Nat Lang Linguist Theory (2016) 34:307–344
DOI 10.1007/s11049-015-9286-8

Experimental syntax and the variation of island effects in English and Italian

Jon Sprouse[1] · Ivano Caponigro[2] · Ciro Greco[3] · Carlo Cecchetto[4]

Received: 28 October 2012 / Accepted: 2 March 2015 / Published online: 10 April 2015

Abstract The goal of this article is to explore the utility of experimental syntax techniques in the investigation of syntactic variation. To that end, we applied the factorial definition of island effects made available by experimental syntax (e.g., Sprouse et al. 2012) to four island types (wh/*whether*, complex NP, subject, and adjunct), two dependency types (wh-interrogative clause dependencies and relative clause dependencies) and two languages (English and Italian). The results of 8 primary experiments suggest that there is indeed variation across dependency types, suggesting that wh-interrogative clause dependencies and relative clause dependencies cannot be identical at every level of analysis; however, the pattern of variation observed in these experiments is not exactly the pattern of variation previously reported in the literature (e.g., Rizzi 1982). We review six major syntactic approaches to the analysis of island effects (Subjacency, CED, Barriers, Relativized Minimality, Structure-building, and Phases) and discuss the implications of these results for these analyses. We also present 4 supplemental experiments testing complex wh-phrases (also called D-linked or lexically restricted wh-phrases) for all four island types using the factorial design in order to tease apart the contribution of dependency type from featural

✉ J. Sprouse
jon.sprouse@uconn.edu

I. Caponigro
ivano@ucsd.edu

C. Greco
ciro.greco@ugent.be

C. Cecchetto
carlo.cecchetto@unimib.it

1 Department of Linguistics, University of Connecticut, Storrs, CT, USA

2 Department of Linguistics, University of California, San Diego, CA, USA

3 Department of English Studies, University of Ghent, Ghent, Belgium

4 Department of Psychology, University of Milan-Bicocca, Milan, Italy

specification. The results of the supplemental experiments confirm that dependency type is the major source of variation, not featural specification, while providing a concrete quantification of what exactly the effect of complex wh-phrases on island effects is.

Keywords Experimental syntax · Island effects · Cross-linguistic variation · wh-Movement · Relative clauses · D-linking

1 Introduction

It has long been the case that the primary empirical foundation of syntactic theories is informally collected acceptability judgments (Chomsky 1965; Schütze 1996). However, in recent years there has been a growing interest in the use of more formal methods for the collection of acceptability judgments, both due to concerns regarding the reliability of informally collected judgments (e.g., Ferreira 2005; Wasow and Arnold 2005; Gibson and Fedorenko 2010), and due to indications that more formal methods may reveal previously unobserved patterns in the data (e.g., Keller 2000; Featherston 2005; Sprouse et al. 2011). While concerns about the reliability of informally collected judgments has turned out to be less pressing than previously thought (Sprouse and Almeida 2012; Sprouse, Schütze, and Almeida 2013), it is still an open question to what extent formal methods may reveal patterns that were previously unobserved with informal data collection methods, especially with respect to complex syntactic phenomena such as syntactic island effects. To that end, we have three goals in this article. First, we want to apply a quantitative definition of island effects that is only licensed by formal experimental methods. We will use a factorial design to isolate island effects over and above other factors (such as processing complexity) that may influence acceptability judgments (Sprouse 2007; Sprouse et al. 2011, 2012). Second, we want to apply the factorial design cross-linguistically to see if it reveals patterns of variation that are distinct from the patterns revealed by informal methods. For this article, we decided to test syntactic island effects in English and Italian because these two languages have been cited as evidence for cross-linguistic variation in syntactic island effects (e.g., Rizzi 1982). Variation is often used as evidence for the construction of specific syntactic theories, therefore we believe it is particularly important to explore the patterns of results that may be revealed using formal experiments and the factorial design. Finally, we want to apply the factorial design across dependency types within each language. For this article, we decided to test both *wh-dependencies*, i.e., dependencies between a wh-word/phrase and a gap within the interrogative clause that the wh-word/phrase introduces, and *relative-clause dependencies* (*rc-dependencies*, in short), i.e., dependencies between the head introducing a headed relative clause and a gap within the relative clause. Both dependencies have been reported to display syntactic island effects, and rc-dependencies have been reported to display cross-linguistic variation (Rizzi 1982). Much like cross-linguistic variation, similarity and variation across constructions has often been used as evidence for the construction of syntactic theories, therefore we believe it is important to explore this dimension

of variation as well. To that end, this article reports the results of a series of eight primary experiments designed to investigate four syntactic island effect types (wh-islands, complex NP islands, subject islands, and adjunct islands), across two languages (English and Italian), and two dependency types (wh-dependencies and rc-dependencies). We also present four supplemental experiments that test the effect of complex wh-phrases (e.g., *which car*) on the four island types in English in an attempt to tease apart the contribution of dependency type (i.e., wh-dependency vs. rc-dependency) and the contribution of featural specification of the head of the dependency (e.g., bare wh-words in wh-dependencies vs. the nominal head and wh-word in rc-dependencies).

The rest of this article is organized as follows. Section 2 reviews the previous empirical observations of island effects in English and Italian (we postpone the discussion of theoretical analyses of these observations to Sect. 6 so that we also discuss our results relative to previous analyses). Section 3 introduces the factorial design of island effects at the heart of this investigation. Section 4 explains the logic and design of the eight primary acceptability judgment experiments. Section 5 presents the results of the eight primary experiments. Section 6 presents four supplemental experiments designed to test the effect of complex wh-phrases and tease apart the contribution of dependency type and the contribution of featural specificity, as well a full replication of our English rc-dependency results (to further establish the reliability of our new empirical observations). Section 7 discusses how our experimental results impact the six major types of syntactic analyses that have been proposed to capture islands effects. Section 8 concludes.

2 Island effects and cross-linguistic variation

2.1 An introduction to island effects

One of the defining characteristics of human language is the existence of *long-distance dependencies* between two (or more) elements in a sentence. For example, the wh-interrogative clauses in (1) illustrate a long-distance dependency between the wh-word or wh-phrase at the beginning of the sentence, which is often called the *antecedent* or the *filler*, and the argument position of an embedded verb, which is often called the *gap position*. Although long-distance dependencies are unconstrained with respect to length as measured in number of words or number of clauses, as in (1), there do appear to be constraints on the types of structures that can contain the gap position, as in (2). In the examples below and throughout the paper, the antecedent will be in italics and the gap position will be indicated with underscores.

(1) a. *What* does Susan think that John bought __?
b. *What* does Sarah believe that Susan thinks that John bought __?
c. *What* does Bill claim that Sarah believes that Susan thinks that John bought __?

(2) a. WHETHER ISLAND
**What* do you wonder [whether John bought __]?

b. COMPLEX NP ISLAND[1]
**What* did you make [the claim that John bought __]?
c. SUBJECT ISLAND
**What* do you think [the speech about __] interrupted the TV show?
d. ADJUNCT ISLAND
**What* do you worry [if John buys __]?
e. RELATIVE CLAUSE ISLAND
**What* did you meet [the scientist who invented __]?
f. SENTENTIAL SUBJECT ISLAND
**What* did [that John wrote __] offend the editor?
g. COORDINATE STRUCTURE ISLAND
**What* did John buy [a shirt and __]?
h. LEFT-BRANCH ISLAND
**Which* did John borrow [__ book]?

Following Ross (1967), the unacceptability that arises when the gap position occurs inside one of the prohibited structures in (2) is often referred to as an *island effect*, which draws on the metaphor that the prohibited structures are *islands* that prevent the wh-words or wh-phrases from *moving* to the front of the sentence. Though island effects are typically exemplified by *wh-dependencies*, as in (2), the same effects crucially arise with several different types of long-distance dependencies in human languages: for instance *rc-dependencies* (3), dependencies between the topicalized pre-posed constituent and the lower gap (4), and dependencies between the preposed adjective and the lower gap in adjective-*though* constructions (5). All the (b) examples in (3)–(5) exemplify an extraction out of a *whether*-island.

(3) RELATIVE CLAUSE FORMATION (RC-DEPENDENCIES)
a. I like the *car* that you think [that John bought __].[2]
b. *I like the *car* that you wonder [whether John bought __].

(4) TOPICALIZATION
a. I don't know who bought most of these cars, but *that car*, I think [that John bought __].
b. *I know who bought most of these cars, but *that car*, I wonder [whether John bought__].

(5) ADJECTIVE-THOUGH CONSTRUCTIONS
a. *Smart* though I think [that John is __], I don't trust him to do simple math.

[1] In the syntactic literature it is common to use the label "complex NP island effects" to refer both to extraction from the clausal complement of a noun, as in (2b), and from a relative clause as in (2e). In this paper we use a stricter terminology and use "complex NP island" to refer only to the clausal complement of a noun. See Cecchetto and Donati (2015) for a nonstandard view of this construction.

[2] We remain neutral regarding the analysis of headed relative clauses: the filler of the gap could either be the head noun (as in raising analyses: e.g., Vergnaud 1974; Kayne 1994; Bianchi 1999; Bhatt 2002; Donati and Cecchetto 2011), or a (potentially null) relative pronoun (e.g., Chomsky 1981; Browning 1987).

b. **Smart* though I wonder [whether John is __], I trust him to do simple math.

As we will see in the next section, the variation in island effects between English and Italian involves both interrogative clause formation and related wh-dependencies (2) and headed relative clause formation and related rc-dependencies (3).

2.2 The cross-linguistic variation of island effects in English and Italian

The seminal study on island effects in Italian is by Rizzi (1982), who first observed that Italian does not exhibit the same set of island effects as English. Specifically, Rizzi observed that whereas English exhibits wh, complex NP, and subject islands as in (2) above, Italian only appears to exhibit complex NP islands (all of the judgments about Italian in the examples below are from Rizzi 1982):

(6) WH ISLAND (Rizzi 1982:50, ex. 6)[3]

Tuo *fratello*, *a cui*$_1$ mi domando *che storie*$_2$ abbiano
your brother, to whom$_1$ myself wonder.1SG what stories$_2$ have.SUBJ.3PL
raccontato __$_2$ __$_1$, era molto preoccupato.
told was very worried
'Your brother, who I wonder what stories they told, was really worried.'

(7) COMPLEX NP ISLAND (Rizzi 1982:51, ex. 9a)

**Questo incarico*, che non sapevo la novità che avrebbero affidato __
this task that not knew.1SG the news that have.IRR.3PL assigned
a te, . . .
to you
'This task, which I didn't know the news that they may have assigned to you . . .'

(8) SUBJECT ISLAND (Rizzi 1982:61, ex. 30a)

Questo autore, *di cui* so che il primo libro __ è stato
this author, by whom know.1SG that the first book has been
pubblicato recentemente, . . .
published recently
'This author, who I know that the first book was published recently, . . .'

Rizzi (1982) did not directly investigate adjunct islands. However, for completeness we can anticipate the results of the current experiments slightly and note that Italian also appears to exhibit adjunct islands (see also Stepanov 2007 for a broad review of the languages that demonstrate adjunct islands):

[3]Here and throughout the paper, italics are ours. Also, both the head and the relative pronoun are italicized when they both occur as a way of being non-committal about which of them is the actual antecedent of the gap. Glosses and translations have been modified or added, if missing in the original examples. English translations of unacceptable sentences are not marked with any diacritic.

(9) ADJUNCT ISLAND

*In campo c'è un *giocatore con il quale* ti dovrebbero dare
in field there's a player with the whom to_you should.3PL give
istruzioni chiare [se gli altri giocatori saranno scorretti __].
instructions clear if the other players be.FUT.3PL unfair
'There's a player playing in the filed towards whom they should give you clear instructions if the other players are unfair.'

