the actor smoked on stage?'

(20) COMPLEX NP ISLAND

a. **Cine** __ a auzit [că a copt Maria un tort]? (non-island/short)
who has heard that has baked Maria a cake
'Who heard that Maria baked a cake?

b. **Cine** __ a auzit [zvonul că a copt Maria un tort]? (island/short)
who has heard rumor.the that has baked Maria a cake
'Who heard the rumor that Maria baked a cake?'

c. **Ce** a auzit îngrijitorul [că a copt Maria __]? (non-island/long)
what has heard janitor.the that has baked Maria
'What did the janitor hear that Maria baked?

d. **Ce** a auzit îngrijitorul [zvonul că a copt Maria __]? (island/long)
what has heard janitor.the rumor.the that has baked Maria
'What did the janitor hear the rumor that Maria baked?'

(21) INTERROGATIVE ISLAND

a. **Cine** __ crede [că a mâncat Paul prăjitura]? (non-island/short)
who believes that has eaten Paul cake.the
'Who believes Paul ate the cake?'

b. **Cine** __ se întreabă [dacă a mâncat Paul prăjitura]? (island/short)
who REFL.3SG wonder whether has eaten Paul cake.the
'Who wonders whether Paul ate the cake?'

c. **Ce** crede mama [că a mâncat Paul __]? (non-island/long)
what believes Mom that has eaten Paul
'What does Mom believe Paul ate?'

d. **Ce** se întreabă mama [dacă a mâncat Paul __]? (island/long)
what REFL.3SG wonder Mom whether has eaten Paul
'What does Mom wonder whether Paul ate?'

(22) SUBJECT ISLAND

a. **Ce** crezi [că a iscat raportul despre dictator __]? (simple/object)
what believe that has sparked report.the of dictator
'What do you believe the report about the dictator sparked?'

b. **Despre ce** crezi [că a iscat o discuție ___ (complex/object)
of what believe that has sparked a conversation
raportul despre dictator]?
report.the of dictator
'What do you believe the report about the dictator sparked a conversation about?

c. **Ce** crezi [că a iscat __ o discuție despre (simple/subject)
what believe that has sparked a conversation of
istorie]?
history
'What do you believe caused a conversation on history?'

d. **Despre cine** crezi [că a iscat raportul ___ o (complex/subject)
of who.ACC believe that has sparked report.the a
discuție despre istorie]?
conversation of history
'Who do you believe the report about __ caused a discussion on history?'

For both sub-experiments, eight items per island type (i.e. sixteen experimental items in total) were distributed over four experimental lists according to a Latin Square design (i.e. eight experimental lists in total). Following Sprouse et al. (2016), twelve grammatical and twenty ungrammatical fillers were added to each experimental list, resulting in a 2:1 ratio of experimental and filler items in each list. The number of grammatical and ungrammatical fillers was chosen on the assumption that three of the four experimental conditions are acceptable, and one unacceptable. Each experimental list thus contained 24 (presumed) acceptable and 24 (presumed) unacceptable items. Sample grammatical and ungrammatical filler items are given in (23); (23b) is ungrammatical because the adjunct *până în ziua de azi* 'up until today' intervenes between the fronted *wh*-phrase and the verb (Alboiu 2002).

(23) a. Când ți-a povestit Adelina că a obținut bursa
when CL.DAT.2SG = has told Adelina that has obtained scholarship.the
pentru anul viitor?
for year coming
'When did Adelina tell you she got the scholarship for next year?'

b. *Ce audiobook până în ziua de azi a ascultat fiecare antreprenor?
which audiobook up until day of today has listened every entrepreneur
'Which audiobook has every entrepreneur listened to, to this day?'

The items were pseudorandomized using the software Mix (van Casteren & Davis 2006) and each list was presented in one of two distinct orders. Each list started with three practice items that were excluded from statistical analysis. The experiments were conducted in Qualtrics.

3.4 Procedure

The experiments were online questionnaires that started with a general information and consent form. Participants provided demographic data and were assigned to one of the experimental lists at random. They were instructed to judge on a 7-point scale how natural the target sentences would sound when uttered by a native speaker of Romanian, where 1 represented *foarte rău* 'very bad' and 7 *foarte bine* 'very good'. The instructions mentioned that participants could rely on their first intuition and that there are no right or wrong answers. Participants could provide their answer using radio buttons, which were presented with the corresponding value (from 1 to 7). The verbal labels *foarte rău* and *foarte bine* were presented at the extremes only (1 and 7). Each trial was presented on a new page. The experiments started with three practice trials before continuing to the actual experimental task, and concluded with the opportunity for participants to provide qualitative comments about the experiment.

3.5 Analysis

We performed a series of linear mixed effect models on the *z*-transformed judgment scores using R (version 4.2.3, R Core Team 2022) and the *lme4* package (Bates et al. 2015). For the items with adjunct, complex NP, and interrogative islands, we entered the two-level variables *distance* (short vs. long) and *construction* (island vs. non-island) into the models as fixed effects; for the items with subject islands we entered the variables *gap site* (subject vs. object) and *complexity* (simple vs. complex). All variables were encoded using deviation contrasts (–0.5; 0.5). We started out with base models that included the predictors of interest and their interaction. By-item and by-participant random intercepts were included in the model only if they improved the model fit (assessed by Likelihood Ratio Tests), followed by by-participant random slopes. The models did not include by-item random slopes, because the experiments tested only eight items per island type (i.e. two per condition). The final model for the adjunct island item set included a by-item random intercept. The final model for the complex NP island item set was a simple linear model, because inclusion of random effects did not improve the model fit. The final model for the subject island item set included by-participant and by-item random intercepts. The final model for the interrogative island item set included by-item and by-participant random intercepts as well as a by-participant random slope for the effect of *construction*. In case an interaction effect

was significant, we ran post-hoc analyses with Bonferroni-corrected *p* values using the *emmeans* package (Lenth 2024).

