The Real-Time Status of Island Phenomena[*]

Colin Phillips
University of Maryland

Abstract

In parasitic gap constructions an illicit gap inside a syntactic island becomes acceptable in combination with an additional licit gap, a result that has interesting implications for theories of grammar. Such constructions hold even greater interest for the question of the relation between grammatical knowledge and real-time language processing. This article presents results from two experiments on parasitic gap constructions in English in which the parasitic gap appears inside a subject island, before the licensing gap. An off-line study confirms that parasitic gaps are acceptable when they occur inside the infinival complement of a subject NP, but not when they occur inside a finite relative clause. An on-line self-paced reading study uses a plausibility manipulation technique to show that incremental positing of gaps inside islands occurs in just those environments where parasitic gaps are acceptable. The fact that parasitic gaps are constructed incrementally in language processing presents a challenge for attempts to explain subject islands as epiphenomena of constraints on language processing, and also helps to resolve apparent conflicts in previous studies of the role of island constraints in parsing.

1. INTRODUCTION. This article is concerned with the relation between grammatical knowledge and the processes involved in real-time language processing. Grammatical theories have typically aimed to account for patterns of acceptability judgments, while remaining relatively agnostic about the real-time processes involved in speaking, understanding and making acceptability judgments. Nevertheless, the role of grammatical constraints in real-time language processes is important from a number of different perspectives. For theories of language processing, it is important to know the extent to which real-time mechanisms reflect the details of grammar. For theories of well-formedness, it is important to know whether some constraints derive from limitations on language processing. And for theories that aim to minimize the distinction between grammatical knowledge and real-time processes, it is important to validate the prediction that real-time processes are grammatically accurate. This article presents evidence on the processing of island constraints on long-distance dependencies that indicates a high degree of grammatical sophistication in real-time language processing, and presents a challenge for the view that island constraints ultimately reflect constraints on processing.

The syntactic dependencies formed by topicalization, relativization, or wh-question formation are potentially unbounded in length, but are also subject to a number of restrictions. For example, a direct object wh-phrase may form a dependency with the verb that assigns its

thematic role across arbitrarily many clauses (1), but such dependencies are impossible when certain types of material intervene between the wh-phrase and the verb. Long-distance dependencies are blocked by relative clauses (2a) and other types of complex NP (2b), interrogative clauses (2c), factive clauses (2d), subject and adjunct clauses (2e-f), among others. Since the seminal work of Ross (1967) these domains have collectively been known as *islands*, and the constraints on long-distance dependency formation are known as *island constraints*. For the sake of expository convenience, the canonical position of the fronted wh-phrase is indicated by a gap in all examples that follow. However, all of the discussion that follows is equally compatible with theories in which a fronted phrase is linked to an underlying argument position as with theories in which the fronted phrase is directly associated with the subcategorizing verb.

(1)   Who did you hope that the candidate said that he admired ___?

(2) a. *Who did the candidate read a book that praised ___?
    b.* Who did the candidate read The Times' article about ___?
    c. *Who did the candidate wonder whether the press would denounce ___?
    d.* Why did you remember that the senator supported the bill ___?
    e. *Who did the fact that the candidate supported ___ upset voters in Florida?
    f. * Who did the candidate raise two million dollars by talking to ___?

Island constraints have attracted interest in a number of different sub-fields of linguistics. In theories of formal syntax, attention has focused on the issue of whether it is either possible or appropriate to give a unified account of all islands. In research on language processing, a number of studies have investigated whether real-time structure-building respects island constraints (see §2). Meanwhile, interest in explaining island phenomena in terms of constraints on processing has appeared in a number of different linguistic approaches spanning both formalist and functionalist traditions (see §3).

In theoretical accounts of island constraints, there is a 30-year history of attempts to unify all islands under a maximally simple generalization, going back at least as far as Chomsky's (1973) *subjacency constraint*, which brings together a number of islands under a constraint that blocks all wh-dependencies that cross two or more bounding nodes (NP or S) in one step. In line with this tradition in the theoretical literature, experimental studies of island constraints have often regarded all syntactic islands as equivalent (one notable exception is Neville, Nicol, Barss, Forster, & Garrett 1991). Meanwhile, the theoretical literature provides a number of reasons to assume that islands are not all created alike, due to differences in their cross-language distribution, due to sensitivity to discourse factors, and due to differential effects on argument and adjunct extraction (e.g. Cinque 1990, Manzini 1992).

The focus of this article is the ability of a subclass of island constraint violations to be 'rescued' by the presence of an additional well-formed wh-dependency, in what is known as *parasitic gap* constructions. The direct object NP gap in (3a) is impossible, as it occurs inside a subject. The gap in (3b) is acceptable, as it is a main clause direct object position. Surprisingly, when the unacceptable gap in (3a) is combined with the acceptable gap in (3b), the result is acceptable (3c). The first gap in (3c) is called a parasitic gap, since its well-formedness is contingent on the presence of another gap. Such constructions are probably not especially frequent in English, and they have often been viewed as marginal in the linguistic literature, but are, in fact, fully acceptable, as Experiment 1 below confirms.