One fact worth noting is that Rizzi (1982) focused on relative clause formation as opposed to wh-interrogative clause formation in his study of island effects in Italian. This contrasts with most studies of island effects in English, which have tended to focus on wh-interrogatives rather than relative clauses. Rizzi (1982) offers a principled explanation for this choice, at least with respect to wh-islands: whereas Italian wh-interrogatives appear to demonstrate wh-islands as illustrated in (10), there is reason to believe that sentences with two or more wh-words in a single clause would be ruled out independently in Italian. Sentences like (11) are reported as unacceptable by Rizzi, despite no obvious island violation. Presumably this degradation is due to the fact that there are two wh-words (originating) in a single clause. Because wh-island effects are predicated upon two wh-words (or phrases) originating in a single clause, what at first glance appears to be a wh-island effect can in fact be captured by the prohibition against multiple wh-words (or phrases) operating in (11) without postulating any independent wh-island effect.[4]

(10) *Chi_1 ti domandi [chi_2 $\_\__2$ ha incontrato $\_\__1$]?
who to-yourself ask.2SG who has met
'Who do you wonder who met?' (Rizzi 1982:51, ex. 7a)

(11) *Mi domando [chi_2 $\_\__2$ ha incontrato chi_1].
to-myself ask.1SG who has met who.
'I wonder who met who.' (Rizzi 1982:51, ex. 8a)

There is a potential problem with Rizzi's (1982) idea that island effects in Italian wh-interrogatives might be illusory: not every island type involves a second wh-word (or phrase). For any island type that does not involve a second wh-word (or phrase) as part of its structural definition, it is possible to construct wh-interrogatives that involve that island type without running afoul of the prohibition against multiple wh-words (or phrases). If these sentences are indeed unacceptable, then that unacceptability either suggests a true island effect, or requires some other additional explanation beyond the prohibition against multiple wh-words (or phrases). As demonstrated in (2a), it is possible to construct a structure that is very similar to a wh-island that involves an embedded interrogative without using a true wh-word by substituting the complementizer *whether* (or *if*) for the wh-word in the specifier position

[4]We are reporting here Rizzi's judgments from the late 70s. Today, interrogative clauses like *Chi ha comprato cosa?* ('Who bought what?') are acceptable in many varieties of Italian, including journalistic jargon, possibly as a borrowing from English. If a syntactic transfer is taking place, it is still on-going. For example, *Chi ha comprato cosa?* sounds better than other combinations of wh-phrases (including those that are acceptable in English). At the present time, the distribution of different wh-in-situ phrases in Italian displays a complex pattern (see Moro 2011).

of the embedded CP. The other examples in (2) demonstrate that most other island types do not involve wh-words or phrases therefore can be investigated in Italian in wh-interrogatives without modification. Because Italian raises the possibility of variation between wh-interrogative and relative clause formation (cross-construction variation), we will investigate both constructions in this study by constructing wh-islands for rc-dependencies with full wh-words, and by constructing wh-islands for wh-dependencies with *whether* in English and *se* (*'if'*) in Italian.

As noted in Sect. 1, we postpone the discussion of theoretical analyses of these observations until Sect. 7. In Sect. 7, we review 6 types of theoretical analyses (covering 10 specific theories), and discuss the modifications to those analyses that would be necessary to accommodate the results of our experiments.

3 Isolating island effects with a factorial design

The factorial design of island effects makes explicit a fact about island effects that has long been implicit in the syntactic literature: sentences that give rise to island effects contain components that could lower acceptability independently of a grammatical island constraint. The idea behind the factorial design is to quantify these extra-grammatical components such that the effect of the grammatical constraint can be isolated (or, if one is agnostic about the source of island effects, the goal is to isolate the acceptability effect that cannot be accounted for by known effects). The factorial design we use in the current study explicitly isolates two non-syntactic components that could impact acceptability: (i) the effect of having a long-distance (often bi-clausal) dependency (e.g., a wh-dependency or a rc-dependency) in the sentence, and (ii) the effect of having a complex syntactic structure (what we call an *island structure*) in the sentence. Crucially, each of these components could potentially give rise to decrements in acceptability for reasons that are independent of syntactic island constraints. Long-distance dependencies tend to be more difficult to process than short-distance dependencies. If this processing difficulty impacts acceptability judgments, then sentences with long-distance dependencies will be rated lower than sentences with short-distance dependencies regardless of whether island constraints are violated. Similarly, island structures often involve more complex structures (e.g., complex NPs) or meanings (e.g., embedded interrogatives). If these structures and/or meanings impact acceptability judgments, sentences containing island structures will be rated lower than sentences that do not contain island structures regardless of whether extraction from islands takes place. What this means in practice is that for an island effect to be a phenomenon in need of a grammatical explanation, the island effect must be defined as a decrease in acceptability over and above the independent decreases caused by the individual components of the sentence. This conceptual definition of island effects has been at the heart of proposals for grammatical island constraints in the syntax and semantics literature. However, traditional studies of island effects have rarely attempted to isolate this decrease. The quantitative nature of experimental syntax techniques makes it possible to create a factorial design that fully instantiates this definition: it isolates the effect of long-distance dependencies, the effect of island structures, and any decrease over and above these two factors (see also Sprouse 2007; Sprouse et al. 2011, 2012).

As the name suggests, the factorial design treats the length of the dependency and the presence of island structures as two factors (GAP-POSITION and STRUCTURE), each with two levels (MATRIX/EMBEDDED and NON-ISLAND/ISLAND). Crossing the levels of these two factors results in four sentence (each sentence is a combination of one level from each factor), exemplified here for *whether*-island:

(12) A factorial design for measuring island effects: STRUCTURE × GAP-POSITION

a.	*Who* __ thinks [that John bought a car]?	NON-ISLAND | MATRIX
b.	*What* do you think [that John bought __]?	NON-ISLAND | EMBEDDED
c.	*Who* __ wonders [whether John bought a car]?	ISLAND | MATRIX
d.	*What* do you wonder [whether John bought __]?	ISLAND | EMBEDDED

The value of the factorial design lies in its ability to isolate each of the effects discussed above. The length effect is captured by [12a − 12b]. The structure effect is captured by [12a − 12c]. The island effect can then be isolated by first calculating the total effect [12a − 12d], and then subtracting the length and structure effect from the total effect. The island effect can also be calculated using a *differences-in-differences* (DD) score (Maxwell and Delaney 2003), which is calculated by subtracting the difference between two conditions related by one factor from the difference between the two conditions related by the other factor, such as DD = [12a − 12c] − [12b − 12d]. Both approaches are mathematically equivalent: if there is no island effect, the result will be 0; if there is an island effect, the result will be positive, and the size of the number will indicate the size of the island effect in the unit of measure of the ratings.

One particularly useful property of the subtraction process inherent in the factorial design is that it allows one to implicitly control for potential acceptability effects above and beyond the two that are explicitly instantiated in the design. As long as the property in question is distributed across two conditions, and those two conditions are distributed across either side of the minus sign in the definition of DD scores, the effect will subtract out when isolating the island effect. This gives two types of control to the factorial design: an explicit type that quantifies the effect of the two chosen factors (in this case, dependency length and structure), and an implicit type that can subtract out the effect of any number of other confounds.

Another useful property of the factorial design is that it lends itself to three procedures for identifying island effects. The first identification procedure is numerical: if the differences-in-differences score is greater than 0, there is an island effect (although exactly how much greater than 0 it must be is a question for syntactic theory). The second identification procedure is visual: if we plot the four conditions according to the two factors (in what is known as an interaction plot), the absence of an island effect will be indicated by two parallel lines as in the left panel of Fig. 1, and the presence of an island effect will be indicated by non-parallel lines as in the right panel of Fig. 1.[5]

The third identification procedure is statistical: if we perform a crossed, two-factor statistical test on the results (i.e., a two-way ANOVA or linear mixed effects model), the presence of an island effect will appear as a statistically significant, super-additive

[5]In Fig. 1 the length effect is represented by the downward slope of the lines, and the structure effect is represented by the vertical separation between the two lines.

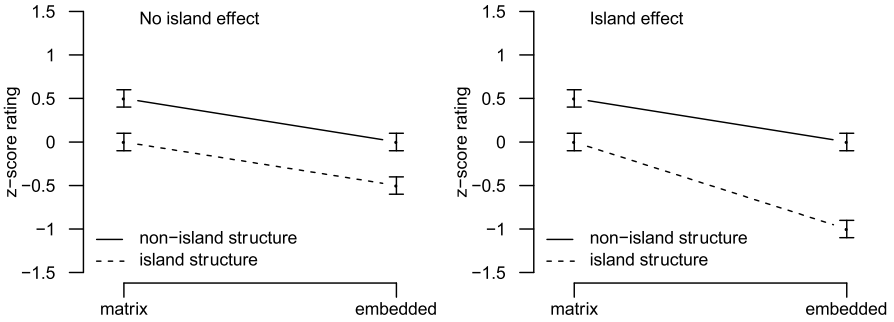

Fig. 1 The *left panel* demonstrates the pattern predicted when no island effect is present. The *right panel* demonstrates the pattern predicted when an island effect is present

interaction between the STRUCTURE and GAP-POSITION factors. We will report all three pieces of information in the results reported below: the DD score as a measure of island effect size, the interaction plots as a visual indicator of the island effects, and p-values derived from linear mixed effects models (using likelihood ratio tests) as a statistical measure of the significance of the interaction.

For completeness we should also mention that this factorial design was originally developed as a way to formalize the definition of island effects as part of an investigation of theories that seek to reduce island effects to an epiphenomenal consequence of processing complexity rather than the consequence of grammatical constraints (e.g., Kluender and Kutas 1993). The logic behind the use of the factorial design in this context is as follows. The simplest reduction of island effects to the processing complexity induced by long-distance dependencies and island structures will result in the linear additivity illustrated in the left panel of Fig. 1. Assuming that the effect of long-distance dependencies and island structures are due solely to processing complexity, the linear additive pattern would be unequivocal evidence that grammatical constraints on extraction from islands are unnecessary to explain the unacceptability of island effects. If instead we observe the super-additivity illustrated in the right panel of Fig. 1, then an additional mechanism (above and beyond the processing cost of long-distance dependencies and island structures) is necessary to explain the effect. This additional mechanism could either be a grammatical constraint, or it could be a mechanism that creates an interaction between the processing of long-distance dependency and island structures (e.g., the limited pool of working memory resources postulated by Kluender and Kutas 1993). Differentiating between these two possibilities requires an explicit theory of the mechanism(s) causing the interaction, and experiments designed to investigate those mechanism(s) (e.g., the working memory studies in Sprouse et al. 2012). These issues are beyond the scope of the present study, therefore here and throughout we assume the standard linguistic position that island effects are the result of grammatical constraints, but admit that this is an interesting question in its own right. Given this assumption, linear additivity suggests the absence of an island constraint, and super-additivity (plus the assumption that islands are the result of a grammatical constraint) suggests the presence of an island constraint.

4 The logic and design of the present study

Our goal is to apply the factorial design to both wh-dependencies and rc-dependencies in both English and Italian in order to probe the cross-linguistic and cross-constructional variation of island effects. We decided to investigate four island types: wh-islands, complex NP islands, subject islands, and adjunct islands. These island types were chosen because they have figured prominently in the literature on variation in island effects, both in Rizzi's (1982) original investigation, and in more recent surveys of cross-linguistic variation (e.g., Stepanov 2007). In this section we discuss the details of the design and construction of the 8 primary experiments.

4.1 The distribution of island types and dependency types across experiments

We constructed 8 primary experiments to test the cross-linguistic and cross-constructional variation in island effects: four in English and four in Italian. Each experiment tested one island type with wh-dependencies and one (distinct) island type with rc-dependencies. The idea was to minimize the number of long-distance dependencies (and island violations) presented to each participant, in order to avoid repetition effects. The distribution of island types and dependency types in each experiment was identical in both languages: As mentioned in Sect. 2.2, two different types of wh-islands were used in order to circumvent the double-wh-prohibition in Italian (Rizzi 1982): for wh-dependencies, which necessarily involve one wh-word, *whether*-islands were used in English and *se*-islands ('if') were used in Italian; for rc-dependencies, which do not unequivocally involve wh-words (or at least not the same wh-words as wh-interrogatives),[6] wh-islands involving full wh-words were used.

4.2 The construction of the target materials

The central challenge in cross-linguistic materials construction is to develop a set of materials that simultaneously respects the specific requirements of each language and minimizes the number of differences between the two languages. For the current study, one global decision that we made was to investigate rc-dependencies that form restrictive relative clauses introduced by a relative pronoun (as opposed to appositive relatives or restrictive relatives introduced by a complementizer, which we leave to future research). This had two major consequences. First, all of the rc-dependencies in Italian had to be constructed with oblique (PP) argument gaps because Italian restrictive relatives can only be introduced by a relative pronoun when the head of the relative is an oblique argument (subjects and direct objects can only be introduced by the complementizer *che* 'that'). This led to a systematic difference between Italian rc-dependencies, which were constructed with oblique argument gaps, and English rc-dependencies, which were constructed with direct object gaps.