3.6 Results

The mean (non-standardized) judgment scores and standard deviations for the four conditions are given for each island type in **Tables 2** and **3**. The mean rating for the grammatical filler items was 6.1 (SD = 1.5) and the mean rating for the ungrammatical filler items was 2.3 (SD = 1.7). The (*z*-transformed) data are visually represented in **Figures 1** and **2**. In what follows, we discuss the results for each island type in turn.

	non-island/ short	island/short	non-island/ long	island/long
Adjunct islands	4.22 (2.08)	4.28 (1.80)	5.00 (2.13)	2.49 (1.72)
Complex NP islands	5.23 (1.80)	5.20 (1.82)	4.64 (2.01)	1.35 (0.79)
Interrogative islands	4.55 (2.13)	5.04 (1.76)	6.19 (1.28)	2.43 (1.65)

Table 2: Mean ratings and SDs (between brackets) by condition in adjunct, complex NP, and interrogative islands.

	simple/object	complex/object	simple/subject	complex/subject
Subject islands	5.37 (1.84)	2.94 (1.94)	5.54 (1.56)	2.15 (1.55)

Table 3: Mean ratings and SDs (between brackets) by condition in subject islands.

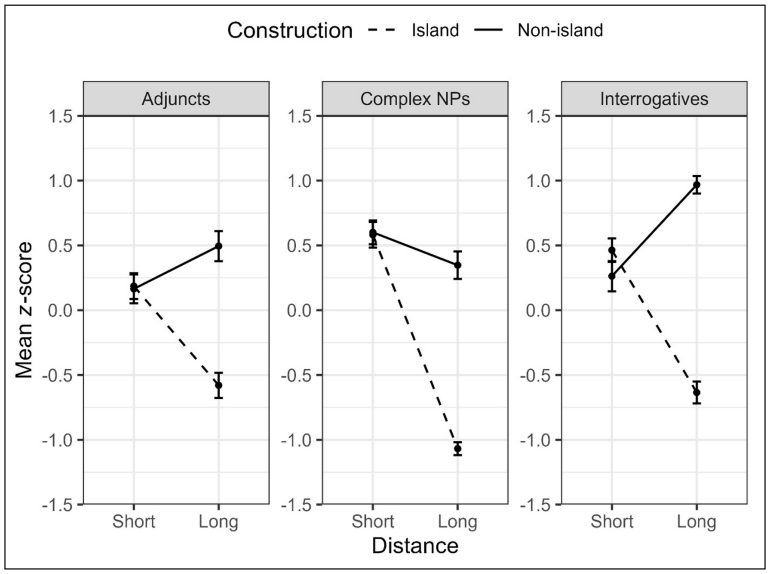

Figure 1: Interaction plots for adjunct, complex NP, and interrogative islands in Romanian (*wh*-dependencies).

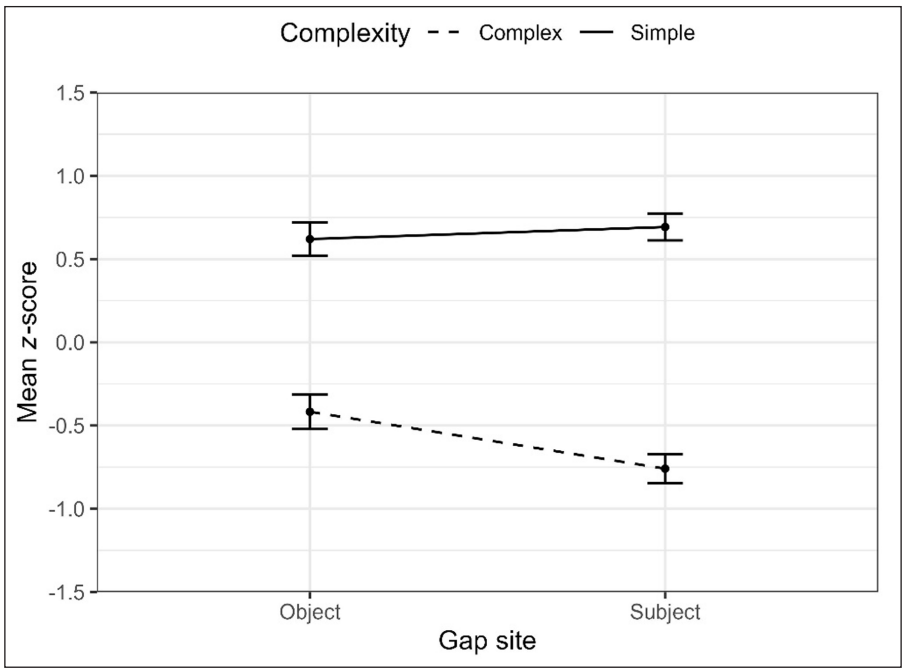

Figure 2: Interaction plots for subject islands in Romanian (*wh*-dependencies).

3.5.1 Adjunct islands

We find a significant island effect in the items with an adjunct island, as the interaction between the two factors was significant ($\beta = 1.11$, $SE = 0.10$, $t = 10.97$, $p < .001$). The statistical model moreover yielded a significant main effect of *distance* ($\beta = –.21$, $SE = 0.05$, $t = –4.21$, $p < .001$) and *construction* ($\beta = 0.52$, $SE = 0.05$, $t = 10.34$, $p < .001$). On average, long-distance dependencies received lower scores than short-distance dependencies, and non-island constructions higher scores than island constructions. Post-hoc analyses reveal that the island/long condition was rated significantly worse than the other three conditions ($t = 10.71$ for island/short, $t = 10.26$ for non-island/short, $t = 15.04$ for non-island/long; $p < .001$ for all). The DD-score for the adjunct island effect is 1.10.

3.5.2 Complex NP islands

We find super-additive effects in the complex NP island item set as well ($\beta = 1.40$, $SE = 0.09$, $t = 15.42$, $p < .001$). Moreover, the effects of *distance* ($\beta = –0.95$, $SE = 0.05$, $t = –21.02$, $p < .001$) and *construction* ($\beta = 0.72$, $SE = 0.05$, $t = 15.84$, $p < .001$) were significant, indicating that short-distance dependencies received higher scores than long-distance dependencies, and non-islands higher scores than islands. The post-hoc analyses again indicate that the island/long condition received significantly lower ratings than the other three conditions ($t = 25.76$ for island/short, $t = 26.06$ for non-island/short, $t = 22.10$ for non-island/long; $p < .001$ for all). The DD-score for complex NP islands is 1.40.