(3) a. * What did the attempt to repair ___   ultimately damage the car?
     b.   What did the attempt to repair the car ultimately damage ___?
     c.   What did the attempt to repair ____{pg}   ultimately damage ___?

The fact that a well-formed dependency can rescue an illicit dependency is an interesting phenomenon in its own right, and has been the subject of intensive investigation in theoretical syntax for the past 25 years (see Culicover & Postal 2001 for a useful anthology). Parasitic gap constructions like (3c) are all the more interesting when viewed in the context of real-time language processing. First, they show that it is an oversimplification to assume that gaps cannot occur inside islands. Second, the parasitic gap in (3c) occurs before the main clause gap that licenses it, creating an apparent look-ahead problem for incremental parsing. This paper investigates how the parser addresses this look-ahead problem and discusses its implications for syntactic and psycholinguistic theories.

2. ISLAND CONSTRAINTS IN LANGUAGE PROCESSING. Studies of on-line sentence comprehension have revealed a good deal of information about the time-course of processing normal long-distance dependencies, and provide a useful starting point for comparison with island environments. Although the parser could, in principle, wait until it encounters direct evidence for a gap position in the form of a verb with an unfilled argument slot (Jackendoff & Culicover 1971; Wanner & Maratsos 1978), a good deal of evidence now indicates that the parser engages in a more active search process. It posits a gap as soon as a potential gap site can be identified, and does not wait to confirm that the gap site is not already occupied. For example, in head-initial languages a primary source of information about potential gap sites comes from verb argument structure and much evidence indicates that object gap sites are posited as soon as an appropriate transitive verb is encountered. Evidence for dependency formation at verb positions comes from a number of different sources, including filled gap effects in reading-time studies (Crain & Fodor 1985, Stowe 1986), eye-movement studies of implausibility detection (Traxler & Pickering 1996), antecedent reactivation effects (Nicol & Swinney 1989, Nicol, Fodor, & Swinney 1994; but cf. McKoon, Ratcliff, & Ward 1994), ERP measures (Garnsey, Tanenhaus, & Chapman 1989, Kaan, Harris, Gibson, & Holcomb 2000, Phillips, Kazanina, & Abada, 2004, Felser, Clahsen, & Münte 2003), or patterns of anticipatory eye-movements (Sussman & Sedivy 2003). Further evidence suggests that the wh-dependency is created immediately at the verb position even if the fronted phrase is an oblique phrase that could only be associated with a gap site that is separated from the verb by a direct object NP (Pickering & Barry 1991).[1] Furthermore, verbs do not provide the only cues to potential gap sites, and in cases where pre-verbal cues are available to signal upcoming argument positions there is evidence for pre-verbal construction of gap sites, both in English (Lee 2004) and in Japanese (Aoshima, Phillips & Weinberg 2004). Taken together, this body of research establishes the time-course of long-distance dependency formation, but does not indicate whether the mechanisms involved in forming these dependencies are restricted by island constraints.

A number of different studies over the past 20 years have investigated the island sensitivity of the parser using a variety of different types of islands and a number of different experimental

---

[1] Such cases were initially presented as evidence for approaches to long-distance dependencies in which the fronted phrase is associated directly with the verb, rather than with a gap/trace position. However, it is now generally agreed that such effects can be captured equally well in gap-based and gap-free theories (Gibson & Hickok 1993).

measures, and they have arrived at apparently conflicting conclusions. These studies are summarized in Table 1.

Study Measure	Structure Example	Island sensitivity
Stowe 1986 self-paced reading	PP in subject NP *The teacher asked what the silly story about <u>Greg's</u> older brother was supposed to mean.*	Yes
Pickering et al. 1994, exp. 2 eye-tracking, self-paced reading	PP in subject NP *I know what a book about <u>the local election</u> discussed the most*	Yes
Traxler & Pickering 1996 eye-tracking	Relative clause in subject NP *We like the city that the author who <u>wrote</u> unceasingly and with great dedication saw while waiting for a contract.*	Yes
Bourdages 1992 self-paced reading	Relative clause in object NP (French) See footnote 2	Yes
McElree & Griffith 1998 speeded grammaticality	Relative clause in object NP *It was the essay that the writer scolded the editor who admired?*	Yes
Yoshida et al. 2004 self-paced reading	*Relative clause in object NP (Japanese)* See footnote 3	Yes
Freedman & Forster 1985 sentence-matching	Complex NP with possessor *Who did the duchess sell Turner's portrait of?*	No
Kurtzman & Crawford 1989 speeded grammaticality	Infinitival in subject NP *Who did your attempt to instruct confuse?*	No
Clifton & Frazier 1989 speeded grammaticality	Relative clause in subject NP *What did John think the girl who always won received?*	No?
Pickering et al. 1994, exp. 1 eye-tracking, self-paced reading	Relative clause in subject NP *I realize what the artist who painted the large mural ate today.*	No?
Kluender & Kutas 1993 event-related potentials	Wh-island *What do you wonder who <u>they</u> caught at by accident?*	boundary detected
McKinnon & Osterhout 1996 event-related potentials	Wh-island *I wonder which of his staff members the candidate was annoyed <u>when</u> his son was questioned by?*	boundary detected
Neville et al. 1991 event-related potentials	Complex NP with possessor *What did the man admire Don's <u>sketch</u> of?*	boundary detected

Table 1: Summary of previous experimental studies of island constraints in language processing. Critical regions of example sentences are underlined where applicable.