Second, the lack of preposition stranding in Italian meant that the gap locations of the oblique (PP) arguments were not unambiguously signaled. For wh-islands, complex NP islands, and adjunct islands, we were able to minimize potential ambiguity

[6] Italian relative clauses are introduced by the same complementizer as embedded declarative clauses or by two different series of relative pronouns that are different from the wh-words that occur in wh-interrogative clauses.

through the careful selection of predicates. However, a pilot experiment suggested that this was not possible for subject islands because of the presence of a second NP (the object) that could be modified by the displaced PP. To remedy this, we decided to modify the factorial definition of subject islands slightly. Whereas the factor GAP-POSITION in the canonical factorial definition in (12) varies between MATRIX-SUBJECT and EMBEDDED-SUBJECT gap positions, the modified subject island definition varies between EMBEDDED-OBJECT and EMBEDDED-SUBJECT gap positions. This allows us to include a PP adjunct with both the subject and object NP, minimizing the potential for the errant interpretation of the gap location of the displaced PP. This modified design closely resembles the definitions used in traditional syntax studies, but crucially results in a non-monotonic interaction rather than a monotonic interaction (the non-parallel lines have slopes in opposite directions). In order to minimize the differences caused by this modification, subject islands in both languages and both dependency types used the modified definition, but all other island types and dependency types used the canonical design. Example materials are provided in (13)–(28) below (and the full set of materials are available on the first author's website).

One anonymous reviewer wonders whether the fact that Italian subject islands involve PP extraction, while English subject islands involve DP extraction, could explain why our results (see Sect. 5) show no evidence for a subject island effect in Italian rc-dependencies. This is a great example of the implicit control afforded by the factorial design that was mentioned in Sect. 3. If we perform a DD-score calculation as discussed in Sect. 3, we can see that each side of the minus sign calculates an effect for the difference between DP extraction and PP extraction: DD = [DP_{object} − PP_{object}] − [$DP_{subject}$ − $PP_{subject}$]. Because these effects are on either side of the minus sign, they will subtract out when we calculate the DD score. This means that the superadditive component we see isolates the difference between extraction from an object and extraction from a subject, which is exactly what the definition of subject islands calls for. The same logic holds for concerns about whether the fact that all four conditions in English subject islands involve DP extraction (never PP extraction). The effect of DP extraction will subtract out when the DD score is calculated, so it is not part of the result. Anticipating our results slightly (see Sect. 5), we can also see empirical corroboration of this logic: the fact that Italian wh-dependencies do show a significant subject island effect while still using PP extraction shows that PP extraction alone is not a viable explanation for the presence/absence of subject island effects in Italian.

The same reviewer also wonders whether our choice to use conditional clauses as the adjunct structure in adjunct islands could have had an impact on the results. First, we agree with the reviewer that it would be interesting to test other types of adjuncts (e.g., causal adjuncts and temporal adjuncts). In this case, we chose conditional adjuncts because previous studies have demonstrated that conditional adjuncts are not transparently reducible to a processing complexity effect (Sprouse et al. 2012), therefore conditional adjuncts are a particularly good candidate for investigations of cross-linguistic variation in grammatical theories of island effects. Second, as the previous paragraph touched upon, the factorial design allows us to test any structure that we may be interested in, and crucially isolate a superadditive component that goes above and beyond the cost of the structure itself. This means that there is no way for the

choice of structure to lead to a confound, as long as we respect the properties of the factorial design. In fact, in principle, the factorial design could be used to identify completely novel island effects this way.

In the end, the differences between English and Italian with respect to pied-piping/preposition-stranding, and our decision to focus on restrictive relatives introduced by relative pronouns led to two systematic differences in materials that should be noted: (i) in order to ensure the correct gap location, subject islands use a modified factorial design (while wh-islands, complex NP islands, and adjunct islands use the canonical factorial definition), and (ii) rc-dependencies in English involved direct object gaps, while rc-dependencies in Italian involved oblique (PP) argument gaps. Beyond these two systematic differences, we attempted to keep all other aspects of the materials uniform whenever possible, and to distribute any known possible confounds across the factorial design to take advantage of the subtraction logic. Each factorial island design consists of 4 lexically matched conditions, which helps to minimize differences across conditions due to lexical content. We constructed 8 distinct quadruplets for each of the four islands, for each of the two dependency types, and for each of the two languages, for total of 128 distinct quadruplets, or 512 total sentences. The full list of materials is available on the first author's website.

ENGLISH: wh-dependencies

(13) WHETHER ISLANDS
- a. *Who* __ thinks that John bought a car?
- b. *What* do you think that John bought __?
- c. *Who* __ wonders [whether John bought a car]?
- d. *What* do you wonder [whether John bought __]?

(14) COMPLEX NP ISLANDS
- a. *Who* __ heard that Jeff baked a pie?
- b. *What* did you hear that Jeff baked __?
- c. *Who* __ heard [the statement that Jeff baked a pie]?
- d. *What* did you [hear the statement that Jeff baked __]?

(15) SUBJECT ISLANDS
- a. *What* do you think the gift prompted __?
- b. *What* do you think __ prompted the rumor?
- c. *Who* do you think the gift from the lobbyist prompted the rumor about __?
- d. *Who* do you think [the gift from __] prompted the rumor about the Senator?

(16) ADJUNCT ISLANDS
- a. *Who* __ thinks that the lawyer forgot his briefcase at the office?
- b. *What* do you think that the lawyer forgot __ at the office?
- c. *Who* __ worries [if the lawyer forgets his briefcase at the office]?
- d. *What* do you worry [if the lawyer forgets __ at the office]?

ENGLISH: rc-dependencies

(17) WH-ISLANDS

a. I take classes with the *professor who* __ thinks that Paul will tutor the struggling student.
b. I take classes with the *struggling student who* the professor thinks that Paul will tutor __.
c. I take classes with the *professor who* __ wonders [when Paul will tutor the struggling student].
d. I take classes with the *struggling student who* the professor wonders [when Paul will tutor __].

(18) COMPLEX NP ISLANDS

a. I know the *fisherman who* __ heard that Laura is dating the boat captain.
b. I know the *boat captain who* the fisherman heard that Laura is dating __.
c. I know the *fisherman who* __ heard [the rumor that Laura is dating the boat captain].
d. I know the *boat captain who* the fisherman heard [the rumor that Laura is dating __].

(19) SUBJECT ISLANDS

a. I voted for the *congressman who* you think the lobbyist offended __.
b. I voted for the *congressman who* you think __ offended the lobbyist.
c. I voted for the *congressman who* you think the gift from the lobbyist prompted the rumor about __.
d. I voted for the *congressman who* you think the gift from __ prompted the rumor about bribery.

(20) ADJUNCT ISLANDS

a. I called the *secretary who* __ thought that the lawyer insulted the client.
b. I called the *client who* the secretary thought that the lawyer insulted __.
c. I called the *secretary who* __ worries [if the lawyer insults the client].
d. I called the *client who* the secretary worries [if the lawyer insults __].

ITALIAN: wh-dependencies

(21) WHETHER ISLANDS

a. *Chi* __ pensa che io abbia letto il libro?
who thinks that I have.SUBJ.1SG read the book
'Who thinks I read the book?'
b. *Cosa* pensi che io abbia letto __?
what think.2SG that I have.SUBJ.1SG read
'What do you think I read?'

c. *Chi* __ si chiede [se io abbia letto il libro]?
who to_himself asks if I have.SUBJ.1SG read the book
'Who wonders if I read the book?'

d. *Cosa* ti chiedi [se io abbia letto __]?
what to_yourself ask.2SG if I have.SUBJ.1SG read
'What do you wonder if I read?'

(22) COMPLEX NP ISLANDS

a. *Chi* __ ha affermato che io avrei rubato una macchina?
who has claimed that I have.IRR.1SG stolen a car
'Who claimed that I stole a car?'

b. *Cosa* hai affermato che io avrei rubato __?
what have.2SG claimed that I have.IRR.1SG stolen
'What did you claim that I stole?

c. *Chi* __ ha fatto [l'affermazione che io avrei rubato una
who has made the-claim that I have.IRR.1SG stolen a
macchina]?
car
'Who made the claim that I stole a car?'

d. *Cosa* hai fatto [l'affermazione che io avrei
what have.2SG made the-claim that I have.IRR.1SG
rubato __]?
stolen
'What did you make the claim that I stole?'

(23) SUBJECT ISLANDS

a. *Chi* pensi che il quadro raffiguri __?
who think.2SG that the painting depict.SUBJ.3SG
'Who do think that the painting portrays?'

b. *Chi* pensi che __ abbia dipinto il quadro?
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 di Maria] raffiguri la
of who think.2SG that the painting on-the wall depicts the
nascita __?
birth
'Who do you think the painting on the wall depicts the birth of?'

d. *Di chi* pensi che [il quadro __] raffiguri la nascita di Venere?
of who think.2SG that the painting depicts the birth of Venus
'Who do you think the painting of depicts the birth of Venus?'

(24) ADJUNCT ISLANDS

a. *Chi* __ dice che io abbia usato il cellulare in classe?
who says that I have.SUBJ.SG used the cell-phone in class
'Who says that I used my cell phone in class?'

b. *Cosa* dici che io abbia usato __ in classe?
what say.2SG that I have.SUBJ.SG used in class
'What do you say that I used in class?'

c. *Chi* __ si infuria [se uso il cellulare in classe]?
who himself infuriates if use.1SG the cell-phone in class
'Who is going to get infuriated if I use my cell phone in class?'

d. *Cosa* ti infuri [se uso __ in classe]?
what yourself infuriates.2SG if use.1SG in class
'What are you going to get infuriated if I use in class?'

ITALIAN: rc-dependencies

(25) WH-ISLANDS

a. Ieri ho visto il *poliziotto del quale* si dice __ che
yesterday have.1SG seen the policeman of_the which one says that
sia innamorato di Lara.
be. SUBJ.SG in_love of Lara
'Yesterday I saw the policeman that they say is in love with Lara.'

b. Ieri ho visto il *poliziotto del quale* si dice __ che
yesterday have.1SG seen the policeman of_the which one says that
Lara sia innamorata.
Lara be.SUBJ.SG in-love
'Yesterday I saw the policeman that they say is in love with Lara.'

c. Ieri ho visto il *poliziotto al quale* ho
yesterday have.1SG seen the policeman to_the which have.1SG
domandato __ [perché sia innamorato di Lara].
asked why be.SUBJ.SG in_love of Lara
'Yesterday I saw the policeman who I asked why he is in love with Lara.'

d. Ieri ho visto il *poliziotto del quale* ti
yesterday have.1SG seen the policeman of-the which to_you
ho domandato [perché Lara sia innamorata __].
have.1SG asked why Lara be.SUBJ.SG in_love
'Yesterday I saw the policeman who I asked why Lara is in love with.'

(26) COMPLEX NP ISLANDS

a. Ho telefonato all'*uomo a cui* hai fatto notare __
have.1SG phoned to_the man to whom have.2SG made notice
che Andrea ha un atteggiamento ostile verso di noi.
that Andrea has a behavior hostile towards of us
'I called the man you pointed out to that Andrea has a hostile attitude towards us.'

b. Ho telefonato all'*uomo verso il quale* hai fatto
have.1SG phoned to_the man towards the whom have.2SG made
notare che Andrea ha un atteggiamento ostile __.
notice that Andrea has an attitude hostile

'I called the man towards whom you pointed out that Andrea has a hostile attitude.'

c. Ho telefonato all'*uomo a cui* hai fatto notare __
have.1SG phoned to_the man to whom have.2SG made notice
[il fatto che Andrea ha un atteggiamento ostile verso di noi].
the fact that Andrea has a behavior hostile towards of us
'I called the man you pointed out to the fact that Andrea has a hostile attitude towards us.'

d. Ho telefonato all'*uomo verso il quale* hai fatto
have.1SG phoned to_the man towards the whom have.2SG made
notare [il fatto che Andrea ha un atteggiamento ostile __].
notice the fact that Andrea has a behavior hostile
'I called the man towards whom you pointed out the fact that Andrea has a hostile attitude.'