3.5.3 Interrogative islands

In the interrogative (*whether*) island item set, we find a significant interaction effect ($\beta = 1.80$, $SE = 0.08$, $t = 21.49$, $p < .001$), as well as significant main effects of *distance* ($\beta = –0.20$, $SE = 0.04$, $t = –4.67$, $p < .001$) and *construction* ($\beta = 0.70$, $SE = 0.05$, $t = 15.24$, $p < .001$). Short-distance dependencies again received higher scores than long-distance dependencies, and non-island constructions higher scores than island constructions. Post-hoc analyses show that the island/long condition received significantly lower scores than the other three conditions ($t = 18.50$ for island/short, $t = 14.40$ for non-island/short, $t = 25.75$ for non-island/long; $p < .001$ for all). The DD-score for interrogative islands is 1.80.

3.5.4 Subject islands

The interaction effect between the factors *gap site* and *complexity* in the subject island item set was significant ($\beta = 0.41$, $SE = 0.09$, $t = 4.57$, $p < .001$), as well as the two main effects (*gap site*: $\beta = –0.14$, $SE = 0.05$, $t = –2.98$, $p = .003$; *complexity*: $\beta = 1.25$, $SE = 0.05$, $t = 27.451$, $p < .001$). Recall that the design of these stimuli was different from the rest. The subject conditions received worse rating than the object conditions, and movement out of complex phrases worse ratings than movement of a single element. Post-hoc analyses reveal that the condition with extraction from a complex subject was rated significantly lower than the other conditions ($t = 5.34$ for complex/object, $t = 21.57$ for simple/object, $t = 22.69$ for simple/subject; $p < .001$ for all). Although the interaction effect is significant, its DD-score is only 0.41.

3.6 Discussion

Our results show that super-additive effects arise in all islands constructions tested in our experiment, providing first experimental evidence for three types of island sensitivity in Romanian *wh*-dependencies (see also Stoica & Schoenmakers 2024 for interrogative islands). We find the strongest island effect in items with an interrogative island (DD 1.80), with gradually weaker effects in items with complex NP (DD 1.40) or adjunct (DD 1.10) islands, and the weakest effect in items with a subject island (DD 0.41). However, recall that the latter item set followed a different design, which has been claimed to underestimate the island effect size (Kush et al. 2018).

Interestingly, closer inspection of **Figure 1** shows that the interrogative island and adjunct island effects are not only driven by the low ratings for *wh*-extraction from island constructions; they are also associated with unexpectedly high ratings for *wh*-extraction from non-island constructions (compared to the other conditions). Post-hoc analyses confirm that the non-island/long conditions received significantly higher ratings than the two short-distance conditions (adjunct islands: $t_{\text{island/short}} = 4.32$, $t_{\text{non-island/short}} = 4.77$; interrogative islands: $t_{\text{island/short}} = 8.10$, $t_{\text{non-island/short}} = 11.90$; all $p < .001$). Although this finding does not dispute the existence of the

island effects in Romanian *per se*, it does call into question the reported DD-scores (cf. Stoica & Schoenmakers 2024). We address this issue in Section 3.6.2.

3.6.1 Addressing our linking hypotheses

That we find all four types of island effects in Romanian is, to a large extent, in line with both the experimental and theoretical literature. Section 3.1 discussed how syntactic theories of island phenomena predict the unacceptability of *wh*-dependencies across complex NP (Subjacency), adjunct, and subject (CED, Barriers) boundaries. Our data confirm this pattern, although the subject island effects are only weak. This might be an artifact of the alternative design of the stimuli necessitated in Romance languages (Sprouse et al. 2016; Abeillé et al. 2020). As argued in Kush et al. (2018), the set-up we used may underestimate the island effect size, because the complex conditions might receive low acceptability ratings to the extent that they cause floor effects, reducing the interaction effect. With average ratings of 2.15 (complex subjects) and 2.94 (complex objects), this is a possible confound in our experiment. The item set, however, also contained a possible definiteness effect, in that the subjects were all definite DPs whereas the objects were all indefinite DPs. Definites are known to be less transparent than indefinites, which opens up the conflicting possibility that the complex subject condition received lower acceptability ratings. This would result in an *overestimation* of the subject island effect. We thus treat the subject island effect with caution, noting that the indicative mood in the embedded clauses may have further reduced acceptability ratings (see Section 2.4). This is a potential route for future research, as Di Sciullo et al. (2003) claim that subjects in Romanian VSO structures can be extracted from, while their examples all include embedded clauses in the subjunctive. We discuss the limitations of our experiment and directions for future research in Section 4.2.

Under the assumption that CP is a cyclic node in Romance languages, the Subjacency Condition does not predict that interrogative clauses are islands for extraction (Sportiche 1981 for French; Rizzi 1982 for Italian; Torrego 1984 for Spanish). Experimental studies on Spanish and Italian (López Sancio 2015; Sprouse et al. 2016) do not corroborate this claim, at least with respect to *wh*-dependencies. Our findings are in line with these experimental studies, in that *wh*-dependencies with the gap located inside an embedded question are unacceptable in Romanian, with a super-additive effect of *distance* and *construction*. This outcome could be an indication that not CP but IP is a cyclic node in Romance languages after all (*pace* theoretical claims).

Recall that discourse-clash accounts predict unacceptability in all cases under investigation, as *wh*-movement foregrounds the filler, while the gap is located inside a backgrounded domain (Goldberg 2006; 2013; Abeillé et al. 2020). However, we did not manipulate the dependency type between focused *wh*-movement of relativization. Further, processing accounts would also

predict super-additivity in each item set, but we did not include a working memory task in our experiments to test the correlation of the island effect sizes and individual participants' cognitive resources such accounts predict (cf. Sprouse et al. 2012a; Hofmeister et al. 2014; Michel 2014; Pham et al. 2020). Our findings may thus be explained by accounts based on information structural conflicts or resource overload, but additional research is required to separate their exact predictions from those by syntactic accounts.