Stowe's argument for the island sensitivity of the parser (Stowe 1986) was based on the fact that the 'filled gap effect' observed at simple licit gap sites was not found in a subject island context. Specifically, the NP *Greg's* in the example in Table 1, which is part of a complex subject NP, was read just as quickly in the example with a wh-phrase as in a control example with no wh-dependency, suggesting that no gap site is posited after the preposition *about*. A similar finding with similar constructions was obtained by Pickering, Barton & Shillcock (1994, expt. 2). The logic of the filled gap effect has also been used to show island sensitivity in French complex NPs (Bourdages 1992)[2] and Japanese relative clauses (Yoshida, Aoshima & Phillips

---

[2] Bourdages adapted the Filled Gap paradigm to French by using constructions with multiple PPs. She showed that in sentences that began with a fronted *with*-PP readers are surprised to encounter a second *with*-PP outside an island, but not when the second *with*-PP is inside an island, as at the underlined region in (i).

2004).[3] Traxler and Pickering (1996) also argue for island sensitivity based on the disappearance of an effect that is observed with well-formed dependencies. In the example in Table 1 the head of the relative clause (*the city*) is an implausible object of the closest verb (*write*), but that verb is embedded inside an additional relative clause. In contexts where the verb could be associated with a grammatical gap site the implausible verb-object combination led to reading time slowdown, relative to a plausible control condition, but no slowdown was observed in the example in Table 1 where the gap would violate the ban on extraction from subjects.

In contrast to the reports of immediate island sensitivity, some studies have argued that the parser is able to construct representations that violate island constraints. Freedman and Forster (1985) base their argument on measurements of the time needed to verify that a pair of sentences are identically matched. Building on the finding that matching times are faster for coherent sentences than for random word strings and certain types of ungrammatical sentences (e.g. agreement errors), Freedman and Forster reason that if the parser is able to represent sentences with island violations, then such sentences should show the same facilitation in sentence matching times as fully acceptable sentences. They show that matching times are indeed facilitated in sentences containing complex NP islands as illustrated in Table 1, and take this as evidence that island constraints act as late filters rather than as immediate constraints on structure building. Kurtzman and Crawford (1991) showed acceptance of gaps inside islands created by infinitival complements of subject NPs in a speeded grammaticality judgment task. We discuss this study in more detail in §4 below, since Kurtzman and Crawford were explicitly interested in parasitic gap constructions, which allow gaps inside a subclass of islands.

A different measure of on-line sensitivity to island constraints is observed in an ERP study by McKinnon and Osterhout (1996). This study shows that when readers enter an island domain while holding an incomplete wh-dependency this elicits the P600 brain response characteristic of syntactic anomaly detection. Two further ERP studies observed left-anterior negativity (LAN) responses one word after the beginning of an island domain (Neville et al. 1991, Kluender & Kutas 1993). These results show that English speakers are sensitive to the boundaries of islands, but since they only indicate that the start of the island domain is detected, they do not provide clear information on whether gaps are posited at potential gap sites inside islands. A related issue arises in a study by McElree & Griffith (1998) that used a speed-accuracy-tradeoff version of a speeded grammaticality task to investigate island violations involving relative clauses on object NPs. The model that best fit their data suggested an onset of sensitivity to the island violation within 100ms of the sentence-final verb, which also was within 350ms of the wh-word that marked the left boundary of the island. This indicates very fast detection of the island environment, whichever word is the trigger for anomaly detection, but is a more indirect measure of whether gaps are posited inside islands.

(i)   Avec qui le voisin a-t-il dit à la petite fille qui jouait <u>avec</u> son amie que sa mere est partie vers trois heures?
      'With whom did the neighbor say to the little girl who was playing with her friend that her mother left around three o'clock.'

[3] Yoshida et al. adapted the Filled Gap paradigm to Japanese by taking advantage of the possibility of scrambling of dative wh-phrases. Aoshima et al. (2004) had previously shown that Japanese speakers prefer to analyze a fronted dative wh-phrase as having undergone long-distance scrambling. One of the measures of this that they provide is a Filled Gap Effect, in which speakers slowdown upon encountering a in-situ dative NP in the embedded clause of a sentence that also contains a fronted dative wh-phrase. Yoshida et al extend this finding by showing that the Filled Gap Effect is not observed when the in-situ dative NP appears inside a relative clause, which is an island for scrambling in Japanese.