(27) SUBJECT ISLANDS

a. Ho incontrato il *giornalista che* pensi che il direttore
have.1SG met the journalist that think.2SG that the director
abbia fatto licenziare __.
have.SUBJ.SG made fired
'I met the journalist that you think that the director fired.'

b. Ho incontrato il *giornalista che* pensi che __
have.1SG met the journalist that think.2SG that
abbia fatto arrabbiare il direttore.
have.SUBJ.SG made angry the director
'I met the journalist that you think pissed off the director.'

c. Ho incontrato il *giornalista del quale* pensi che
have.1SG met the journalist of_the whom think.2SG that
[l'articolo del direttore] abbia causato il
the article of_the director have.SUBJ.SG caused the
licenziamento __.
firing
'I met the journalist who you think the director's article has causes the firing of.'

d. Ho incontrato il *giornalista del quale* pensi che
have.1SG met the journalist of_the whom think.2SG that
[l'articolo __] abbia causato il licenziamento del
the article have.SUBJ.SG caused the firing of_the
direttore.
director
'I met the journalist who you think that the article of caused the firing of the director.'

(28) ADJUNCT ISLANDS

a. Hanno intervistato il *manager a cui* ho detto __ che
have.3PL interviewed the manager to whom have.1SG said that
Andrea ha fatto nuove rivelazioni su di lui.
Andrea has made new revelations on of him
'They interviewed the manager who I told that Andrea made new revelations about him.'

b. Hanno intervistato il *manager su cui* mi hanno detto
have.3PL interviewed the manager on whom to_me have.3PL said
che Andrea ha fatto nuove rivelazioni __.
that Andrea has made new revelations
'They interviewed the manager they told me that Andrea has made new revelations about.'

c. Hanno intervistato il *manager a cui* abbasseranno lo
have.3PL interviewed the manager to whom lower.FUT.3PL the
stipendio __ [se Andrea farà nuove rivelazioni].
stipend if Andrea make.FUT.3SG new revelations
'They interviewed the manager whose salary they will lower the salary if Andrea makes new revelations.'

d. Hanno intervistato il *manager su cui* ci abbasseranno lo
have.3PL interviewed the manager on whom to-us lower.FUT.3PL the
stipendio [se Andrea farà nuove rivelazioni __].
stipend if Andrea make.FUT.3SG new revelations
'They interviewed the manager who they will lower our salary if Andrea makes new revelations on.'

4.3 The construction of the fillers

Each experiment contained 32 filler items, which led to a 2:1 ratio of filler to target items (there were 16 target items per experiment: 2 island types × 4 conditions × 2 tokens per condition). To construct the filler items for the four English experiments, we selected 32 sentence types (out of 300) from Sprouse, Schütze, and Almeida (2013). The 32 sentence types were selected such that they evenly span the complete range of acceptability observed in those experiments, with the additional constraint that 12 of the filler types came from the acceptable side of the spectrum, and 20 came from the unacceptable side of the spectrum. Under the assumption that three conditions per island type are acceptable, and one condition per island type is unacceptable, this distribution of filler items would lead to a 1:1 ratio of acceptable items to unacceptable items in each experiment. Because Sprouse, Schütze, and Almeida (2013) created multiple tokens of each sentence type, we were able to use a distinct token for each filler sentence type for each of the four English experiments. While this introduces a small amount of variability across the four English experiments, it has the benefit of making each of the four experiments completely (lexically) distinct. This is useful when running experiments on Amazon Mechanical Turk, as it is possible for participants to participate in more than one experiment. For the four Italian experiments, we translated each of the 32 English fillers into Italian and then used our own

Table 1 The distribution of island types and dependency types tested per experiment (in both English and Italian)

Experiment	wh-Dependency	rc-Dependency
1	*whether* island	subject island
2	subject island	wh-island
3	complex NP island	adjunct island
4	adjunct island	complex NP island

judgments to verify that the distribution of acceptability remained as intended (1:1 acceptable to unacceptable, relatively evenly distributed across the complete range of acceptability). Because the four Italian experiments were administered in person, only one set of filler items was used in all four experiments. The full list of fillers is available along with the materials on the first author's website.

4.4 The construction of the surveys

We distributed each island and dependency type into 8 experiments according to the schema in Table 1. For each experiment, we distributed the 8 tokens per conditions into 4 lists using a Latin Square design such that each list contains 2 tokens of each condition, and none of the tokens of target conditions were lexically related. The 16 target items in each list were then combined with 32 filler items for a total of 48 items per list. The 48 items in each list were then pseudo-randomized such that target items from the same island type never appeared in immediate succession. The same 6 practice items were added to the beginning of each list to allow participants to familiarize themselves with the rating scale before rating target items. These 6 items were distributed across the range of acceptability (two acceptable, two moderate, two unacceptable). These practice items were not marked as practice items, so from the perspective of the participants they were simply part of the survey. This construction procedure resulted in 4 surveys for each of the 8 experiments (32 surveys total), with each survey being 54 items long, and containing two ratings for each of the four conditions for two island types. The task in all eight experiments was a 7-point Likert scale task, with 1 at the low end and 7 at the high end of acceptability. The instructions for the task were the same as Sprouse, Schütze, and Almeida (2013), which are available on the first author's website.

4.5 Participants

The four English experiments were conducted using Amazon Mechanical Turk (see Sprouse 2011 for an evaluation of AMT for acceptability judgment collection). 56 participants were recruited for each experiment (224 total). Participants were paid $2 for their participation. Participant selection criteria were enforced as follows. First, the AMT interface automatically restricted participation to AMT users with a US-based location. Second, we included two questions at the beginning of the experiment to assess language history: (1) Were you born and raised in the US?, (2) Did both of your parents speak English to you at home? These questions were not used to

determine eligibility for payment, consequently there was no incentive to lie. Four participants were excluded from analysis, one in each of the four experiments, for failing to answer yes to both language history questions. All of the analyses reported below are on the 55 participants remaining in each experiment.

The four Italian experiments were conducted as part of a large undergraduate class at the University of Milano-Bicocca. 195 participants were recruited during a single class session. Because Italian educational policies prevent payment for participation in research experiments, all participants were volunteers. Two participants were excluded from the analysis for being non-native speakers of Italian. The distribution of participants across the four experiments is: 50, 49, 47, and 47.

5 Results

The results of all eight experiments were analyzed identically. First, the raw ratings from each participant were z-score transformed. The z-score transformation eliminates certain kinds of scale biases between participants (e.g., using one end of the scale, or using a larger or smaller range of values) by converting each participant's ratings to a standardized scale (each transformed rating represents the number of standard deviations the raw rating was from the participant's mean rating). Next, we constructed linear mixed effects models with items and participants included as random factors on each of the island types using GAP-POSITION and STRUCTURE as fixed factors. This is comparable to a repeated-measures two-way ANOVA, but with participants and items entering the model simultaneously. We then calculated p-values for the two main effects and the interaction term using likelihood ratio tests. We decided to use linear mixed effects models because of their current popularity among some experimentalists; however, it should be noted that the theoretical appropriateness of treating the items in acceptability judgment experiments as a random effect is far from settled (see Wike and Church 1976 and other articles in that volume). As such, these statistical tests may be overly conservative (i.e., the p-values reported here may be too high). Finally, we calculated differences-in-differences scores for each participant, and then calculated mean differences-in-differences scores for each island as a (non-standardized) effect-size measure for each island type.

5.1 English wh-dependencies

For English wh-dependencies, the four experiments revealed significant super-additive interactions for *whether*, complex NP, and adjunct islands, and a nearly-significant interaction ($p = .062$) for subject islands. Given that subject islands with English wh-dependencies have demonstrated significant interactions in at least three previous studies (Sprouse 2007; Sprouse et al. 2011, 2012), we take this nearly significant result to be a consequence of the modified factorial design used here. We therefore choose to interpret the nearly significant result as theoretically significant. As such, these results replicate previous findings using the factorial definition of island effects (Sprouse 2007; Sprouse et al. 2011, 2012). Figure 2 presents interaction plots for each of the island types, along with p-values for the interaction term of the linear mixed effects models, and DD scores as a measure of effect size. The full set of raw data is available on the first author's website.

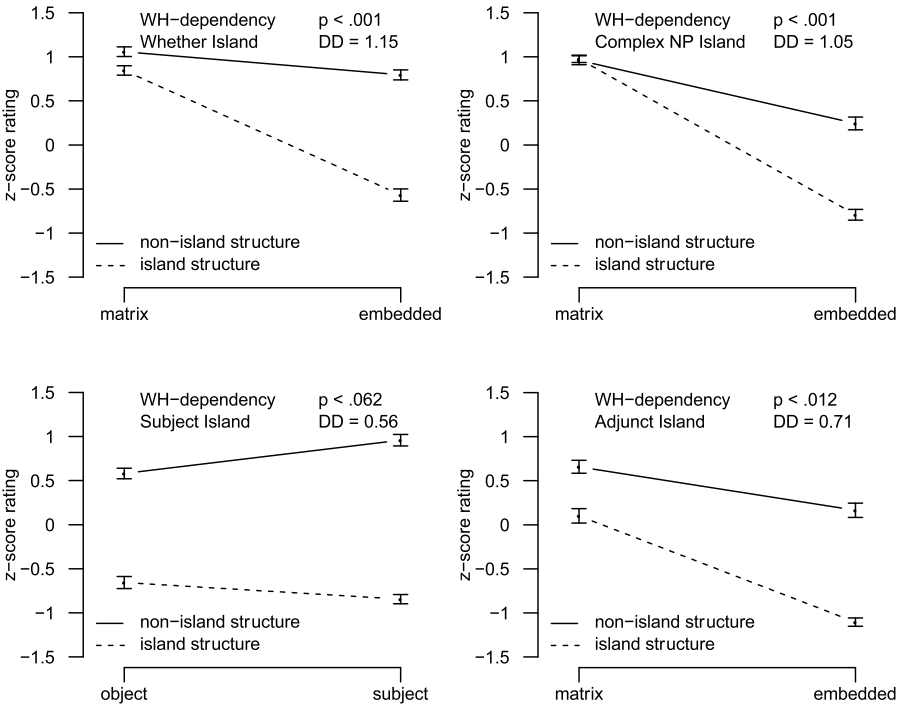

Fig. 2 English wh-dependencies. Interaction plots for each island type, with p-values (based on likelihood ratio tests) and DD scores

5.2 English rc-dependencies

For English rc-dependencies, the four experiments revealed significant super-additive interactions for wh, complex NP, and subject islands. In contrast, adjunct islands demonstrated nearly perfect linear additivity (a p-value of .992 and DD score of .01). One interesting property to note is that the island-violating sentence in the adjunct island design is rated relatively unacceptable (around $-.75$). This unacceptability could explain why adjunct islands have been assumed to be present for English rc-dependencies, as it is only after using the factorial design that it becomes clear that this unacceptability can be completely explained by the linear sum of the effect of long-distance dependencies and the effect of island structures. Figure 3 presents interaction plots for each of the island types, along with p-values for the interaction term of the linear mixed effects models, and DD scores as a measure of effect size. The full set of raw data is available on the first author's website.

5.3 Italian wh-dependencies

For Italian wh-dependencies, the four experiments revealed significant super-additive interactions for all four island types. Figure 4 presents interaction plots for each of the island types, along with p-values for the interaction term of the linear mixed effects

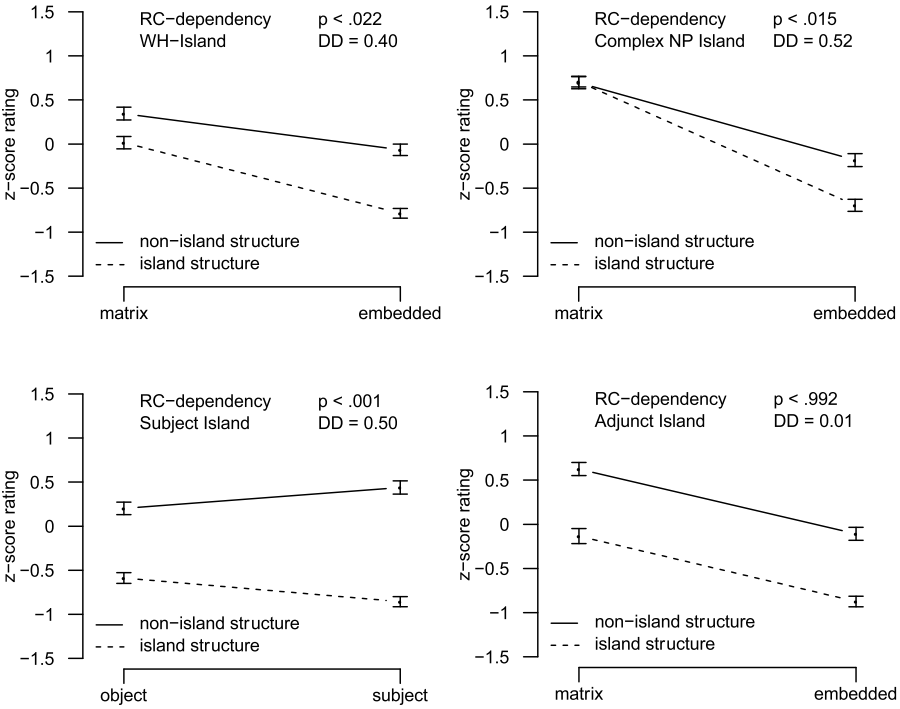

Fig. 3 English rc-dependencies. Interaction plots for each island type, with p-values (based on likelihood ratio tests) and DD scores

models, and DD scores as a measure of effect size. The full set of raw data is available on the first author's website.