3.6.2 A comparison of experimental results and the definition of 'island effects'

Sprouse et al. (2016) and López Sancio (2015) find the same island effects in their investigation of *wh*-dependencies in Italian and Spanish as we did for Romanian; the relative effect sizes for each language are presented in **Table 4**. The three languages have in common that interrogative island effects are the strongest. Further, subject island effects are the weakest in Spanish and Romanian by a considerable margin. Before making further comparisons between these languages, however, a closer look at the data pattern in **Figure 1** is in order.

	Language	Adjuncts	Complex NPs	Interrogatives	Subjects
Sprouse et al. (2016)	Italian	1.31	0.89	1.69	1.37
López Sancio (2015)	Spanish	1.71	1.41	1.65	0.80
This study	Romanian	1.10	1.40	1.80	0.41

Table 4: DD-scores (based on *z*-scores) per island type reported in experimental work on *wh*-extraction in Romance languages with a comparable experimental design.

Examining the data pattern in **Figure 1** closely, we recognize a certain trend in some of the data-sets: the theoretically unmarked non-island/long condition receives remarkably high ratings, ratings that are significantly higher, in fact, than those for the other theoretically unmarked conditions. This trend can be observed particularly clearly in our adjunct and interrogative island data, but also in experimental papers from others (based on visual inspection, disregarding subject islands because of their alternative design), such as Stoica and Schoenmakers (2024) for Romanian *whether- and when*-islands with a *wh*-dependency, with bare *wh* and *what* N fillers, and Kush et al. (2018) for Norwegian conditional adjunct and *whether*-islands with a *wh*-dependency, with complex fillers.

The pattern we observe is not problematic under the common definition of an island effect (see Sprouse & Villata 2021), because it does not rely on low ratings for the island/long condition alone, but rather on all four conditions. That is, it is possible that increased distance between the filler and the gap improves acceptability for both the island and the non-island condition, which amounts to a main effect of the factor *distance*. The non-island/long condition then receives

higher ratings than its short-distance counterparts, and so does the island/long condition, although this condition also suffers from the penalty commonly theorized to be due to subject islandhood (which is arguably strong enough to push the mean rating below the mean ratings of the short-distance conditions). This data pattern yields super-additivity, represented by a large DD-score.[14] As noted, this does not pose a problem from a strictly statistical point of view. However, this explanation of the findings revolves around 'reverse distance' effects, according to which long-distance dependencies receive higher ratings than their short-distance counterparts. From a psycholinguistic perspective, this is highly unexpected: it has frequently been observed that the parser attempts to resolve *wh*-dependencies as soon as possible (e.g. Phillips et al. 2005 and references therein). And while it is possible, of course, that the effect of dependency length is not the same across languages (cf. Vasishth et al. 2010), the observed reverse distance effects appear to be unstable even within languages: unlike our adjunct and interrogative island item sets, our complex NP item set moves in the expected direction (see **Figure 1**). Moreover, Stoica & Schoenmakers (2024) report reverse distance effects with different magnitudes for *wh*-fillers with different complexities (i.e. bare *wh* vs. *what* N vs. *which* N) in Romanian, and while the reverse pattern found in Norwegian conditional adjuncts and *whether*-islands with a *wh*-dependency (Kush et al. 2018) was replicated in Kobzeva et al. (2022) with *wh*-dependencies (for conditional adjuncts only; Kobzeva et al. tested different types of embedded question islands), it was not replicated in Kush et al. (2019) or Kobzeva et al. (2022) with topicalization and RC-dependencies. It thus appears that the island type, the filler type, and the dependency type may influence (reverse) distance effects in experiments investigating island phenomena.

That such factors may influence the acceptability of long-distance dependencies is in line with processing accounts (see Section 2.3), which claim that a combination of factors contributing processing costs may constrain ease of parsing, and that these factors may influence acceptability in different ways. Kluender & Kutas (1993b) show, for example, that different types of embedded clause (*that, if, wh*) are not equally transparent, which is reflective of the processing cost associated with maintaining a *wh*-dependency across them. However, we stress that the observed distance effects are *reversed* from what we would expect from a psycholinguistic perspective, and therefore the unexpected pattern is, as of yet, without formal explanation. We consequently repeat Michel's (2014: 118) recommendation to visualize data patterns and perform post-hoc pairwise comparisons to "clarify what a reasonable interpretation

[14] Note that a large DD-score may also be arrived at in other ways, that is, they are not necessarily due to lower acceptability of the island/long condition (Michel 2014). Stoica & Schoenmakers (2024: 184) argue that, on the common definition of super-additivity, "we could speak of an island effect which is due to differences in the acceptability of sentences *without an island structure*, while at the same time nothing particularly interesting is happening in sentences *with an island structure.*" According to Michel (2014: 118), "it would not be appropriate to conclude that such a pattern showed an 'island effect'," and he stresses that visualizations of the data as well as post-hoc paired comparisons are crucial for the interpretation of island effects.

of a DD score should be". With regard to the DD-scores in **Table 4**, we note that the data plots in Sprouse et al. (2016) and López Sancio (2015) follow comparable patterns, and the Romanian data from complex NP and subject islands do, too. As noted, however, the adjunct and interrogative island effects follow a pattern distinct from the other languages, but similar to the interrogative island pattern reported in Stoica & Schoenmakers (2024). It could be that a confounding factor the experimental design does not control for contaminates the acceptability ratings in some conditions but not others, and consequently impacts the DD-score (cf. Keshev & Meltzer-Asscher 2019). Possible reasons for the unexpected pattern can be explored in future research.

