1 Experimental syntax and island effects: Toward a comprehensive theory of islands

Jon Sprouse and Norbert Hornstein

1 Introduction

One of the most pervasive properties of human language is the existence of *dependencies*: necessary relationships that hold between two elements in a sentence. The primary objects of study in this volume are *long-distance "filler-gap" dependencies* – a special subset of dependencies that are not constrained by standard measures of length such as number of words or number of clauses. For example, wh-questions in English (1) contain a long-distance dependency between the wh-word at the beginning of the sentence and a theta-role assigning element (such as a verb or preposition) later in the sentence that selects the wh-word as a semantic argument. We will adopt a relatively theory-neutral terminology and call the end of this dependency the *gap position*, indicated by an underscore in examples. The pattern in (1a–1c) suggests that long-distance dependency between wh-words and gap positions in English can be separated by any number of embedded sentences:[1]

(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 __?

Although wh-dependencies tend to be used as the canonical example of long-distance dependencies in the linguistic literature, there are many different constructions in the world's languages that contain long-distance dependencies. For example, other English long-distance dependencies include relativization (2a), topicalization (2b), adjective-though constructions (2c), and various types of clefts (2d), among others:

[1] Clearly there is an upper limit to the length of sentences that a human speaker can understand. However, this is true of sentences that contain long-distance dependencies and sentences that do not, suggesting that the eventual parsing failure is not due to the presence of long-distance dependencies. Instead, it is likely a consequence of the limited memory resources available for tracking the entities and relationships described by the sentence.

(2) a. I like the car that John bought __.
b. I like most of these cars, but that car, I love __.
c. Fast though the sports car is __, I prefer the hybrid.
d. This is the car that John bought __.

Although long-distance dependencies are not constrained by standard measures of length, this does not mean that all long-distance dependencies are acceptable. Crucially, when the gap position of a long-distance dependency is inside certain structures, the sentence becomes unacceptable:[2]

(3) a. *What do you wonder [$_{CP}$ whether John bought __]?
b. *What did you make [$_{NP}$ the claim that John bought __]?
c. *What do you think [$_{NP}$ the speech about __] interrupted the TV show?
d. *What do you worry [$_{CP}$ if John buys __]?
e. *What did you meet [$_{RC}$ the scientist who invented __]?
f. *What did [$_{CP}$ that John wrote __] offend the editor?
g. *What did John buy [$_{ConjP}$ a shirt and __]?
h. *Which did John borrow [$_{NP}$ __ book]?

Ross (1967) used the metaphorical term *island* to refer to these "gap-resistant" structures, evoking the idea that the wh-word could not *move* from the gap-position inside the island to the front of the sentence.[3] Building on this, we will use the term *island effect* to refer to the unacceptability that arises when a gap position occurs within an island.[4] It is also common in the literature to refer to island effects based on the structure that creates them: WH-islands (3a), Complex Noun Phrase islands (3b), Subject islands (3c), Adjunct islands (3d), Relative Clause islands (3e), and Sentential Subject islands (3f), although some island types are more commonly referred to based on the proposed constraint that they violate, as in Coordinate Structure Constraint violations (3g) and Left Branch Extraction violations (3h).

[2] Some terminology: "(un)acceptability" is used when describing speaker judgments. These are often referred to in the literature as "grammaticality judgements." However, this is misleading for at least two reasons. First, grammaticality is at most one factor determining acceptability. Second, acceptability is a descriptive predicate that describes the observable data, while grammaticality is a predicate that describes the mental representations that linguists theoretically invoke to (at least partially) explain this data. Speakers have privileged access to their acceptability judgments. Nobody's grammaticality judgments are dispositive. See below for further discussion.

[3] While it is true that this was originally a theory-laden metaphor (invoking the idea of movement that is central to transformational grammar), the term itself has been adopted by nearly all linguistic theories, therefore we will continue to use it here. A historical note: Ross (1967) attributed island effects to the illicit application of "chopping" rules within islands. Movement from islands was permitted. The prohibition against movement from islands is proposed in later accounts that built on Ross's earlier work, most especially Chomsky's Subjacency Theory (1973, 1981, 1986).

[4] We have chosen the term *island effect* over the more common *island violation* because the former is agnostic about the source of the unacceptability (the primary question driving this volume), while the latter specifically refers to the violation of a specific (likely grammatical) constraint.