5.4 Italian rc-dependencies

For Italian rc-dependencies, the four experiments revealed significant super-additive interactions for wh, complex NP, and adjunct islands. In contrast, subject islands revealed nearly perfect linear additivity (a p-value of .84 and a DD score of −.07). Figure 5 presents interaction plots for each of the island types, along with p-values for the interaction term of the linear mixed effects models, and DD scores as a measure of effect size. The full set of raw data is available on the first author's website.

5.5 Summary

To sum up, the goal of this study was to apply the factorial definition of island effects to four island types (wh/*whether*, complex NP, subject, and adjunct), two dependency types (wh and rc) and two languages (English and Italian) in order to investigate the variation of island effects across constructions and across languages. To that end, we conducted 8 acceptability judgment experiments testing these 8 combinations of island types and dependency types. The results of these 8 experiments are summarized in Table 2.

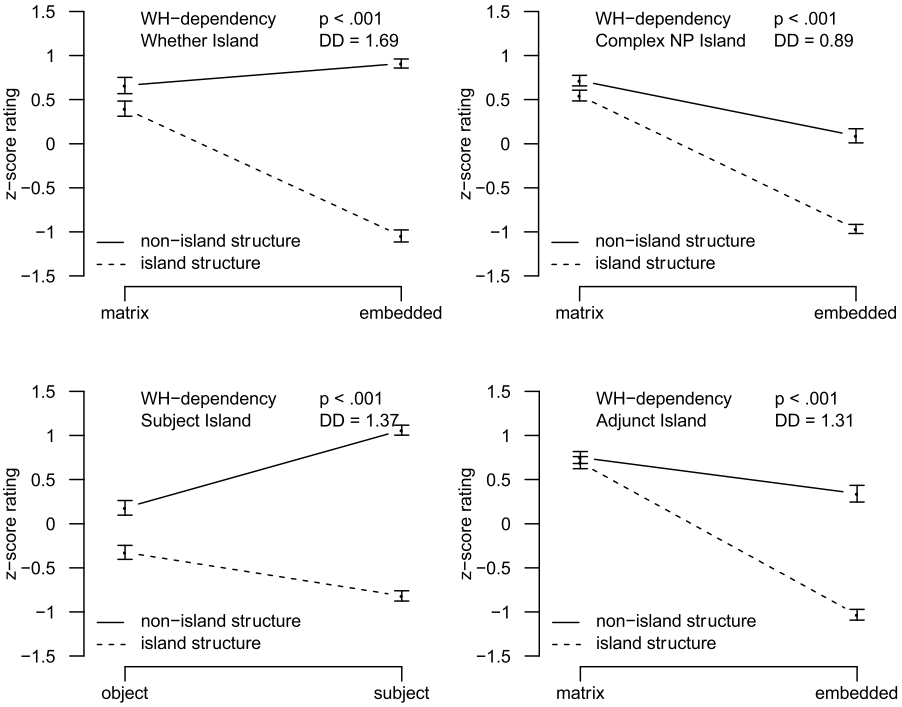

Fig. 4 Italian wh-dependencies. Interaction plots for each island type, with p-values (based on likelihood ratio tests) and DD scores

Table 2 Summary of the results of the English and Italian experiments. A plus sign (+) indicates a significant super-additive interaction suggesting an island effect; a minus sign (−) indicates no evidence of a super-additive interaction suggesting no island effect. Minus signs are in *bold*

Language	Dependency	wh/*whether*	Complex NP	Subject	Adjunct
English	wh	+	+	+	+
	rc	+	+	+	**−**
Italian	wh	+	+	+	+
	rc	+	+	**−**	+

As Table 2 indicates, the results of these experiments suggest that there is variation both across dependency types and across languages: while both English and Italian show evidence of an island effect for all four island types with wh-dependencies, English rc-dependencies do not show adjunct island effects, and Italian rc-dependencies do not show subject island effects.[7]

[7]One logically possible explanation for the variation observed in subject and adjunct islands is that the materials were confounded in the items that showed the variation. For example, the lack of subject island effects with Italian rc-dependencies could be explained if those materials (and only those materials)

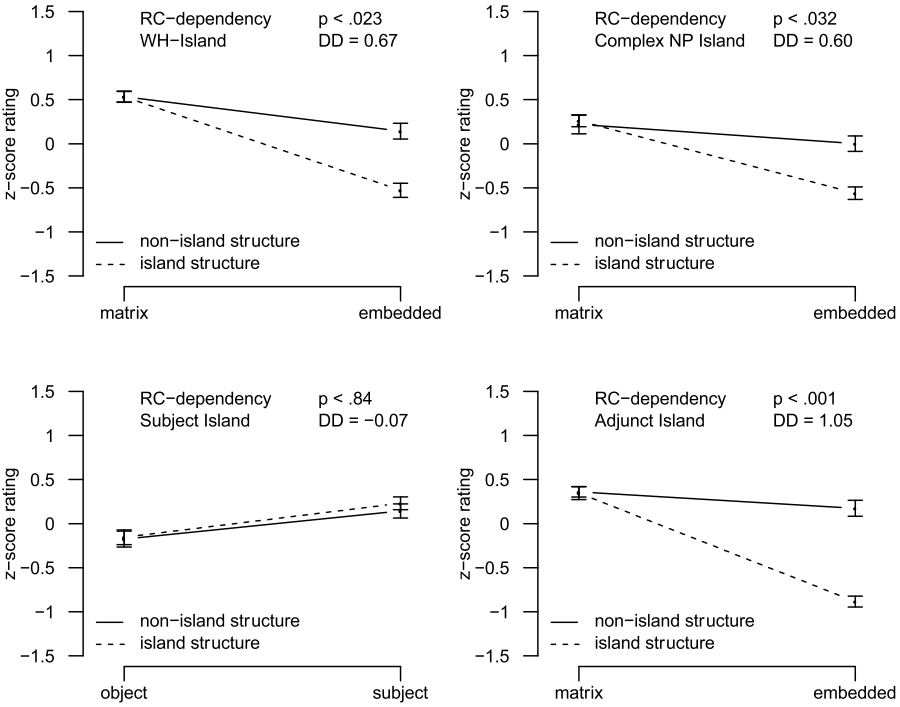

Fig. 5 Italian rc-dependencies. Interaction plots for each island type, with *p*-values (based on likelihood ratio tests) and DD scores

6 Dependency type versus featural specification: further experimental evidence

The two dependency types tested in our experiments vary not only by type (wh vs. rc), but also by featural specification of the head of the dependency. In particular, wh-dependencies are introduced by bare wh-words while rc-dependencies are introduced by a combination of a nominal head followed by a wh-word. This raises the question: Was the variation in island effects observed in our experiments due to the dependency type or due to the featural specification of the heads?[8] To attempt to disentangle these two properties in English, we decided to test wh-dependencies that are introduced by complex wh-phrase in which a wh-word combines with a nominal (e.g., *which car*) with all four island types. In principle, complex wh-phrases might

contained post-verbal subjects instead of pre-verbal subjects. And the lack of adjunct island effects with English rc-dependencies could be explained if those materials (and only those materials) contained complement *if*-clauses instead of adjunct *if*-clauses. We believe that this sort of explanation (in which the results are the consequence of a confound) is extremely unlikely due to the careful nature of our materials construction, therefore in the discussion that follows we will take the results at face value. We have posted the entire set of materials on the first author's website so that interested readers can assess the likelihood of these confounds for themselves.

[8]Thanks to Norbert Hornstein (p.c.) and one anonymous reviewer for bringing this point to our attention.

behave differently from bare wh-words for two reasons. First, they are more likely to be D-linked, i.e., their interpretation crucially depends on contextually salient sets of individuals. Second, intervention effects for complex wh-phrases have been observed to be different from intervention effects with bare wh-words, at least in child grammar (see Friedmann et al. 2009). These two properties of complex wh-phrases have led to the use of at least two terms in the literature to refer to these phrases: *D-linked* wh-phrases, which tends to be used to highlight the D-linking property, and *lexically restricted* wh-phrases, which tends to be used to highlight the contribution of the noun to Relativized Minimality effects (e.g., Friedmann et al. 2009). Teasing apart these properties is an interesting research topic in its own right; however, for the current study we do not attempt to do so. Instead, we simply rely on the featural similarity between complex wh-phrases and relative clause head nouns. Hence we use the term *complex* wh-phrases to remain neutral with respect to the property driving any potential effect on island effects.

The overlapping similarities between complex wh-phrases and bare wh-words on one hand, and complex wh-phrases and relative clause heads on the other, make the following predictions. If it is the dependency type that is driving the variation, then wh-dependencies with complex wh-phrases will display the same pattern of island effects as wh-dependencies with bare wh-words; if it is the featural specification that is driving the variation, then wh-dependencies with complex wh-phrases will display the same pattern of island effects as rc-dependencies (regardless of whether it is ultimately D-linking, intervention, or even processing difficulty issues that are underlying the featural specification effect). To test this question we simply changed the bare wh-words in the first four English experiments to complex wh-phrases (the full set of materials is on the first author's website). We kept all other details of the experiment identical, including the number of participants tested (56 for each of 4 experiments), the method of recruitment (Amazon Mechanical Turk), the details of the surveys (e.g., the order of presentation of items), and the analysis of the results (z-score transformations and linear mixed effects models). One welcome consequence of this minimal change in experimental materials is that these four experiments serve as both a test of complex wh-dependencies and a replication of the rc-dependency results reported in Sect. 5 (because the rc-dependency materials are unchanged in these new experiments). Given that it was the rc-dependency results that differed from previous reports in the literature, this replication is an important step in establishing the reliability of these results. We report both the D-linked wh-dependency results and the rc-dependency replication results in the two subsections below.

6.1 wh-Dependencies with complex wh-phrases and island effects

In the first set of experiments, English wh-dependencies displayed significant interactions for all four island types, while English rc-dependencies displayed significant interactions for all but adjunct islands. As Fig. 6 indicates, English complex wh-dependencies pattern with bare wh-word wh-dependencies in displaying significant interactions for all four island types, including adjunct islands, and do not pattern with rc-dependencies. This suggests that the variation observed in adjunct island results between wh-dependencies and rc-dependencies is indeed due to the dependency type, and not due to the featural specification of the heads of the dependencies.