4 General discussion

4.1 Results

Despite the unexpected patterns discussed in Section 3.6.2, based on the full data pattern in **Figure 1** we conclude that Romanian displays island sensitivity: the theoretically marked island/long conditions received lower acceptability ratings in each item set, as confirmed by post-hoc pairwise comparisons. More specifically, Romanian exhibits at least interrogative, adjunct, and complex NP island effects. While the adjunct and complex NP island effects were expected based on the literature, the interrogative island effect contradicts the original claim that the language does not obey the *Wh*-island Constraint (Comorovski 1986, Rudin 1988). Our stimuli were presented free of context, however, and the fronted elements were bare *wh*-phrases. It has been claimed that extraction from interrogative clauses is possible on the condition that the extracted element is D-linked (Comorovski 1989; Alboiu 2002) and a recent experimental study (Stoica & Schoenmakers 2024) shows that D-linking (bare *wh* vs. *which* N) indeed ameliorates extraction from *whether*- and *when*-islands in Romanian (although the island effect in the D-linked conditions is still significant). The degree of discourse embedding may thus impact the island effect size, which may also play a role in other island types (as claimed in Sevcenco 2006 for Romanian complex NP islands).

We also find a weak but significant subject island effect, but we treat this effect with caution. The alternative experimental design we chose may have underestimated the size of the island effect, because of floor effects in the complex conditions (Kush et al. 2018), while a definiteness effect may have overestimated it.

Our results are in line with earlier experimental work on Italian (Sprouse et al. 2016) and Spanish (López Sancio 2015), although there is variation between the languages with regard to the island effect sizes (which are in part due to the unexpected pattern in the Romanian data, see Section 3.6.2). In what follows, we discuss potential avenues for future research.

4.2. Limitations and future directions

Our experiments provide initial evidence for island sensitivity in Romanian, based on experimental data from four construction types. This outcome was predicted by most theories of island phenomena. Future research can be geared towards more precise claims from linguistic theories. We suggest various ways forward.

First, all our stimuli contained island structures in the indicative. Indicatives clauses, however, have been argued to be less permeable to extraction than subjunctive clauses (see Section 2.4; Baunaz & Puskás 2014). This claim can be tested experimentally in future work.

Second, recall that Abeillé et al.'s (2020) *Focus-Background Conflict* criterion in (6) predicts that a focused element should not be part of a backgrounded constituent. As such, the condition rules out *wh*-dependencies, but not other types of dependencies which involve non-focused movement, such as relativization or topicalization (*pace* Goldberg 2006; 2013). Our experiment tested *wh*-dependencies only. All else being equal, the FBC criterion predicts that other dependency types are not resisted in the same way as *wh*-dependencies. Future work could therefore investigate other dependency types, as has been done for other languages as well (e.g. López Sancio 2015; Sprouse et al. 2016; Kush et al. 2019; Abeillé et al. 2020; Kobzeva et al. 2022; Cuneo & Goldberg 2023).

Third, our experiment did not include a working memory task, so we cannot investigate the potential correlation between the island effect size and working memory capacity, as would be predicted by processing accounts of island phenomena. Future experiments may include a working memory task, although the design of the experiment as well as the choice of working memory task are matter of debate (Hofmeister 2012a,b; 2014; Sprouse et al. 2012a,b). Earlier experimental investigations have found that individual cognitive resources explain only little variation in the observed island effects (Sprouse et al. 2012a; Michel 2014; Pham et al. 2020).

Processing accounts also argue that processing is easier with stronger discourse embedding or accessibility (see Hofmeister & Sag 2010). Complex DP fillers (e.g. *which book* instead of *what*) are therefore expected to yield higher acceptability ratings than bare *wh* fillers, because they are D-linked (Pesetsky 1987) (or because there are less demanding interference effects, Villata et al. 2016; Chesi et al. 2023; Stoica & Schoenmakers 2024). Experimental studies on various languages have included the filler's complexity as a factor in their design and indeed found amelioration effects (Goodall 2015; Sprouse et al. 2016 for English; Kush et al. 2018; Stoica & Schoenmakers 2024), but the island effect persisted when the fillers were complex instead of bare *wh*-phrases. As noted, Stoica & Schoenmakers (2024) tested the claim for Romanian interrogative islands, but the question has not been addressed for other island types in the language.

Finally, another approach to testing claims from processing accounts, suggested by a reviewer, is to investigate the impact of verb frame frequencies in the experimental items (Liu

et al. 2022a). The [V + *that*] frame may be more frequent than the [V + *whether*] frame, for example, which may in turn increase the acceptability of the conditions with an embedded declarative, on the assumption that these sentences resemble the stored template of a prototypical structure more closely (after Dąbrowska 2008). While post-hoc computations of this kind would be a viable option for our interrogative island item set, our experiment did not manipulate the verb forms (we only tested *crede că* 'believe that' and *se întreaba dacă* 'wonder whether'). The results would consequently not be theoretically informative. Verb frame frequences may provide potentially insightful information, however, as they did for English in Liu et al. (2022a), and so this hypothesis can be tested for other languages in dedicated investigations in the future.

We conclude this section with a methodological note. Our experiments as well as several earlier experimental studies tested only eight items per island type. In a 2×2 design, this amounts to participants rating only two distinct items per condition. We urge future researchers to test larger numbers of items, so as to avoid that the acceptability ratings reflect lexical properties of the stimuli or other idiosyncratic factors. More generally, statistical models may struggle to provide accurate predictions when only a small number of items are tested (i.e. low statistical power). Although it is difficult to specify an ideal number of items required in experiments like the one presented in this paper, it is worth noting that larger numbers of items increase the reliability of the data-set. As Maxwell & Delaney (2003: 25) illustrate, the chances of passing a multiple-choice quiz by simply guessing the answers are slimmer when the test is longer. Increased numbers of items in experiments investigating island phenomena would thus allow for closer scrutiny of the variation patterns in them.