As an acceptability-based phenomenon, the source of island effects has long been a topic of debate within the linguistic and psycholinguistic literature. The problem lies in the fact that acceptability judgments are a behavioral response that is the result of successful sentence processing (Chomsky 1965, Schütze 1996, Sprouse and Almeida 2013), and as such could be influenced by any of the cognitive systems that are implicated in successful sentence processing, from the multiple mental representations that can be used to characterize a sentence (e.g., phonological, morphological, syntactic, semantic, pragmatic), to the different components of the parsing system that must be deployed during normal sentence comprehension (e.g., structure-building operations, ambiguity resolution heuristics, working memory systems). In short, this is the classic problem of cognitive science (mapping observable behavioral responses to unobservable cognitive constructs), exacerbated by the complexity and multi-level nature of human language. The primary empirical goal of this volume is to bring the techniques of experimental syntax (broadly construed) to bear on this particular instantiation of the cognitive science problem, and move the field one step closer to identifying the source of island effects.

2 Components of a comprehensive theory of island effects

Even from the brief introduction to island effects presented above, it should be clear that identifying the source of island effects requires much more than a simple catalog of the constructions that demonstrate them. Experimental syntax provides a set of tools that goes beyond the traditional acceptability judgment experiments that have been used (to good success) in the existing literature. Our hope is that these tools may reveal new types of data that bear on this question. The logical place to begin the search for new evidence is by reviewing the complex patterns of island effects, both across languages and across construction types, that have been previously reported in the syntax literature. These patterns are useful to our goals in at least two ways. First, they provide a starting point for thinking about the types of evidence that could be used to isolate the source of island effects. It is our hope that the chapters in this volume will use some of these complex patterns to tease apart the role of different levels of linguistic representation and processing in explaining the unacceptability of island effects. Second, these patterns present a list of phenomena that any comprehensive theory of island effects must explain. It is not enough for a theory of island effects to simply explain the unacceptability of island effects in one language or one construction; a comprehensive theory must also explain the complex pattern that is observed across languages and across constructions.

Table 1.1 *Cross-linguistic variation in island effects*

	WH	Complex NP	Subject	Adjunct	Relative Clause
English	*	*	*	*	*
Italian	–	*	?	*	*
Spanish	–	*	?	*	*
Portuguese	–	*	?	*	*
French	?	*	*	*	*
German	*	*	?	*	*
Russian	*	*	?	*	*
Scandinavian	–	–	–	–	?
Hungarian	?	*	?	*	*

2.1 *Variation in languages with overt long-distance dependencies*

Perhaps the most obvious fact that a comprehensive theory must explain is the cross-linguistic variability of island effects. As illustrated in (3) above, English demonstrates at least eight different types of island effects; however, several languages demonstrate fewer. Table 1.1 presents nine languages that are known to employ wh-movement in questions, and five of the most studied island effects: WH-islands (3a), Complex Noun Phrase islands (3b), Subject islands (3c), Adjunct islands (3d), and Relative Clause islands (3e). The diacritics indicate whether the specified language demonstrates that particular island effect: asterisks indicate that the island effect arises in that language, dashes indicate that the island effect does not arise in that language, and question marks indicate that the island effect arises for some sentence types, but not others. We should note that Table 1.1 idealizes the empirical results to a considerable extent. There has been considerable work on these cross-linguistic differences and the differences noted here are not nearly as categorical as displayed. For example, many English speakers treat the wh-island violations discussed in Rizzi 1982b as acceptable (c.f. Grimshaw 1986). Furthermore, it has long been noted that the degrees of unacceptability substantially differ across the various islands. For example, violations of the WH-island condition are generally less unacceptable than violations of the relative clause version of the Complex Noun Phrase Constraint. This said, the table offers a good approximation of the received wisdom, though the information it encodes should be treated as a potential object of study. For discussion of these matters see chapter 3 (Hofmeister *et al.*), chapter 4 (Phillips), chapter 11 (Kush *et al.*), chapter 12 (Jurka), chapter 13 (Polinsky *et al.*).

To the extent that Table 1.1 is accurate, the cross-linguistic variation it reports raises some very interesting questions for theories of island effects. For

example, it has proven relatively difficult to characterize precisely the variability indicated by question marks; that is, island effects in certain languages that appear arise for some sentences, but not others. Furthermore, the mere existence of variability has proven challenging for approaches to island effects that postulate a source that is outside of the grammar (e.g., components of the sentence processing system), as grammatical theories have traditionally been the sole locus of cross-linguistic variation.