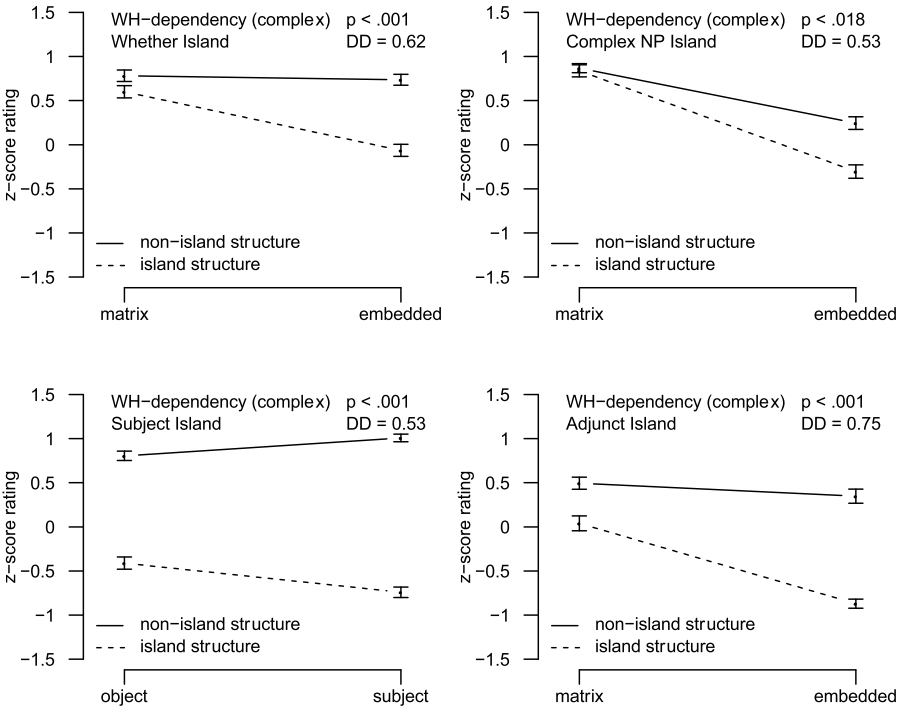

Fig. 6 English wh-dependencies with complex wh-phrases. Interaction plots for each island type, with *p*-values (based on likelihood ratio tests) and DD scores

To our knowledge this is the first published report of the results of using the factorial design to test island effects with complex wh-phrases. As such, a couple of notes about these results are in order. First, everything else being equal, a complex wh-phrase is more D-linked than the corresponding bare wh-word. Although it is occasionally claimed in the literature that D-linked wh-phrases ameliorate island effects, the factorial design reveals that superadditive interactions are present for all four island types tested (whether, complex NP, subject, and adjunct islands). This suggests that whatever the amelioration of D-linking is, it is not enough to completely eliminate the superadditive island effect. That being said, *whether* islands and complex NP islands do show a specific type of amelioration effect: the island-violating sentences for *whether* and complex NP islands are rated higher with complex wh-phrases (mean z-scores near 0) than they are with bare wh-words (mean z-scores near −.5; see Fig. 2). These rating increases lead to a concomitant decrease in effect sizes (DD scores): whether and complex NP islands with bare wh-words have DD scores of about 1.15 and 1.05 respectively, while with D-linked wh-phrases they have DD scores of about .6 and .5 respectively. This suggests that there is a type of amelioration effect on island-violating sentences, but not enough to completely eliminate the island effect itself. This amelioration appears to be specific to the island-violating sentences, as there does not appear to be much of a difference in the other three (grammatical) sentences in the factorial design. In addition, there appears to be no

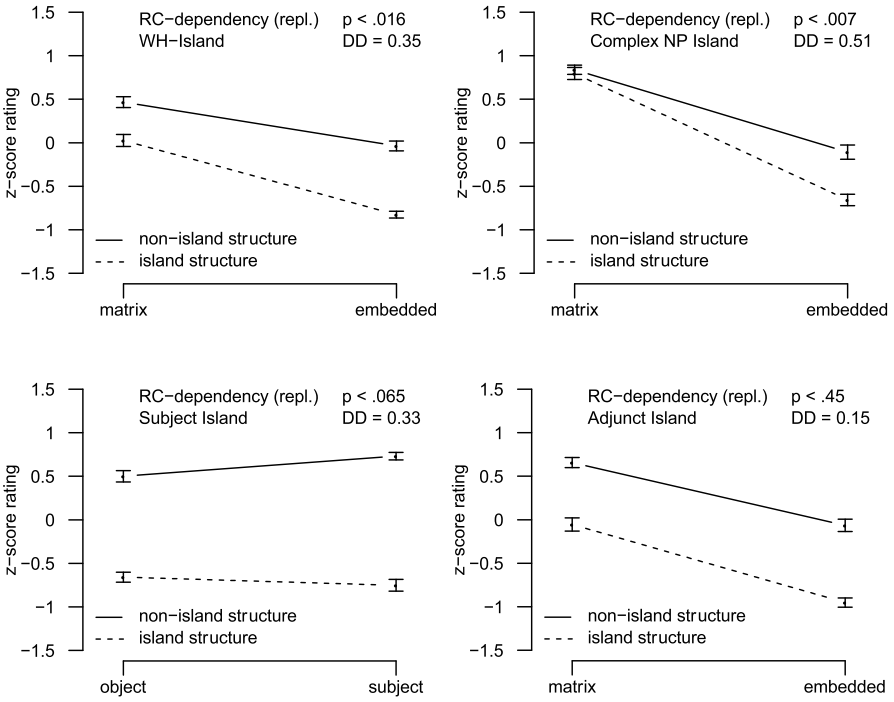

Fig. 7 Replication of English rc-dependency results. Interaction plots for each island type, with p-values (based on likelihood ratio tests) and DD scores

amelioration effect of any kind for subject and adjunct islands. Taken together with the specific effect on island-violating sentences, this result accords well with the distinction between strong and weak islands in the literature (for a review see Szabolcsi 2006). It is also interesting to note that rc-dependencies do not show this amelioration effect. This result suggests that the amelioration is specific to D-linking, and not a general effect of featural specification. A complete characterization of the D-linking effect under the factorial design is beyond the scope of this article; however, one potentially interesting consequence of the pattern of results obtained here is that this D-linking effect may be specific to island-violating sentences and not a general effect of featural specification on sentence processing (as has been claimed in the literature, e.g., Hofmeister and Sag 2010). We leave a detailed investigation of this effect to future research.

6.2 Replication of rc-dependency results

Turning next to the replication of the rc-dependency results that were collected along with results from wh-dependencies with complex wh-phrases, we see in Fig. 7 that the results nearly perfectly replicate. The replication shows significant interactions for wh-islands and complex NP islands, with similar effect sizes. The replication also shows the same lack of adjunct island effect. The only small change in the replication is that the p-value for the subject island interaction is .065, which is slightly

above the standard .05 criterion, and instead in the "marginal" range (similar to the marginal result for subject islands with English wh-dependencies in the first set of experiments). This result may be a consequence of the altered factorial design for subject islands (necessitated by the lack of pied-piping in Italian), which may contain a small garden-path effect in the third condition (island-object), resulting in an underestimate of the size of the subject island effect. This small difference aside, the replication of the wh-island and complex NP island effects, and the crucial lack of evidence for an adjunct island effect, provides additional confidence in the results of our first set of experiments.

7 Consequences for existing theories of island effects

In this section we present a brief review of several syntactic theories of island effects and discuss the consequences of our experimental results for each of them. First, we focus solely on *syntactic* theories of islands effects, setting aside non-syntactic theories, such as semantic and pragmatic theories of island effects, because the variation observed in our experiments does not revolve around island types that have figured prominently in the non-syntactic islands literature.[9] Second, the literature on island effects is large and ever expanding, and we won't be able to discuss all of the analyses that have been proposed. In anticipation of the results of our experiments, we focus on theories that are particularly relevant for subject islands and adjunct islands, as these are the points of variation that we observe in our results. To this extent, many existing proposals can be divided into two major groups with respect to how they account for subjects and adjunct island effects. On the one hand, there are approaches which tie island effects to the structural distinction between complements and non-complements (such as CED and structure-building approaches; cf. Sects. 7.1 and 7.2 below); on the other hand, there are approaches that trace subject and adjunct islands to locality constraints, adopting the idea that the application of movement operations are limited over certain portions of the syntactic structure (such as Subjacency, Barriers, Phases and possibly Relativized Minimality; cf. Sects. 7.3–7.6 below). We will show that our results call for modifications of all existing syntactic theories of island effects, with some of the locality-based theories looking more promising (Subjacency, Barriers, Phases, Relativized Minimality) than theories based on the complement/non-complement distinction (CED, Structure-building).

7.1 Consequences for the Condition on Extraction Domains (CED)

Huang (1982) observed that extraction out of adjunct clauses leads to unacceptability similar to the other known island effects. This new adjunct island effect, coupled

[9] Although we did test wh-islands, which have figured prominently in semantic approaches to island effects such as Szabolcsi and Zwarts (1993) and Abrusán (2011), we did not find any variation in their presence, so our results do not impact debates between syntactic and semantic approaches. We did not test relative clause islands (which have figured prominently in pragmatic approaches to island effects such as Erteschik-Shir 1973 and Goldberg 2006), so our results do not contribute to that discussion. Finally, the conditional adjunct islands that we tested are not the same type of adjunct island in the semantic approach of Truswell (2007).

with the previously observed subject island effect, suggested that the distinction between complements (which are not typically islands) and non-complements (subjects, adjuncts) was relevant for the theory of island effects. This insight led Huang to propose the Condition on Extraction Domains (CED) using the Government and Binding framework. The CED states that extraction out of a phrase is only possible if that phrase is properly governed, where proper government is established under a local relationship with a lexical head (the exact formulation of proper government varies from analysis to analysis). In this way, the CED explains subject and adjunct island effects because non-complements (subjects and adjuncts) are typically not properly governed. Furthermore, the CED can be extended to capture variation in subject and adjunct island effects by allowing for (i) variation in the structural position of subjects, depending on the language and on the thematic and structural properties of the predicate (e.g., post-verbal subjects may be in the correct relationship to the verb to allow for proper government in unaccusative and passive structures, which may account for some examples of successful extraction from subjects across languages; see Haegeman et al. 2014 for a review of the literature on this topic; see also Jurka 2010 and Polinsky et al. 2013 for experimental investigations); and (ii) cross-linguistic variation in the definition of proper government (e.g., one could also capture extraction from subjects by defining proper government such that even pre-verbal subjects are properly governed by the verb; see for example Spyropoulos and Stamatogiannis 2011). Although the notion of government fell out of favor with the rise of the Minimalist Program, the insight of the CED that the complement/non-complement distinction may be relevant for island effects has survived in many Minimalist analyses (see Stepanov 2007 for a more detailed review of the evolution of CED analyses).

Under a CED analysis, the variability that we observed in subject and adjunct islands would suggest that the possibility of proper government of subjects and adjuncts varies both across languages (as already noted by Huang 1982) and also across dependencies. This latter conclusion seems inconsistent with the guiding idea of proper government, as proper government is a local relationship between a governing head (such as V) and a phrase, with no obvious connection to the type of dependency that happens to be moving an item out of the phrase. Furthermore, while variability in whether subjects are proper governed by V seems plausible, the idea that adjuncts could be properly governed by V (as required by English rc-dependencies) also seems inconsistent with the idea of proper government. Taken together, these concerns suggest that our results provide (further) empirical evidence that a government-based approach to island effects may not be the right avenue to pursue.

7.2 Consequences for structure-building approaches

A second type of island theories that hinge on the distinction between complement and non-complement is represented by structure-building approaches. These proposals derive the impossibility of movement out of certain structures from the way the structure is constructed. Perhaps the first example of a structural-building approach is the multiple spell-out theory of Uriagereka (1999) and Nunes and Uriagereka (2000), which derives subject islands and adjunct islands (i.e., CED effects) from the fact that non-complements (such as subjects and adjuncts) must be constructed in a different workspace from the main spine of the sentence, in contrast with complements

(such as objects), which are constructed in the primary workspace. The syntactic objects in the secondary workspaces must be spelled-out prior to being merged with the main structure in order for linearization to be unambiguous, in effect collapsing these secondary workspace structures into something akin to complex words, and rendering them opaque to extraction. Under this approach, languages like Italian that do not demonstrate subject island effects must either (i) not spell-out subjects prior to merging them with the main sentence, or (ii) involve post-verbal subjects that can be correctly linearized with the main spine without spell-out.

The structure-building approach in Johnson (2003) is predicated upon a similar insight that subjects and adjuncts are assembled separately from the main spine of the sentence, but implements that insight with a different architecture. Johnson's architecture consists of a numeration (initially containing all of the lexical items that will comprise the sentence) and a syntactic workspace for constructing trees. Subjects and adjuncts are constructed first (before the main spine of the sentence), and then placed back in the numeration in a process called *renumeration*. The renumeration process triggers irrevocable linearization, preventing future movement out of renumerated items (i.e., creating islands). The renumerated items are then selected from the numeration when it is time to merge them into the main sentence (just like a standard lexical item). Languages, like Italian, that do not demonstrate subject island effects must not renumerate subjects, and therefore must achieve correct linearization in some other way (again, perhaps through constructing subjects in a post-verbal position).

The eclectic approach to subject and adjunct islands proposed by Stepanov (2007) is another example, at least in part, of a structure-building approach to island effects. Stepanov observes that there is cross-linguistic variation in the presence of subject island effects, but no reported variation in the presence of adjunct island effects (except for the specific adjunct type investigated in Truswell 2007). From this he argues that the two island effects should derive from different properties of the grammar, hence the descriptor 'eclectic'. Under the eclectic approach, subject islands are captured by a type of freezing effect (Wexler and Culicover 1981), such that languages with pre-verbal (or moved) subjects should show island effects, and languages with post-verbal (or in-situ) subjects should not (see also Takahashi 1994). In contrast, adjunct islands, which are presumed to be universal, derive from the fact that adjuncts must be constructed in a second workspace and linearized prior to being merged with the main spine of the sentence (similar to Uriagereka 1999 and Nunes and Uriagereka 2000).