Two further studies have presented evidence that suggests that island constraints may be violated in parsing, but both studies are open to alternative interpretations. First, in a speeded grammaticality judgment task Clifton & Frazier (1989) presented sentences that contained a wh-phrase and an optionally transitive verb inside a relative clause ('won' in the example in Table 1), and compared this with sentences that replaced the optionally transitive verb with an obligatorily intransitive verb (e.g. 'excelled'). Results indicated slower judgments for sentences with an optionally transitive verb, and Clifton and Frazier interpret this as evidence that participants attempted to link the wh-phrase with the optionally transitive verb, ignoring the island constraint. However, the slowdown may have simply reflected readers' uncertainty over the argument structure of the optionally transitive verb. Second, in paired eye-tracking and self-paced reading studies, Pickering et al. (1994) used a filled gap paradigm to test for effects of active gap creation inside relative clauses. In both studies they found a slowdown at the position of the verb inside the relative clause, rather than at the subsequent 'filled gap' as in standard studies in this paradigm. Pickering and colleagues propose that the slowdown at the verb reflects formation of an illicit wh-dependency, but concede that it may simply reflect the overall processing load of the target sentences at that region, rather than an effect of dependency formation.

The prevailing opinion in psycholinguistics has been that the evidence supports the position that island constraints are immediately effective in parsing, and that contrary findings may be due to flaws in experimentation. The sentence matching studies, in particular, have been extensively criticized on methodological grounds (Crain & Fodor 1987, Stowe 1992). However, as can be seen from Table 1, it would not be easy to divide the studies on both sides of the debate in terms of the experimental measures used. An alternative possibility that is explored here is that variation in results may reflect differences in the specific islands used across studies. In particular, the existence of parasitic gap paradigms such as (3) above suggests that it is an oversimplification to assume that the grammar of English does not allow wh-dependencies to enter islands. Parasitic gap constructions provide a good reason why a parser might actively search for a gap position in at least some types of island.

3. PROCESSING ACCOUNTS OF ISLAND CONSTRAINTS. In contrast to purely formal accounts of island constraints, it has often been suggested that at least some island constraints may ultimately derive from constraints on language processing. These accounts can be divided into two broader classes. In one approach, island constraints are still explicitly represented in the grammar of a particular language, but the constraints are assumed to be the grammaticization of constraints on language processing (Fodor 1978, 1983, Berwick & Weinberg 1984, Hawkins 1999). In another approach, constraints on long-distance dependencies are not explicitly represented as part of a speaker's grammatical knowledge, and are merely reflections of structures that are difficult or impossible to process (Deane 1991, Pritchett 1991, Kluender & Kutas 1993).

For example, in the domain of subject islands, the primary focus of this article, Pritchett (1991) argues that it is impossible to incrementally construct a gap inside a subject island, as a consequence of the 'head driven' parsing architecture that he assumes. In Pritchett's parser, which strictly limits predictive structure building operations, the specifier and complement of a phrase cannot be attached into the parse tree until the head of the phrase has been attached. A consequence of this is that subject NPs cannot be attached into the larger parse tree until the head of the clause has been reached, i.e., an auxiliary or verbal inflection. In combination with the additional assumption that filler-gap dependencies require the filler and the gap to be a part of

the same syntactic tree, it follows that it is impossible to incrementally construct a filler-gap dependency into a subject NP.

Hawkins (1999) presents a processing-based account of subject islands that relies on the same configurational properties of subject NPs as Pritchett, although Hawkins assumes that island constraints must still be explicitly represented in some form in the grammars of individual languages. Hawkins' primary thesis is that there are close parallels across languages in which types of filler-gap dependencies are difficult to process, and that island constraints are closely tied to these scales of difficulty, but that languages are free to select different degrees of difficulty to 'conventionalize' as ungrammatical. Since ungrammaticality in this approach reflects the choice of a threshold of processing difficulty that is conventionalized differently in different languages, island constraints must still be explicitly represented in individual grammars, albeit in a different format than in formal accounts of islands. In the case of subject islands, Hawkins proposes that these filler-gap dependencies are more difficult to process because the subject NP that contains the gap precedes the verb that subcategorizes the subject NP, in contravention of his 'valence completeness' preference (p. 278).

Pritchett's and Hawkins' accounts of subject islands are based on a special property of subjects and therefore predict that subject islands have a different cause from other island phenomena. In contrast, Berwick & Weinberg (1984) propose that the subjacency constraint reflects an architectural property of the parser, and hence that all islands that are accounted for by subjacency ultimately reflect a common limitation on the processing of filler-gap dependencies. Berwick & Weinberg argue that when the parser constructs a gap position it must identify an appropriate antecedent phrase, and that the parser's search for the antecedent is restricted to crossing no more than one bounding node. Assuming that NP and S, or their equivalents, are bounding nodes, this means that gap inside a subject NP cannot be linked to an antecedent external to that NP during parsing.

All accounts that derive island constraints from limitations on language processing share the following straightforward prediction: island constraints should not be violated in real-time language processing. If the parser is able to posit gaps inside islands, then it is difficult to argue that island constraints reflect restrictions on the parser. This prediction has clear implications for the processing of parasitic gap constructions, and it is explored in detail in what follows.