Although Table 1.1 suggests that there is a good amount of variation cross-linguistically, at least two patterns are suggested by this (limited) sample of languages. The first is the relatively robust correlation between a lack of WH-island effects and at least a partial lack of Subject island effects, as seen in several Romance languages. The facts underlying this correlation led Rizzi (1982b) to propose what is now one of the most famous theories of island variation (see also Torrego 1984 for an extension to Spanish). Rizzi's proposal was predicated upon Chomsky's (1973) analysis of island effects, which postulated a constraint on the wh-movement operation that Chomsky called the Subjacency Condition. The Subjacency Condition held that wh-movement must target landing sites (primarily spec, CPs) that are *subjacent* to the gap position, where *subjacent* was defined as crossing fewer than two *bounding nodes*, and bounding nodes were defined as any NP or IP that dominated the gap position. In effect, the Subjacency constraint meant that a single instance of wh-movement only operated within a single CP: a wh-word could move from its gap position within the clause to the specifier of the local CP, but not farther, because this one-clause movement would cross one bounding node (the IP), but two-clause movement would necessarily cross at least two bounding nodes (the embedded IP and the matrix IP).

Although this proposal that wh-movement is bounded amounted to a radical shift from Ross's (1967) analysis, in which wh-movement was an unbounded operation, it crucially provided a mechanism for capturing island effects. Although we won't demonstrate the analyses here, the Subjacency analysis captured WH-islands by making the first spec, CP position unavailable as a landing site (because it is filled with a wh-word), thus forcing wh-movement to violate Subjacency by crossing two bounding nodes (the IP of the embedded clause and the IP of the matrix clause). The Subjacency analysis also captured Subject islands (which were discovered later by Huang 1982a), as movement out of the subject NP required wh-movement to cross two bounding nodes (the subject NP and the IP).[5]

[5] The original Subjacency analysis was not without problems, even for English. It could not account for Complex NP, Relative Clause, and Adjunct islands without additional assumptions. Furthermore, it wrongly predicted that movement out of NPs in object position should be unacceptable (i.e., an Object island to parallel Subject islands). Chomsky (1986) attempted to

Rizzi (1982b) argued that the success of the Subjacency analysis at capturing WH-islands (and later Subject islands) was a problem for languages like Italian, as Italian allows long-distance dependencies (in this case, relative clause dependencies) to have a gap position inside of embedded questions, which is the canonical WH-island configuration:

(4) Absence of WH-island effects in Italian
   a. Il solo incarico che non sapevi a chi avrebbero affidato è poi finito proprio a te.
      the only task that not knew.2SG to whom have.3PL.COND assigned is then ended_up right to you
      'The only task that you didn't know who they would assign_to was then given right to you.'
   b. Tuo fratello, a cui mi domando che storie abbiano raccontato, era molto preoccupato.
      your brother, to whom myself ask.1SG what stories have.3PL told, was very worried
      'Your brother, who I wonder what stories they told __, was very worried.'

Crucially, Rizzi did observe other island types in Italian, such as Complex NP islands:

(5) Presence of NP island effects in Italian
   a. *Questo incarico, che non sapevo la novità che avrebbero affidato a te, . . .
      this task, that not knew.1SG the news that have.3PL.COND assigned to you
      'This task, which I didn't know the news that they would assign __ to you, . . . '
   b. *Tuo fratello, a cui temo la possibilità che abbiano raccontato tutto, . . .
      your brother, to whom fear.1SG the possibility that have.3PL told everything
      'Your brother, who I am afraid of the possibility that they have told __ everything . . . '

From these facts Rizzi argued that Italian relative clause formation must involve successive cyclic movement and the Subjacency Condition; however, Rizzi also argued that the bounding nodes for Italian could not be NP and IP, as that would lead to WH-islands. Instead, Rizzi proposed that the choice of bounding nodes be a parameter that could vary across languages. In order to capture the presence of Complex NP islands and the absence of WH-islands, Rizzi proposed that the bounding nodes in Italian should be NP and CP. Torrego (1984) later confirmed that this parametric theory of Subjacency would correctly account for the island facts in Spanish, as well as the apparent correlation between WH-island effects

correct these problems, as well as unify the definition of *bounding node* from the Subjacency Condition and *barrier* from the Empty Category Principle. Although this attempt is now generally considered a failure, it remains a classic example of two of the primary goals of high-level syntactic theorizing: correcting empirical inadequacies of previous analyses while reducing the number of objects in the ontology of the theory.

(which are predicated upon IP bounding nodes) and Subject island effects (which are also predicated upon IP bounding nodes) that is suggested by the Italian and Spanish facts (see also Sportiche 1981 for French).