5 Conclusion

In this paper, we reported on two acceptability judgment experiments in which we investigated whether *wh*-dependencies are unacceptable in Romanian when the gap is located inside four types of structures (islands). In line with earlier experimental studies, we tested four island types (adjuncts, complex NPs, interrogative clauses, subjects) with *wh*-dependencies and find significant island effects in each of the data-sets. The subject island effects are somewhat weaker than the others, which is presumably due to the alternative stimulus creation procedure. We treat this island effect with care. That we find significant interrogative island effects was not predicted by Subjacency accounts which adopt parametrization and assume that CP and not IP is a cyclic node in Romance languages; this approach did however predict the adjunct and complex NP island effects we found. Note, however, that our findings do not exclude the possibility that the source of the observed island effects is syntactic. Our findings are also compatible with discourse-based and processing accounts; more research is required to disentangle the exact claims from these competing theories. Our suggestions for future experimental work include investigations

of the influence of the mood of the embedded clause, dependency type, filler type, verb frame frequencies, and cognitive resources available for individual participants.

Our findings provide an empirical anchor for the discussion on island sensitivity in Romanian, which corroborates earlier experimental findings for other Romance languages (e.g. López Sancio 2015; Sprouse et al. 2016) at least with regard to *wh*-dependencies. We note, however, that many experimental studies into island effects (including ours) tested rather small numbers of items. This may impede the reliability and the generalizability of the results. The precision of measurement can thus be improved, in order to achieve a better understanding of the vast empirical landscape of island effects, in Romanian and beyond.

Abbreviations

1SG = first person singular; 2SG = second person singular; 3SG = third person singular; 3PL = third person plural; ACC = accusative; CL = clitic; DAT = dative; F = feminine; REFL = reflexive; SUBJ = subjunctive

Data availability

The data files, stimuli, and analysis scripts are available in our Open Science Framework repository at https://osf.io/23wrg/.

Ethics statement

The experiments were approved by the Ethics Assessment Committee Humanities of Radboud University (number 2022-5980).

Acknowledgments

We would like to thank all Romanian participants for their participation in the experiment, and three anonymous reviewers for their valuable comments.

Competing interests

The authors have no competing interests to declare.

References

Abeillé, Anne & Hemforth, Barbara & Winckel, Elodie & Gibson, Edward. 2020. Extraction from subjects: Differences in acceptability depend on the discourse function of the construction. *Cognition* 204. Article 104293. DOI: https://doi.org/10.1016/j.cognition.2020.104293

Alboiu, Gabriela. 2002. *The features of movement in Romanian*. University of Bucharest Press.

Bates, Douglas & Mächler, Martin & Bolker, Ben & Walker, Steve. 2015. Fitting linear mixed effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: https://doi.org/10.18637/jss.v067.i01

Baunaz, Lena & Puskás, Genoveva. 2014. On subjunctives and islandhood. In Côté, Marie-Hélène & Mathieu, Eric (eds.), *Variation within and across Romance languages: Selected papers from the 41st Linguistic Symposium on Romance Languages*, 233–254. John Benjamins. DOI: https://doi.org/10.1075/cilt.333.16bau

Boeckx, Cedric. 2012. *Syntactic islands*. Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139022415

Chaves, Rui & Putnam, Michael. 2020. *Unbounded dependency constructions: Theoretical and experimental perspectives*. Oxford University Press. DOI: https://doi.org/10.1093/oso/9780198784999.001.0001

Chesi, Cristiano & Bressan, Veronica & Belletti, Adriana. 2023. The role of D-linking and lexical restriction in locality violations. *Quaderni di Linguistica e Studi Orientali* 9. 29–43. DOI: https://doi.org/10.36253/qulso-2421-7220-15162

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

Chomsky, Noam. 1977. On wh-movement. In Culicover, Peter & Wasow, Thomas & Akmajian, Adrian (eds.), *Formal syntax,* 71–132. Academic Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Foris.

Chomsky, Noam. 1986. *Barriers*. MIT Press.

Chomsky, Noam. 1995. *The minimalist program*. MIT Press. DOI: https://doi.org/10.7551/mitpress/9780262527347.001.0001

Comorovski, Ileana. 1986. Multiple wh movement in Romanian. *Linguistic Inquiry* 17(1). 171–177.

Comorovski, Ileana. 1989. Discourse-linking and the Wh-island Constraint. *Proceedings of East Linguistics Society* 19(1). 78–96.

Coopmans, Cas & Ligtenberg, Kars & Suijkerbuijk, Michelle & Schoenmakers, Gert-Jan. 2024. Comparing syntactic and discourse accounts of adjunct island effects: Experimental evidence from acceptability judgments. Preprint available at https://lingbuzz.net/lingbuzz/008551.

Cuneo, Nicole & Goldberg, Adele. 2023. The discourse functions of grammatical constructions explain an enduring syntactic puzzle. *Cognition* 240. Article 105563. DOI: https://doi.org/10.2139/ssrn.4355084

Dąbrowska, Ewa. 2008. Questions with long-distance dependencies: A usage-based perspective. *Cognitive Linguistics* 19(3). 391–425. DOI: https://doi.org/10.1515/COGL.2008.015

Deane, Paul. 1991. Limits to attention: A cognitive theory of island phenomena. *Cognitive Linguistics* 2(1). 1–63. DOI: https://doi.org/10.1515/cogl.1991.2.1.1

den Dikken, Marcel & Lahne, Antje. 2013. The locality of syntactic dependencies. In den Dikken, Marcel (ed.), *The Cambridge handbook of generative syntax*, 655–698. Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511804571.023

Di Sciullo, Anna Maria & Paul, Ileana & Somesfalean, Stanca. 2003. The clause structure of extraction asymmetries. In Di Sciullo, Anna Maria (ed.), *Asymmetries in grammar*, 285–306. John Benjamins. DOI: https://doi.org/10.1075/la.57.13dis

Dobrovie-Sorin, Carmen. 1990. Clitic doubling, wh-movement, and quantification in Romanian. *Linguistic Inquiry* 21(3). 351–397.

Erteschik-Shir, Nomi. 1973. On the nature of island constraints. Ph.D. dissertation at Massachusetts Institute of Technology.