4. PARASITIC GAPS. Parasitic gap constructions represent an important exception to the generalization that gaps may not occur inside islands, as seen in (3), repeated from above. A parasitic gap is a gap that is acceptable inside a syntactic island just in case (i) the sentence contains another gap that is not inside an island and (ii) both gaps are linked to the same wh-phrase. This is how the subject island violation in (3a) ceases to be problematic when it is combined with the simple object extraction in (3b) to create the parasitic gap construction in (3c).

(3) a.   * What did the attempt to repair ___   ultimately damage the car?
     b.   What did the attempt to repair the car ultimately damage ___?
     c.   What did the attempt to repair ___$_{pg}$ ultimately damage ___?

Parasitic gaps are interesting linguistic phenomena in their own right, but parasitic gaps inside subject islands are especially interesting from the perspective of real-time language processing, due to the fact that the illicit gap precedes the gap that licenses it. There are a number

of alternative ways in which the parser might approach the challenge posed by the subject parasitic gap in (3c).

A conservative approach would be for the parser to defer construction of the parasitic gap inside the subject island until after it has confirmed the presence of a licensing gap in the main clause. Under such an approach the parser could minimize the risk of constructing a dependency that might turn out to be ungrammatical, but this would come at the expense of fully incremental structure building. Note that in order to confirm the presence of the licensing gap it is not sufficient to detect an appropriate main clause verb, such as *damage* in the examples in (3), since this verb could be followed by an overt object NP, as in example (3a).

A variant of the conservative approach is one in which the parser posits a gap in (3c) as soon as it encounters clear evidence of the presence of a gap inside the subject island. In the case of (3c), this evidence could become clear when the adverb *ultimately* is reached, indicating that the obligatorily transitive verb *repair* lacks an overt object NP. Although this approach does not confirm the presence of a licensing gap before constructing a parasitic gap, it is still a conservative approach on the part of the comprehender, who may assume that his interlocutor produces a gap inside an island only if it will be followed by a subsequent licensing gap. Note that this approach runs into difficulties if the verb inside the island is optionally intransitive, since the presence of a gap is doubtful in such cases.

A third approach is one in which the parser constructs the parasitic wh-dependency in (3c) in just the same manner that it constructs other wh-dependencies, actively positing a gap as soon as it encounters a verb that can license a licit gap site. In the case of (3c), this means that a gap would be posited as soon as the parser reaches the verb *repair* inside the subject island, and before it encounters direct evidence about the presence or absence of a gap. This would then put the parser in a state where it must encounter a subsequent gap in the main clause in order for the sentence to be well formed. This is the one approach that allows the parser to preserve full incrementality, but it also requires that the parser incorporate the details of the grammar of parasitic gaps. Note also that in order to pursue this approach, the parser must be able to construct wh-dependencies that cross subject islands. This would appear to contradict suggestions that syntactic islands constrain the parser's search for gaps, and would therefore also challenge claims that island constraints derive from limitations on real-time processing.

The next section presents the results of an on-line experiment that distinguishes among these alternatives.

It is important to note that the distribution of parasitic gaps is restricted, and that there are many island-violating wh-dependencies that cannot be rescued by the presence of an additional licensing gap. For example, a finite relative clause inside a subject NP creates an island for wh-fronting (4a), similar to the subject island in (3a). However, when this gap is combined with the main clause gap in (4b) to create the combination in (4c), the result is still unacceptable.[4] This paradigm is confirmed with a large number of native speakers in Experiment 1 below. The illicit gap in (4a) differs from the illicit gap in the infinitival clause in (3a) in the respect that the environment of the gap in (3a) contains only one island-inducing component, the subject NP, whereas (4a) contains multiple island-inducing components, the subject NP, the adjunct status of

---

[4] This generalization appears to be at odds with some classic examples of acceptable parasitic gaps discussed by Kayne (1983) in which the parasitic gap is inside a finite relative clause, e.g. *She is the kind of person that everyone who meets ends up falling in love with.* Although I cannot provide a full account here of why this example is more acceptable than examples like (4c), it should be noted that all of Kayne's examples of acceptable parasitic gaps inside relative clauses used relative clauses with a quantificational head NP.

the relative clause, and the wh-island created by relativization of the subject of the relative clause.

(4) a. *What did the reporter that criticized ___ eventually praise the war?
     b.   What did the reporter that criticized the war eventually praise ___?
     c. *What did the reporter that criticized ___ eventually praise ___?

Examples like (4c) show that preservation of accuracy and incrementality does not require the parser to actively posit gaps inside all types of syntactic islands. If the parser were to actively construct a gap upon encountering the embedded verb *criticize* in (4c), there would be no possibility for a grammatical completion to the sentence. A fully accurate and incremental parser should actively posit a gap inside an island only when the gap may be licensed as a well-formed parasitic gap.