Whereas the first pattern in Table 1.1 primarily concerns Romance languages, the second pattern concerns Scandinavian languages (Swedish, Norwegian, Danish, and Icelandic). As first observed by Engdahl (1980) for Swedish, Scandinavian languages do not demonstrate any of these five island effects.[6] Engdahl (1980) argued this leaves only two options. Option one is to allow unboundedness in the parametric theory, such that wh-movement in Scandinavian can be unbounded. The second option is to allow Scandinavian CPs to have multiple specifier positions. The discovery of Subject islands in the intervening years means that this second option must be augmented by setting the bounding nodes in Scandinavian to NP and CP (because Subject islands are caused by NP and IP being bounding nodes, regardless of the number of spec, CP positions available). In either case, the existence of apparently island-less languages such as modern Scandinavian languages raises interesting challenges for any comprehensive theory of island effects.

All Subjacency-like accounts of islands have four parts: (i) a proposal for measuring the size of a given movement step, (ii) a principle that limits the step size of any given movement, (iii) a specification of which nodes permit escape ("escape hatches"), and (iv) a specification of how many slots an escape hatch has. The variation noted in Table 1.1 is accounted for by parameterizing one (or more) of these four basic features. Much theoretical work has concentrated on trying to rationalize these four features (e.g., why some nodes count for measuring step size and some do not) and exploring the analytic options of varying one or another assumption. Perhaps the most interesting consequence of Subjacency-like accounts of islands is how they tied together island effects with the requirement that all long-distance movement be successive cyclic. In other words, Subjacency implies successive cyclicity. In fact, early work by Kayne and Pollock (1978) and Torrego (1984) attempted to provide empirical evidence for successive cyclic movement.

2.2 *Wh-in-situ and the argument/adjunct distinction*

A comprehensive theory of island effects must not only account for the pattern of variation observed in languages with overt long-distance dependencies (e.g., wh-movement), but also for the pattern of island effects that have been observed

[6] Swedish is not bereft of apparent island effects. Rather it does not display island effects in all contexts where they are theoretically expected to appear (and as they do appear in English). For example, there are some unacceptable instances of extracting out of complex noun phrases, but others seem perfectly fine. For some discussion see chapter 11 (Kush *et al.*).

in languages without overt long-distance dependencies. The primary example of this is the argument/adjunct distinction that has been observed in so-called wh-in-situ languages such as Chinese (Huang 1982a), Japanese (Lasnik and Saito 1984), and Sinhala (Hagstrom 1998). In wh-in-situ languages, question formation does not involve displacement of the wh-word: the wh-word appears in the same position that the questioned constituent would appear in a declarative sentence (i.e., the gap-position in wh-movement languages). Given that there is no overt long-distance dependency in wh-in-situ languages, one might expect that there are also no island effects in wh-in-situ languages. At least for wh-arguments such as *who* and *what*, this prediction appears to hold (illustrated here for an Adjunct island in Japanese):

(6) [John-wa [$_{ADJ}$ kare-no okusan-ga nani-o katta kara] okoru-to no]?
J-top he-gen wife-nom what-acc bought because get.angry-Q?
'What would John get angry because his wife bought t_{what}?'

However, wh-adjuncts such as *why* and *how* cannot appear inside of island structures in wh-in-situ languages:

(7) *[John-wa [$_{ADJ}$ kare-no okusan-ga naze atarasii doresu-o katta kara]
J-top he-gen wife-nom why new dress-acc bought because
okoru-to no]?
get.angry Q?
'Why would John get angry because his wife bought a new dress t_{why}?'

This pattern suggests that wh-in-situ languages do indeed display island effects, but that the underlying source of island effects in wh-in-situ languages is conditioned by the argument/adjunct distinction.

Much like the variation observed with wh-movement languages, the most famous analysis of the argument/adjunct distinction in wh-in-situ languages is predicated upon the Subjacency analysis of Chomsky (1973). Huang (1982a) proposed that wh-words in wh-in-situ languages do in fact undergo a type of wh-movement, but instead of being *overt* as is the case with wh-movement languages, the wh-movement in wh-in-situ languages is *covert*. Huang further proposed restricting the Subjacency condition to *overt* wh-movement (as a condition on S-Structure).[7] In this way, the covert wh-movement in wh-in-situ languages would be constrained by Subjacency, correctly accounting for the absence of island effects with wh-arguments in wh-in-situ languages. In order to account for the presence of island effects with wh-adjuncts in wh-in-situ languages, Huang proposed that the Empty Category Principle, which stated that every empty category (such as the trace hypothesized to exist in the gap position

[7] This relies on the definition of movement as an operation holding between an antecedent and a phonetically null trace; see Chomsky (1977).

of wh-movement) must be *properly governed*, holds for traces of both overt and covert wh-movement (as a condition on LF). Huang then defined two ways in which a trace can be properly governed: (i) wh-argument traces are properly governed by virtue of being theta-marked (i.e., assigned a thematic role) by a predicate (called *head government*), and (ii) wh-adjunct traces, which are not theta-marked, can be properly governed by residing in the correct configurational relationship with the displaced wh-adjunct (called *antecedent government*). By defining the configurational relationship of antecedent government properly, the island effects observed with in-situ wh-adjuncts could be captured as a violation of the Empty Category Principle.