Structure-building approaches are strongly universalist: it is the nature of how subjects and adjuncts are constructed that leads to their island status, with no regard to the type of dependency under consideration, and very little room for cross-linguistic grammar differences. Therefore the variability in subject and adjunct islands observed in our results poses a substantial problem for structure-building approaches to island effects, as altering the way in which subjects and adjuncts are constructed based on the type of dependency does not appear to be in line with the spirit of structure-building approaches. For example, the multiple spell-out approach of Uriagereka (1999) and Nunes and Uriagereka (2000) would have to say that English rc-dependencies cause adjuncts to be constructed in the primary syntactic workspace

(while other adjuncts are constructed in a secondary workspace), and that Italian rc-dependencies similarly cause subjects to be constructed in the primary workspace (while other subjects are constructed in a secondary workspace). The renumeration approach of Johnson (2003) would have to say that English rc-dependencies cause adjuncts to no longer require renumeration for successful linearization; and the same goes for subjects and Italian rc-dependencies.

While the eclectic approach of Stepanov (2007) will encounter essentially the same problems as other structure building approaches with respect to the variation of adjunct islands, it could offer a more flexible solution to subject islands variation, assuming that rc-dependencies and wh-dependencies target different subject positions. To account for subject island variation, it might be possible to assume that rc-dependencies can access the subject in-situ even when this appears in a preverbal position (as proposed, for instance, by Chomsky 2008), while interrogative *wh*-dependencies can only target the subject in its surface position (leading to a violation of the freezing effect).

In conclusion, while it may be possible to encode this variability in these systems, it runs contrary to the leading idea of the structure-building approach (i.e., that islands arise due to the way the phrase is constructed), therefore our results seem to suggest that structure-building approaches may not be the most fruitful avenue to pursue.

7.3 Consequences for the Subjacency Condition

Rizzi's (1982) original observations were analyzed using the Subjacency Condition (Chomsky 1973), which in effect prevents the grammatical operation *movement* (thought to be the source of both wh-dependencies and rc-dependencies) from crossing two or more bounding nodes. Prior to Rizzi's observations of Italian island effects, the bounding nodes were assumed to be NP and IP (or DP and TP in more modern terms). Rizzi proposed parameterizing the set of possible bounding nodes such that languages like English, which display both wh-islands and subject islands, have one set of bounding nodes (NP and IP), and that languages like Italian, which do not display either wh-islands or subject islands, have a different set of bounding nodes (NP and CP). This analysis received some initial support from Torrego's (1984) observation that Spanish wh-dependencies show the same pattern of island effects as Italian rc-dependencies: complex NP and adjunct island effects are present, but wh-islands and subject island effects are absent. Crucially, the Subjacency framework was designed to capture wh-islands, subject islands, and complex NP islands, as well as the cross-linguistic variation of those islands, with a single constraint (adjunct islands would not be discussed in detail until Huang 1982).

Because we observed different patterns of island effects for wh-dependencies and rc-dependencies (even within a single language), our results suggest that different bounding nodes should be postulated for each dependency type. One possible analysis would be as follows. First, English wh-dependencies and Italian wh-dependencies could set IP and NP as bounding nodes, with the additional assumption that the specifier of conditional clauses cannot be a landing site for wh-movement (to account for adjunct island effects). Second, English rc-dependencies could also set IP and NP as bounding nodes, but allow rc-movement to land in the specifier of conditional

clauses. Finally, Italian rc-dependencies could set only IP as a bounding node, with the additional assumption that the specifier of conditional *if*-clauses cannot be a landing site for rc-movement, and the assumption that the specifier of a noun-complement CP cannot be a landing site for rc-movement either.

While the above analysis captures the observed facts, a note may be in order about the two landing-site assumptions that are required to make the analysis work. The first assumption, that the specifier of adjunct CPs is not a viable landing site, has always been necessary to accommodate adjunct islands within the Subjacency framework (recall that adjunct islands were first observed after the Subjacency condition was first formulated). What our results add to the picture is the possibility that there is variation across dependency types when it comes to this prohibition. Ideally both this prohibition, and its variability, would be derivable from deeper principles (but given that very few researchers currently work in the Subjacency framework, we do not pursue this question further). The second assumption, that the specifier of CP complements of NP is not a viable landing site, has also always been necessary to accommodate complex NP islands within the original Rizzi (1982) analysis of Italian. This assumption has had no empirical consequence for English, because movement from a complex NPs would cross at least two bounding nodes in English either way. Our results do not impact this assumption at all: Italian requires this assumption under both Rizzi's original analysis and our revised analysis.

7.4 Consequences for the Barriers framework

The barriers framework (Chomsky 1986) has several goals. First, it continues the program begun by the Subjacency condition to capture several, if not all, island effects within a single constraint. Second, it attempts to recast the Subjacency Condition in terms of *barriers to movement* instead of bounding nodes, evoking a parallelism with *barriers to government* in formulations of the Empty Category Principle (e.g., Huang 1982; Lasnik and Saito 1984). Third, it attempts to correct some incorrect predictions of the original Subjacency Condition, such as the prediction that extraction from any complex NP should be unacceptable in English, including NPs in object position (e.g., an "object island"). Finally, it attempts to subsume the CED by incorporating the complement/non-complement distinction into the definition of barrier (specifically, the definition of barrier contains the notion L-marking, the definition of L-marking contains the notion theta-government, and the definition of theta-government is predicated upon theta-marking under sisterhood, which is another way of saying complement). In order to accommodate the complete range of acceptable movement dependencies in English, the barriers framework also introduced the idea that moved elements could adjoin to an XP that is a barrier in order to avoid crossing that barrier (most notably adjunction to VP in acceptable extractions from VP). In this way, the presence or absence of island effects was the result of an interaction between the definition of barrier for that language, and the constraints on adjunction in that language. For example, both subjects and adjuncts are barriers to movement, therefore languages like English which demonstrate both island effects must prohibit adjunction to subjects and adjuncts, and languages like Italian that do not demonstrate subject island effects must allow adjunction to subjects.

In some ways the barriers framework can accommodate our results the most straightforwardly of those reviewed thus far. Because the presence of subject and adjunct islands in the barriers framework is predicated upon disallowing adjunction to subjects and adjuncts, the variation we observed can be captured by simply allowing adjunction to subjects in Italian rc-dependencies (but no other dependencies, including English rc-dependencies), and allowing adjunction to adjuncts in English rc-dependencies (but no other dependencies, including Italian rc-dependencies). As such, there is no need to alter the definition of barrier to accommodate our results. In fact, our results simplify the theory of barriers somewhat: in order to accommodate the lack of wh-islands with Italian rc-dependencies reported by Rizzi (1982), Chomsky (1986) proposes cross-linguistic variation in whether embedded IPs are barriers (presence of wh-islands) or embedded CPs are barriers (absence of wh-islands). Because we observed wh-islands in both types of dependencies in Italian, there is no need to introduce this variation. Of course, other languages may motivate this variation, but we leave that to future research. On a similar note, it would be ideal if both the constraints on adjunction, and the variation in those constraints, could be derived from deeper properties of the grammar, but we leave that to future research as well.

7.5 Consequences for phase-based approaches

Phase-based approaches to island effects build on the idea of cyclicity in grammatical derivations by defining small chunks of structure, called phases, within which certain syntactic operations can (or must) occur. Phases help to reduce the overall computational complexity of syntactic derivations by restricting syntactic operations to local domains (and thus encoding the local nature of syntactic dependencies). Under the original phase theory (Chomsky 2001), phases are defined as vP and CP, and the Phase Impenetrability Condition (PIC) ensures that only the *edge* of the phase (the phase head and its specifier) is available to syntactic operations involving higher phases. This system straightforwardly predicts the existence of wh-islands: the lower wh-word (or phrase) cannot move to the edge of its phase (the embedded CP) because the specifier of the phase is already filled with a wh-word. Therefore the lower wh-word is not visible to higher phases (e.g., matrix CP). While the original phase theory capture wh-islands easily, it cannot explain complex NP, subject, or adjunct islands without additional assumptions, because in each case, it appears as if the specifier positions of each phase are available for movement.

As part of a broader investigation of extraction in Tagalog, Rackowski and Richards (2005) propose a phase-based interpretation of subject and adjunct islands (i.e., CED effects): phrases that enter into an agreement relationship with a phase head (in their analysis, little v) are transparent to movement dependencies, and phrases that do not enter into an agreement relationship with a phase head are opaque to movement dependencies. Rackowski and Richards present evidence that, in Tagalog, CP complements of VP show agreement, whereas CP adjuncts do not, corresponding with their phase-based theory of CED effects.

To account for our results, this approach would have to state that adjuncts in English rc-dependencies agree (albeit without morphological consequence) with little v, while adjuncts in English wh-dependencies do not. Similarly, subjects in Italian

rc-dependencies must agree (again, without morphological consequence) with little v, while subjects in Italian wh-dependencies do not. Agreement, much like structure-building algorithms, is generally assumed to be independent of the nature of the other dependencies in the sentence, therefore the dependency-based variation observed in our results is not the most natural state of affairs for an agreement-centered theory of islands. That being said, one benefit of the agreement-centered theory is that it makes testable predictions: one could look for corroborating evidence of agreement (or lack-of-agreement) that correlates with the presence and absence of islands, or one could look for languages that show overt morphological agreement in one dependency type and not the other.

Müller (2010) proposes a different phase-based approach to CED effects. Under this theory, items can only move to the edge of a phase, and thus escape the PIC, if an edge feature is added to the phase head. Edge features can be freely added to any phase head, but only if the phase head is still active in the derivation. Active phase heads are defined as phase heads that have specifiers waiting to be merged in the future. This means that last-merged specifiers signify the deactivation of phase heads, which in turn prevents the addition of an edge feature to the head, and consequently prevents the movement of items out of the last-merged specifier to the edge of the phase. In this way, last-merged specifiers become islands to movement. This system straightforwardly explains the existence of subject islands, as subjects are last-merged specifiers. Adjunct islands require the additional assumption that adjuncts are the last-merged specifier of a special (covert) functional head.

In order to accommodate our results, Müller's analysis must posit an additional (presumably covert) specifier above the subject in Italian rc-dependencies (but not in Italian wh-dependencies). Similarly, adjuncts in English rc-dependencies must either not be specifiers at all (perhaps they are traditional adjuncts), or must not be the last-merged specifier of the adjunct head (while adjuncts in wh-dependencies must still be last-merged specifiers). These modifications are similar to the modifications necessary in the Rackowski and Richards (2005) theory in that they are not the most natural state of affairs, but they do make testable predictions: there should be detectable consequences of the existence of the (presumably covert) true last-merged specifiers in these constructions.

7.6 Consequences for Relativized Minimality

The Relativized Minimality (RM) framework (Rizzi 1990, 2004) is designed to capture the fact that syntactic dependencies tend to be sensitive to the presence of similar elements. As the name suggests, RM defines local in relative terms: a dependency cannot hold between two items in a sentence if there is a third item that (i) intervenes between the two items in terms of c-command, and (ii) can potentially engage in the same dependency. In this way, only the smallest possible dependencies are licit. Exactly what counts as intervention, and exactly what it means to 'potentially engage in a dependency', are open areas of research in the RM framework, although there is common agreement that intervention is likely based on shared syntactic features.

Among A′-elements, operators naturally constitute a class of interveners, in that they all share a quantificational feature at the interface. The original formulation of

RM was in fact meant to deal mostly with *wh*- and negative islands, where the potential intervener is easily identifiable. While the original version of RM does not apply to islands where the potential interveners are not obvious, such as complex NP, subject, and adjunct islands, this is not necessarily true for more recent developments. Greco (2013) proposes that certain kinds of subjects participate in RM restrictions. Haegeman (2012) proposes that some adverbial clauses normally involve the presence of an operator in their left-periphery. Taking into consideration specifically *if*-clauses, which constitute the empirical case for adjunct islands in our experiments, the relevant representation is provided by adopting Bhatt and Pancheva's (2006) proposal, according to which *if*-clauses are derived through the movement to the left periphery of an operator over possible worlds:

(29) a. [. . .] if John arrives late [. . .]
b. [. . .] [$_{CP}$ OP_w C [John arrives late in w]] [. . .]