There has been only one previous experimental study of subject islands that support parasitic gaps. Kurtzman & Crawford (1991) present the results of a number of speeded grammaticality judgment studies in which participants read sentences in a self-paced reading paradigm and gave acceptability judgments on what they had read so far as soon as a beep sounded. Across a number of studies they found that when the beep sounded at the final verb of sequences like *Who did your attempt to instruct confuse* participants judged the string to be acceptable in 60-68% of trials, suggesting acceptance of a gap inside a subject island. Although this shows less than complete acceptance, it is almost identical to the level of acceptance found in unambiguously acceptable sentences like *Who did your attempt to instruct Jim confuse?* Kurtzman & Crawford also showed that acceptability of parasitic gaps was lower when the subject NP contained a relative clause, as in *Who did your statement that you instructed confuse?* The results of these studies suggest that speakers may be able to accept subject parasitic gap constructions after only brief processing time. However, since the judgments were always given after the main verb was presented, the results provide no indication of whether the parser engages its normal active gap creation mechanisms inside islands, or whether the parasitic gaps are constructed retroactively, after encountering the main verb.

The syntactic literature on parasitic gaps contains extensive discussion of different possible analyses of the parasitic gap phenomenon. This literature encompasses questions of whether the parasitic gap is the same as other gaps created by extraction in English or whether it is a null pronominal, whether parasitic gap phenomena are related to other multiple-gap constructions such as Across-the-Board extractions, and whether the parasitic gap is directly linked to the overt antecedent or linked to an additional null operator. The reader is referred to Culicover (2001) for an excellent review of syntactic literature on parasitic gaps. However, since the focus of this article is on the basic question of whether any kind of dependency is formed between a wh-phrase and a verb inside a syntactic island, and all theoretical analyses assume that the parasitic gap is in some way dependent on the presence of the licensing gap, the current study has similar relevance across all of the different theoretical approaches.

Summarizing, parasitic gaps inside subject islands present an interesting challenge for models of real-time structure building. A good deal of evidence indicates that language processing is highly incremental and that long-distance filler-gap dependencies are constructed 'actively', meaning that the parser posits gap sites before receiving confirmation of an open argument slot. In order to accommodate parasitic gap constructions the parser either must sacrifice incrementality or grammatical accuracy or it must be equipped to actively posit gaps

inside the subclass of syntactic island environments that support parasitic gaps. We now turn to experimental studies that investigate what the parser does in such cases.

5. EXPERIMENT 1: OFF-LINE JUDGMENTS. A first step in assessing the relation between off-line acceptability judgments and on-line structure building is to confirm that uninitiated native speakers share the judgments assumed by professional linguists. This step is particularly relevant in the case of parasitic gaps, since even linguists have often viewed them as somewhat marginal in acceptability. The primary aim of this study was to compare the acceptability of parasitic gap constructions with the acceptability of the individual gaps that are combined in the parasitic gap constructions. A secondary aim was to compare the acceptability of the islands investigated here with the island contexts investigated in other studies of the island constraints in language processing. The results of this second comparison are presented in the General Discussion section below.

5.1. PARTICIPANTS. Acceptability ratings were collected from 51 undergraduate students at the University of Maryland, all of whom gave informed consent and were paid $10/hour for their participation. None of the participants had any familiarity with syntactic theory or psycholinguistics, and all were naïve to the purpose of the study.

All of the participants completed the acceptability rating questionnaire after first completing the reading time study in Experiment 2. However, we report the results of the rating study first, since they establish a premise for the reading-time study. The reading time study contained no examples of parasitic gaps or other syntactic island violations, so it was unlikely to have affected the outcome of the acceptability rating study.

5.2. MATERIALS. The materials for this study consisted of 24 sets of items that tested the components of parasitic gap constructions and 60 additional items that tested a variety of acceptable *wh*-dependencies and syntactic island violations. The parasitic gaps items consisted of 6 conditions in a 2 x 3 factorial design that manipulated gap-type (good gap, bad gap, both gaps) and finiteness (infinitival complement vs. finite relative clause). All conditions began with a main clause that contained an interrogative complement clause that started with the wh-word *what* or *who*. The embedded clause contained a complex subject NP and at least one gap associated with the wh-word. In the 'good gap' conditions, there was a direct object gap in the same clause as the wh-word. This condition provided a baseline measure of the ratings for an unambiguously acceptable wh-dependency in a moderately complex sentence. In the 'bad gap' conditions there was a direct object gap inside the complex subject NP. This condition provided a measure of the informants' sensitivity to subject islands. The 'both gaps' conditions combined the good gap and the bad gap to create a potential parasitic gap construction. The complex subject NP contained an infinitival complement clause in the infinitival conditions, and a finite relative clause in the finite conditions. A sample set of items is shown in Table 2. The items for this study were not matched with the items in the reading-time study. The primary reason for this is that it was necessary in the acceptability rating study to use unambiguously transitive verbs in the embedded clauses, in order to provide clear cues to the presence of syntactic gaps. In contrast, the design of the on-line study required different types of verbs to be used.