The ECP approach to the pattern of island effects in wh-in-situ languages, though empirically adequate, required two assumptions that could be viewed as introducing redundancy into the theory. First, the role of the Subjacency Condition was conditioned upon the type of wh-movement (overt or covert), suggesting that there were in fact two different movement operations that applied to the same types of elements (wh-words), only at different points in the syntactic derivation. Second, the structural definitions for antecedent government under the ECP were identical to the structural definitions of island effects under the Subjacency Condition, suggesting that there were two syntactic constraints (Subjacency and the ECP) with the same content. These redundancies led several syntacticians to propose alternative analyses of wh-in-situ island effects. For example, Nishigauchi (1990) proposed an analysis in which Subjacency constrains both overt and covert movement, thus capturing the existence of island effects for in-situ wh-adjuncts in the same way as wh-movement in English. In order to account for the absence of island effects for in-situ wh-arguments, Nishigauchi proposed a covert pied-piping operation in which the entire island structure moved to spec, CP rather than just the wh-word. Because the wh-word never crossed the island boundary, Subjacency was not violated. Tsai (1994) and Hagstrom (1998) took a different tack altogether, rejecting the idea that in-situ wh-words move covertly, and instead focusing on the relationship between the question particle that appears in C in languages such as Japanese and Sinhala and the in-situ wh-word. Although no consensus was ever reached about the correct analysis of the argument/adjunct distinction in wh-in-situ languages, it is clear from these proposals that accounting for these facts is no easy task, and should be a high priority for any comprehensive theory of island effects.

2.3 *Resumptive pronouns*

Although most of the languages discussed so far exclusively employ gap positions as the foot of long-distance dependencies, about half of the world's languages appear to allow a second option: resumptive pronouns. Resumptive

pronouns are lexically indistinguishable from regular pronouns, but appear in the position that under other circumstances would be the gap position of a long-distance dependency (McCloskey 2006). When it comes to the interaction of resumptive pronouns and island effects, McCloskey (2006) identifies three types of languages.

Type 1: Free-variation languages

In the first type of language, exemplified here by Irish (McCloskey 1990, 2006), resumptive pronouns are essentially in free variation with gaps, as long as the gap/pronoun appears outside of an island structure:

(8) a. an ghirseach a ghoid na síogaí __
the girl C stole the fairies
'the girl who the fairies stole'
b. an ghirseach ar ghoid na síogaí í
the girl C stole the fairies her
'the girl who the fairies stole'

Inside of island structures, gaps and resumptive pronouns are in complementary distribution: gaps cannot appear inside of island structures, but resumptive pronouns can:

(9) a. teach nach n-aithneochthá cá rabh sé
house neg recognize where was it
'A house that you wouldn't recognize where it was'
b. *teach nach n-aithneochthá cá rabh __
house neg recognize where was __
'A house that you wouldn't recognize where __ was'

The relevance of resumptive pronouns in free-variation languages for the theory of island effects rests in their dual nature (which McCloskey (2006) describes as Janus-like): whereas true gaps are canonically associated with long-distance dependencies that are sensitive to island effects, and true pronouns are canonically associated with a type of long-distance dependency that is not sensitive to island effects (i.e., binding relations), resumptive pronouns fall in between by allowing non-binding long-distance dependencies to cross island structures.[8]

Type 2: Restricted distribution languages

In the second type of language, exemplified here by Vata (a Kru language of North Africa: Koopman 1984, Koopman and Sportiche 1986), resumptive pronouns and gaps do not freely vary outside of island structures. For example,

[8] It should be noted that the relevance of resumptive pronouns to island effects was first observed by Ross (1967). He noted that resumptive pronouns obviate island effects when present. As he also assumed that they are related to their antecedents via movement, he concluded that movement per se could not be island-sensitive. The approach discussed by McCloskey (2006) inverts this logic: binding is different from movement and the latter is island-sensitive while the former is not. Importantly both point to a conclusion of current interest: that overt gaps make a difference even if the dependency looks similar. The problem is not the dependency but how it is formed.