Erteschik-Shir, Nomi & Lappin, Shalom. 1979. Dominance and the functional explanation of island phenomena. *Theoretical Linguistics* 6(1–3). 41–86. DOI: https://doi.org/10.1515/thli.1979.6.1-3.41

Goldberg, Adele. 2006. *Constructions at work: The nature of generalization in language*. Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199268511.001.0001

Goldberg, Adele. 2013. Backgrounded constituents cannot be “extracted”. In Sprouse, Jon & Hornstein, Norbert (eds.), *Experimental syntax and island effects*, 221–238. Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139035309.012

Goldberg, Adele & Cuneo, Nicole & Fergus, Abigail. 2024. Addressing a challenge to the Backgroundedness account of islands. DOI: https://doi.org/10.31234/osf.io/hmc9n

Goodall, Grant. 2015. The D-linking effect on extraction from islands and non-islands. *Frontiers in Psychology* 5. Article 1493. DOI: https://doi.org/10.3389/fpsyg.2014.01493

Hofmeister, Philip & Sag, Ivan. 2010. Cognitive constraints and island effects. *Language* 86(2). 366–415. DOI: https://doi.org/10.1353/lan.0.0223

Hofmeister, Philip & Staum Casasanto, Laura & Sag, Ivan. 2012a. How do individual cognitive differences relate to acceptability judgments? A reply to Sprouse, Wagers, and Phillips. *Language* 88(2). 390–400. DOI: https://doi.org/10.1353/lan.2012.0025

Hofmeister, Philip & Staum Casasanto, Laura & Sag, Ivan. 2012b. Misapplying working-memory tests: A reductio ad absurdum. *Language* 88(2). 408–409. DOI: https://doi.org/10.1353/lan.2012.0033

Hofmeister, Philip & Staum Casasanto, Laura & Sag, Ivan. 2014. Processing effects in linguistic judgment data: (super-)additivity and reading span scores. *Language and Cognition* 6(1). 111–145. DOI: https://doi.org/10.1017/langcog.2013.7

Huang, James. 1982. Move *wh* in a language without *wh* movement. *The Linguistic Review* 1. 369–416. DOI: https://doi.org/10.1515/tlir.1982.1.4.369

Keshev, Mayaan & Meltzer-Asscher, Aya. 2019. A processing-based account of subliminal wh-island effects. *Natural Language & Linguistic Theory* 37(2). 621–657. DOI: https://doi.org/10.1007/s11049-018-9416-1

Kluender, Robert. 1991. *Cognitive constraint on variables in syntax*. Ph.D. dissertation at University of California, San Diego.

Kluender, Robert. 1992. Deriving island constraints from principles of predication. In Goodluck, Helen & Rochemont, Michael (eds.), *Island constraints: Theory, acquisition, and processing*, 223–258. Kluwer. DOI: https://doi.org/10.1007/978-94-017-1980-3_8

Kluender, Robert. 1998. On the distinction between strong and weak islands: A processing perspective. In Culicover, Peter & McNally, Louise (eds.), *Syntax and Semantics 29: The limits of syntax*, 241–279. Academic Press. DOI: https://doi.org/10.1163/9789004373167_010

Kluender, Robert. 2004. Are subject islands subject to a processing account. In Chand, Vineeta & Kelleher, Ann & Rodríguez, Angelo & Schmeiser, Benjamin (eds.), *Proceedings of the 23rd West Coast Conference on Formal Linguistics*, 475–499. Cascadilla Press.

Kluender, Robert & Kutas, Marta. 1993a. Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience* 5(2). 196–214. DOI: https://doi.org/10.1162/jocn.1993.5.2.196

Kluender, Robert & Kutas, Marta. 1993b. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8(4). 573–633. DOI: https://doi.org/10.1080/01690969308407588

Kobzeva, Anastasia & Sant, Charlotte & Robbins, Parker & Vos, Myrte & Lohndal, Terje & Kush, Dave. 2022. Comparing island effects for different dependency types in Norwegian. *Languages* 7(3). Article 97. DOI: https://doi.org/10.3390/languages7030197

Kush, Dave & Lohndal, Terje & Sprouse, Jon. 2018. Investigating variation in island effects: A case study of Norwegian wh-extraction. *Natural Language & Linguistic Theory* 36(3). 743–779. DOI: https://doi.org/10.1007/s11049-017-9390-z

Kush, Dave & Lohndal, Terje & Sprouse, Jon. 2019. On the island sensitivity of topicalization in Norwegian: An experimental investigation. *Language* 95(3). 393–420. DOI: https://doi.org/10.1353/lan.2019.0051

Lenth, R. 2024. emmeans: Estimated marginal means, aka least-squares means. R package version 1.10.2. DOI: https://doi.org/10.32614/CRAN.package.emmeans

Liu, Yongting & Ryskin, Rachel & Futrell, Richard & Gibson, Edward. 2022a. A verb-frame frequency account of constraints on long-distance dependencies in English. *Cognition* 222. Article 104902. DOI: https://doi.org/10.1016/j.cognition.2021.104902

Liu, Yongting & Winckel, Elodie & Abeillé, Anne & Hemforth, Barbara & Gibson. Edward. 2022b. Structural, functional, and processing perspectives on linguistic island effects. *Annual Review of Linguistics* 8. 495–525. DOI: https://doi.org/10.1146/annurev-linguistics-011619-030319

López Sancio, Sergio. 2015. *Testing syntactic islands in Spanish*. MA thesis at University of the Basque Country.

Maxwell, Scott & Delaney, Harold 2003. *Designing experiments and analyzing data: A model comparison perspective* (2nd edition). Psychology Press. DOI: https://doi.org/10.4324/9781410609243

Michel, Daniel. 2014. *Individual cognitive measures and working memory accounts of syntactic island phenomena*. Ph.D. dissertation at University of California, San Diego.