Infinitival	*good*	The outspoken environmentalist worked to investigate what the local campaign to preserve the important habitats had harmed ___.
	bad	The outspoken environmentalist worked to investigate what the local campaign to preserve ___ had harmed the annual migration.
	both	The outspoken environmentalist worked to investigate what the local campaign to preserve ___ had harmed ___.
Finite	*good*	The outspoken environmentalist worked to investigate what the local campaign that preserved the important habitats had harmed ___.
	bad	The outspoken environmentalist worked to investigate what the local campaign that preserved ___ had harmed the annual migration.
	both	The outspoken environmentalist worked to investigate what the local campaign that preserved ___ had harmed ___.

Table 2: Sample set of items for acceptability ratings in Experiment 1.

The 24 sets of items were distributed among 6 lists in a Latin Square design, such that each list contained one version of each item and equal numbers of each condition. The lists were combined with the 60 filler items and two randomizations of each list were generated to create a total of 12 versions of the questionnaire. Participants were asked to rate each sentence on a scale from 1 (unacceptable) to 5 (acceptable) using pencil and paper. The questionnaire began with a small number of sample items and instructions that distinguished acceptability from plausibility. The questionnaire took around 15-20 minutes to complete.

5.3. RESULTS. Results of the rating study confirm the generalizations about parasitic gaps presented in section 4 and also suggest that parasitic gaps do not deserve their reputation as merely marginally acceptable constructions. Mean ratings for the 6 parasitic gaps conditions are shown in Figure 1.
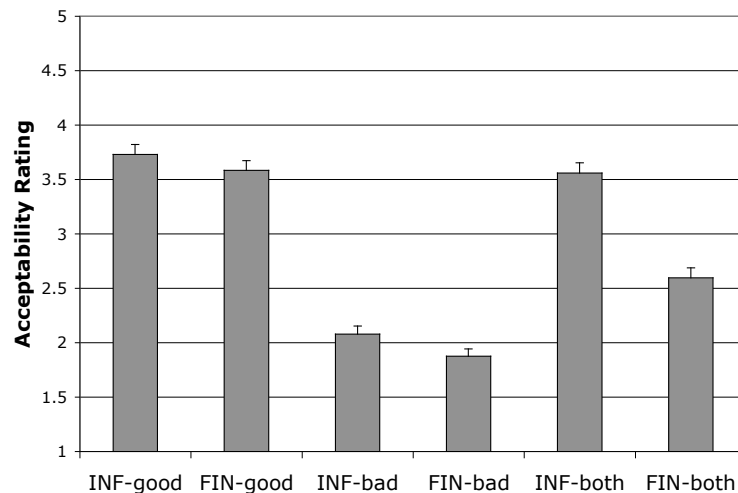

Figure 1: Mean acceptability ratings, Experiment 1.

A repeated measures ANOVA showed reliable main effects of gap-type ($F_1(2,100) = 117.49$, $p < .0001$; $F_2(2,100) = 142.21$, $p < .0001$) and finiteness ($F_1(1,50) = 37.74$, $p < .0001$; $F_2(1,50) = 53.38$, $p < .0001$), and a significant interaction of gap-type and finiteness ($F_1(2,100) = 19.88$, $p < .0001$; $F_2(2,100) = 17.39$, $p < .0001$). Planned comparisons within each level of the gap-type

factor showed that manipulation of finiteness did not affect ratings for the good gap conditions (F1(1,50) = 1.97, p = .17; F2(1,50) = 1.08, p = .31), and had a small but reliable effect on ratings for the bad gap conditions (F(1,50) = 5.49, p < .05; F2(1,50) = 5.56, p < .05), due to scores that were 0.2 lower for the finite conditions. In contrast, the effect of finiteness was highly reliable for the both gap conditions (F(1,50) = 51.85, p < .0001; 106.05, p < .0001).

Comparison of ratings for the infinitival good gap and both gaps conditions showed no significant difference (F1(1,50) = 2.77, p = .103; F2(1,50) = 1.54, p = .23). A similar comparison within the finite conditions showed a highly reliable difference (F1(1,50) = 51.63, p < .0001; F2(1,50) = 54.45, p < .0001), due to ratings that were on average 0.96 points lower in the finite bad gap condition. This pattern of results indicates that addition of the gap inside the subject island has little or no impact upon acceptability when the subject contains an infinitival complement. This confirms that parasitic gaps are a real phenomenon, and are not marginally acceptable for uninitiated English speakers. The sharply reduced ratings for the finite both gaps condition is consistent with the claim that parasitic gaps are restricted to a subclass of island types. Note, however, that even in the finite conditions there is a partial 'rescuing' effect when the bad gap and the good gap are combined, with the both gaps condition receiving a higher rating than the bad gap condition.

There is a possible concern over that the mean ratings for even the acceptable conditions in this study were rather low, in the 3.5-3.7 range on a 1-5 scale. However, experience with studies of this kind indicates that mean scores are always lowered when any degree of complexity is introduced into the test items. A baseline measure is provided by the 6 filler items in this study that contained simple 2-clause *wh*-dependencies, such as *The used-car salesman remembered what the racecar driver said that the skillful mechanic had fixed*. These sentences received a mean rating of 4.01.