According to Haegeman (2012), the presence of such an operator triggers RM effects with respect to some cases of A′-movement *inside* the adjunct clause, explaining some restrictions on main clause phenomena in these environments. Extending this logic to cases of movement that target a position *outside* the adverbial clauses, the representation in (29) might become the basis to account for adjunct islands in terms of RM: operator movement targeting a position outside the adjunct clause cannot take place because of the intervention of the operator over possible worlds which sits at the edge of the adjunct structure.

Of course, in order to account for our results, this proposal must include some way to distinguish between rc- and wh-dependencies. Recent versions of RM can address this difference independently since it has been proposed that intervention is sensitive to the featural composition of the involved elements. The generalization that emerges is that movement of a category which is featurally more complex is not blocked by the intervention of a featurally simpler category (Starke 2001; Rizzi 2004). In the same spirit, Abels (2012) observes that rc-dependencies also appear to be less constrained than wh-dependencies in other configurations involving locality restrictions and proposes to trace this difference back to the richer featural composition of relative pronouns compared to wh-elements (see also Haegeman 2012; Haegeman and Ürögdi 2010a, 2010b). This logic might also apply to the asymmetry between rc and *wh*-dependencies that we observed in English adjunct islands. According to this logic, an intervention effect in rc-dependencies can be avoided if the relative operator (or whatever category moves in rc-dependencies) has a richer featural composition than the intervening category, namely the operator over possible worlds at the left edge of the *if*-clause.

We discussed the role of featural composition for island effects in Sect. 6.1. Complex wh-phrases do not seem to ameliorate adjunct island effects in English. So, in order to extend the RM approach to adjunct islands, relative operators should bear a featural specification that (i) is not overtly manifested by the distinction between complex wh-phrases and bare wh-words and (ii) specifically interacts with the operator over possible worlds at the left edge of the *if*-clauses. This would require further enriching (complicating?) the RM taxonomy to distinguish between different subclasses of mutual interveners (see in particular the discussion in Abels 2012).

Table 3 Summary of the qualitative results of the 12 experiments. The presence of island effects is indicated with a plus sign (+), the absence of island effects is indicated with a minus sign

Language	Dependency	wh/*whether*	Complex NP	Subject	Adjunct
English	wh (bare)	+	+	+	+
	wh (complex)	+	+	+	+
	rc	+	+	+	−
Italian	wh	+	+	+	+
	rc	+	+	−	+

Moreover, the featural composition of either the moving category or the intervener must be parameterized, since Italian does not display the same asymmetry between rc- and wh-dependencies in adjunct islands. Therefore, either the featural composition of the operator on the edge of the adjunct clause or the featural composition of the relative operator must be made variable across different languages.

7.7 Summary

The upshot of the new variation observed in our experiments is that some syntactic theories of island effects can more naturally accommodate the new facts (Subjacency, Barriers, and Phase-based approaches being the most natural, and CED and Structure-building approaches being the least), and some even make new predictions that should be testable (especially the Phase-based approaches). In all cases, the modifications required by this new pattern of variation call for deeper explanations (e.g., the variability in landing sites under Subjacency, the variability in adjunction under Barriers, and the variability in agreement or specifiers under Phases). Although we leave these deeper questions to future research, we would like to note that these questions serve as an interesting example of how cross-linguistic experimental work can lead to a series of new research directions in theoretical syntax.

8 Conclusion

The results of the 12 experiments presented here, summarized in Table 3, strongly suggest variation in island effects both across languages, and across dependency types. This variation requires modifications to all existing syntactic theories of island effects, with some theories accommodating the facts fairly well (Subjacency, Barriers, Phases), and some theories accommodating the facts less well (CED, Structure-building). The results also suggest that the variation truly is due to the nature of the dependencies, and not the featural specification of the heads of the dependencies. Furthermore, the results suggest that complex wh-phrases do not entirely ameliorate any of the island effects, but instead raise the acceptability of island-violating sentences of wh-island and complex NP islands.

Taken as a whole, these results suggest that cross-linguistic experimental work has the potential to both reveal previously unobserved effects (e.g., the lack of adjunct islands with rc-dependencies in English), as well as better isolate previously

observed effects (e.g., the D-linking effect on certain island-violating sentences). As such, we believe that formal experimental work deserves a prominent place in the cross-linguistic syntactician's toolkit.

Acknowledgements This work was supported in part by National Science Foundation grants BCS-0843896 and BCS-1347115 to JS. CG was supported by the Fonds Wetenschappelijk Onderzoek (FWO)—FWOproject-2009-Ofysseus-Heageman-G091409 during the final revision of the manuscript. We would like to thank Michela Marchesi for assistance collecting data for the Italian wh-dependencies experiment. We would like to thank Jeremy Hartman, Norbert Hornstein, Luigi Rizzi, and two anonymous reviewers for helpful comments and suggestions on earlier versions of this article. We would also like to thank audiences at Harvard University, Johns Hopkins University, and the University of Illinois Chicago for helpful comments at various stages of the development of this study. All errors remain our own.

References

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

Abels, Klaus. 2012. The Italian left periphery: a view from locality. *Linguistic Inquiry* 43: 229–254.

Bhatt, Rajesh. 2002. The raising analysis of relative clauses: evidence from adjectival modification. *Natural Language Semantics* 10: 43–90.

Bhatt, Rajesh, and Roumyana Pancheva. 2006. Conditionals. In *The Blackwell companion to syntax, Vol. 1*, eds. Martin Everaert and Henk van Riemsdijk, 638–687. Oxford: Blackwell.

Bianchi, Valentina. 1999. *Consequences of antisymmetry: headed relative clauses*. Berlin: de Gruyter.

Browning, Marguerite. 1987. Null operator constructions. Ph.D. thesis, MIT.

Cecchetto, Carlo, and Caterina Donati. 2015. *(Re)labeling*. Cambridge: MIT Press. ISBN 9780262028721.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge: MIT Press.

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

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

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

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

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

Donati, Caterina, and Carlo Cecchetto. 2011. Relabeling heads. A unified account for relativization structures. *Linguistic Inquiry* 42: 519–560.

Erteschik-Shir, Nomi. 1973. On the nature of island constraints. Ph.D. thesis, MIT.

Featherston, Sam. 2005. Magnitude estimation and what it can do for your syntax: some wh-constraints in German. *Lingua* 115: 1525–1550.

Ferreira, Fernanda. 2005. Psycholinguistics, formal grammars, and cognitive science. *The Linguistic Review* 22: 365–380.

Friedmann, Naama, Adriana Belletti, and Luigi Rizzi. 2009. Relativized relatives: types of intervention in the acquisition of A-bar dependencies. *Lingua* 119: 67–88.

Gibson, Edward, and Evelina Fedorenko. 2010. Weak quantitative standards in linguistics research. *Trends in Cognitive Sciences* 14: 233–234.

Goldberg, Adele. 2006. *Constructions at work*. London: Oxford University Press.

Greco, Ciro. 2013. Subjects and arguments in Ā-syntax. Ph.D. thesis, University of Milan-Bicocca.

Haegeman, Liliane. 2012. *Adverbial clauses, main clause phenomena, and the composition of the left periphery: the cartography of syntactic structures*, Vol. 8. London: Oxford University Press.

Haegeman, Liliane, and Barbara Ürögdi. 2010a. Referential CPs and DPs: an operator movement account. *Theoretical Linguistics* 36: 111–152.

Haegeman, Liliane, and Barbara Ürögdi. 2010b. Operator movement, referentiality and intervention. *Theoretical Linguistics* 36: 233–246.

Haegeman, Liliane, Ángel L. Jiménez-Fernández, and Andrew Radford. 2014. Deconstructing the Subject Condition in terms of cumulative constraint violation. *The Linguistic Review* 31: 73–150.

Hofmeister, Philip, and Ivan A. Sag. 2010. Cognitive constraints and island effects. *Language* 86: 366–415.

Huang, C.-T. James. 1982. Logical relations in Chinese and the theory of grammar. Ph.D. thesis, MIT.

Johnson, Kyle. 2003. Towards an etiology of Adjunct islands. In *19th Scandinavian Conference of Linguistics*, Vol. 31 of *Nordlyd Tromso University working papers on language and linguistics*.

Jurka, Johannes. 2010. The importance of being a complement: CED effects revisited. Ph.D. thesis, University of Maryland.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge: MIT Press.

Keller, Frank. 2000. Gradience in grammar: experimental and computational aspects of degrees of grammaticality. Ph.D. thesis, University of Edinburgh.

Kluender, Robert, and Marta Kutas. 1993. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8: 573–633.

Lasnik, Howard, and Mamoru Saito. 1984. On the nature of proper government. *Linguistic Inquiry* 15: 235–289.

Maxwell, Scott E., and Harold D. Delaney. 2003. *Designing experiments and analyzing data: a model comparison perspective*. Mahwah: Lawrence Erlbaum Associates.

Moro, Andrea. 2011. Clause structure folding and the "wh-in situ effect". *Linguistic Inquiry* 42: 389–411.

Müller, Gereon. 2010. On deriving CEG effects from the PIC. *Linguistic Inquiry* 41: 35–82.

Nunes, Jairo, and Juan Uriagereka. 2000. Cyclicity and extraction domains. *Syntax* 3: 20–43.

Polinsky, Maria, Carlos Gómez Gallo, Peter Graff, Ekaterina Kravtchenko, Adam Milton Morgan, and Anne Sturgeon. 2013. Subject islands are different. In *Experimental syntax and island effects*, ed. Jon Sprouse, 286–309. Cambridge: Cambridge University Press.

Rackowski, Andrea, and Norvin Richards. 2005. Phase edge and extraction: a Tagalog case study. *Linguistic Inquiry* 36: 565–599.

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

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

Rizzi, Luigi. 2004. Locality and left periphery. In *The cartography of syntactic structures*, ed. Adriana Belletti. Vol. 3 of *Structures and beyond*, 223–251. New York: Oxford University Press.

Ross, John Robert. 1967. Constraints on variables in syntax. Ph.D. thesis, MIT.

Schütze, Carson T. 1996. *The empirical base of linguistics: grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.

Sprouse, Jon. 2007. A program for experimental syntax. Ph.D. thesis, University of Maryland.

Sprouse, Jon. 2011. A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods* 43: 155–167.

Sprouse, Jon, and Diogo Almeida. 2012. Assessing the reliability of textbook data in syntax: Adger's Core Syntax. *Journal of Linguistics* 48: 609–642.

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

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

Sprouse, Jon, Carson Schütze, and Diogo Almeida. 2013. A comparison of informal and formal acceptability judgments using a random sample from Linguistic Inquiry 2001–2010. *Lingua* 134: 219–248.

Spyropoulos, Vassilios, and Nikolaos Stamatogiannis. 2011. Subextraction from subjects in Greek. Paper presented at the conference on Islands in Contemporary Linguistic Theory, University of the Basque Country Vitoria-Gasteiz, 16–18 November 2011.

Starke, Michal. 2001. Move dissolves into Merge: a theory of locality. Ph.D. thesis, University of Geneva.

Stepanov, Arthur. 2007. The end of CED? Minimalism and extraction domains. *Syntax* 10: 80–126.

Szabolcsi, Anna. 2006. Strong vs. weak islands. In *The Blackwellcompanion to syntax, Vol. 4*, eds. Martin Everaet, Henk van Riemsdijk, Rob Goedemans, and Bart Hollebrandse, 479–531. New York/Oxford: Blackwell.

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

Takahashi, Daiko. 1994. Minimality of movement. Ph.D. thesis, University of Connecticut, Storrs, CT.

Torrego, Ester. 1984. On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15: 103–129.

Truswell, Robert. 2007. Extraction from adjuncts and the structure of events. *Lingua* 117: 1355–1377.

Uriagereka, Juan. 1999. Multiple spell-out. In *Working minimalism*, eds. Samuel David Epstein and Norbert Hornstein, 251–282. Cambridge: MIT Press.

Vergnaud Jean Roger. 1974. French relative clauses. Ph.D. thesis, MIT.
Wasow, Thomas, and Jennifer Arnold. 2005. Intuitions in linguistic argumentation. *Lingua* 115: 1481–1496.
Wexler, Kenneth, and Peter Culicover. 1981. *Formal principles of language acquisition*. Cambridge: MIT Press.
Wike, E. L., and J. D. Church. 1976. Comments on Clark's "The language-as-fixed-effect fallacy". *Journal of Verbal Learning and Verbal Behavior* 15: 249–255.