Momma, Shota & Dillon, Brian. 2023. Discourse factors do not explain islands. DOI: https://doi.org/10.2139/ssrn.4635713

Neal, Anissa & Dillon, Brian. 2021. Definitely islands? Experimental investigation of definite islands. In Beltrama, Andrea & Schwarz, Florian & Papafragou, Anna (eds.), *Proceedings of Experiments in Linguistics Meaning 1*, 237–248. Linguistic Society of America. DOI: https://doi.org/10.3765/elm.1.4885

Newmeyer, Frederick. 2016. Nonsyntactic explanations of island constraints. *Annual Review of Linguistics* 2. 187–210. DOI: https://doi.org/10.1146/annurev-linguistics-011415-040707

Pană Dindelegan, Gabriela. 2013. *The grammar of Romanian*. Oxford University Press.

Pañeda, Claudia & Kush, Dave. 2021. Spanish embedded question island effects revisited: An experimental study. *Linguistics* 60(2). 463–504. DOI: https://doi.org/10.1515/ling-2020-0110

Pañeda, Claudia & Lago, Sol & Vares, Elena & Veríssimo, João & Felser, Claudia. 2020. Island effects in Spanish comprehension. *Glossa: A Journal of General Linguistics* 5. Article 21. DOI: https://doi.org/10.5334/gjgl.1058

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

Pham, Catherine & Covey, Lauren & Gabriele, Alison & Aldosari, Saad & Fiorentino, Robert. 2020. Investigating the relationship between individual differences and island sensitivity. *Glossa: A Journal of General Linguistics* 5. Article 94. DOI: https://doi.org/10.5334/gjgl.1199

Phillips, Colin. 2013. Some arguments and nonarguments for reductionist accounts of syntactic phenomena. *Language and Cognitive Processes* 28(1/2). 156–187. DOI: https://doi.org/10.1080/01690965.2010.530960

Phillips, Colin & Kazanina, Nina & Abada, Shani. 2005. ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research* 22(3). 407–428. DOI: https://doi.org/10.1016/j.cogbrainres.2004.09.012

Pritchett, Bradley. 1991. Subjacency in a principle-based parser. In Berwick, Robert & Abney, Steven & Tenney, Carol (eds.), *Principle-based parsing: Computation and psycholinguistics*, 301–345. Kluwer. DOI: https://doi.org/10.1007/978-94-011-3474-3_12

R Core Team. 2022. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rizzi, Luigi. 1982. *Issues in Italian syntax*. Foris. DOI: https://doi.org/10.1515/9783110883718

Rizzi, Luigi. 1990. *Relativized Minimality*. MIT Press. DOI: https://doi.org/10.1016/0024-3841(91)90047-9

Rodríguez, Alejandro & Goodall, Grant. 2023. Clitic left dislocation in Spanish: Island sensitivity without gaps. Presentation at the *36th Annual Conference on Human Sentence Processing*. Pittsburgh.

Ross, John. 1967. *Constraints on variables in syntax*. Ph.D. dissertation at Massachusetts Institute of Technology.

Rudin, Catherine. 1988. On multiple questions and multiple wh-fronting. *Natural Language & Linguistics Theory* 6. 445–501. DOI: https://doi.org/10.1007/BF00134489

Sevcenco, Anca. 2006. *Terms of binding*. Ph.D. dissertation at Utrecht University.

Shen, Zheng & Lim, Meghan. 2024. Definite DP islands effects across dependency types. Preprint available at https://ling.auf.net/lingbuzz/007747

Sportiche, Dominique. 1981. Bounding nodes in French. *The Linguistic Review* 1(2). 219–246. DOI: https://doi.org/10.1515/tlir.1981.1.2.219

Sprouse, Jon. 2007. *A program for experimental syntax: Finding the relationship between acceptability and grammatical knowlege* Ph.D. dissertation at University of Maryland.

Sprouse, Jon & Caponigro, Ivano & Greco, Ciro & Cecchetto, Carlo. 2016. Experimental syntax and the variation of island effects in English and Italian. *Natural Language & Linguistic Theory* 34(1). 307–344. DOI: https://doi.org/10.1007/s11049-015-9286-8

Sprouse, Jon & Hornstein, Norbert. 2013. *Experimental syntax and island effects*. Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139035309

Sprouse, Jon & Villata, Sandra. 2021. Island effects. In Goodall, Grant (ed.), *The Cambridge handbook of experimental syntax*, 227–258. Cambridge University Press. DOI: https://doi.org/10.1017/9781108569620.010

Sprouse, Jon & Wagers, Matt & Phillips, Colin. 2012a. A test of the relation between working memory capacity and syntactic island effects. *Language* 88(1). 82–123. DOI: https://doi.org/10.1353/lan.2012.0004

Sprouse, Jon & Wagers, Matt & Phillips, Colin. 2012b. Working memory capacity and island effects: A reminder of the issues and the facts. *Language* 88(2). 401–407. DOI: https://doi.org/10.1353/lan.2012.0029

Starke, Michal. 2001. *Move dissolves into merge: A theory of locality*. Ph.D. dissertation at University of Geneva.

Stoica, Irina & Schoenmakers, Gert-Jan. 2024. Interrogative islands in Romanian. In Sevcenco, Anca & Stoica, Irina & Stoicescu, Ioana & Tigău, Alina & Tomescu, Veronica & Tănase-Dogaru, Mihaela (eds.), *Exploring linguistic landscapes: A Festschrift for Larisa Avram and Andrei Avram*, 163–189. Bucharest University Press.

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

van Casteren, Maarten & Davis, Matthew. 2006. Mix, a program for pseudorandomization. *Behavior Research Methods* 38(4). 584–589. DOI: https://doi.org/10.3758/BF03193889

Vasishth, Shravan & Suckow, Katja & Lewis, Richard & Kern, Sabine. 2010. Short-term forgetting in sentence comprehension: Crosslinguistic evidence from verb-final structures. *Language and Cognitive Processes* 25(4). 533–567. DOI: https://doi.org/10.1080/01690960903310587

Villata, Sandra & Rizzi, Luigi & Franck, Julie. 2016. Intervention effects and relativized minimality: New experimental evidence from graded judgments. *Lingua* 179. 76–96. DOI: https://doi.org/10.1016/j.lingua.2016.03.004