Having established using off-line judgments that normal speakers do allow *wh*-dependencies to enter syntactic islands in parasitic gap constructions, the next step is to use on-line measures to investigate the time-course of the construction of these dependencies.

6. EXPERIMENT 2: ON-LINE READING-TIME STUDY. The aim of this study was to test whether the active gap-creation mechanisms used in the processing of normal wh-dependencies operate in the same manner in parasitic gap contexts. Specifically, the experiment was designed to check whether a wh-dependency is constructed as soon as the parser encounters the verb inside the subject island in sentences like (3c), without waiting to confirm the presence of a parasitic gap or a licensing gap. In addition, the study was designed to test whether the parser is able to immediately distinguish between infinitival subject islands, which support parasitic gaps, and finite relative clause subject islands, which do not.

The experimental measure of the construction of a wh-dependency in this study involved the implausibility detection technique previously used in a number of studies using the 'stops making sense task' (Tanenhaus, Carlson, & Trueswell 1989, Boland, Tanenhaus, Garnsey, & Carlson 1995), event-related potentials (Garnsey et al. 1989) and eye-tracking (Traxler & Pickering 1996).

6.1. PARTICIPANTS. Participants for this study were the same 51 University of Maryland undergraduates that participated in Experiment 1. They were paid $10 for the roughly one hour that it took to complete all of the tasks in the study. The reading-time data from 3 participants was not included due to technical problems, leaving a total of 48 participants.

6.2. MATERIALS. Experimental items were 24 sets of 4 conditions, organized in a 2 x 2 factorial design that manipulated the factors plausibility (plausible vs. implausible) and finiteness (infinitival complement vs. finite relative clause). In all conditions the main clause was followed by an embedded interrogative clause that contained a wh-phrase and a direct object gap in the same clause, and the wh-phrase was a plausible direct object of the verb associated with the gap. For example, in the examples in Table 3, it is plausible to overburden a school and also plausible to motivate high school students. Thus, all sentences were globally ambiguous and plausible.

However, each target sentence also contained an additional verb that was embedded inside the subject NP following the wh-phrase, and the primary concern of the study was to test whether speakers attempt to construct a wh-dependency upon encountering this verb. To this end, the wh-phrase was manipulated such that it was either a plausible or an implausible direct object of this verb. For example, in the examples in Table 3 it is plausible to expand a school, but implausible to expand high school students. Previous findings suggest that if speakers attempt to construct a wh-dependency upon reaching this verb there should be a slowdown in reading times associated with detection of the implausibility (Traxler & Pickering 1996). The clause containing this critical verb was either an infinitival complement of the subject NP or a finite relative clause modifier of the subject NP. If the parser constructs wh-dependencies into islands where allowed by the grammar of parasitic gaps, then a plausibility-related slowdown is expected in the infinitival conditions but not in the finite conditions.

Infinitival	*plausible*	The school superintendent learned which schools the proposal to expand drastically and innovatively upon the current curriculum would overburden ___ during the following semester.
	implausible	The school superintendent learned which high school students the proposal to expand drastically and innovatively upon the current curriculum would motivate ___ during the following semester.
Finite	*plausible*	The school superintendent learned which schools the proposal that expanded drastically and innovatively upon the current curriculum would overburden ___ during the following semester.
	implausible	The school superintendent learned which high school students the proposal that expanded drastically and innovatively upon the current curriculum would motivate ___ during the following semester.

Table 3: Sample set of experimental conditions, Experiment 2. The highlighted regions indicate the wh-phrase and the verb inside the complex subject NP.

Note that although the interest of the study was to test whether potential parasitic wh-dependencies are constructed at the verb inside the subject island, none of the target items or fillers contained an actual parasitic dependency. In all target sentences, the only gap was in the same clause as the wh-phrase. Therefore, any evidence for the construction of parasitic wh-dependencies must reflect the participants' own parsing biases, rather than priming effects from the experimental materials.

The subject NPs that followed the wh-phrase were chosen such that they could both take an infinitival complement and be the subject of a finite relative clause. The 9 nouns used were *idea*, *plan*, *effort*, *campaign*, *scheme*, *request*, *attempt*, *struggle*, and *proposal*. The verbs inside the subject NP all allowed either an NP or a PP complement, and it was always the PP complement

that was used in the target sentences. The PP was separated from the verb by a three-word adverbial phrase that delayed confirmation of the argument structure of the verb.

The 24 sets of experimental items were distributed among 4 lists in a Latin Square design and combined with 72 filler items of comparable length and difficulty and presented in pseudorandom order. A full set of experimental items is available on request.

6.3. PLAUSIBILITY QUESTIONNAIRE. An off-line plausibility rating questionnaire was conducted in order to confirm that the implausible verb-object combinations were indeed judged to be implausible, and to confirm that the implausible conditions were equally implausible in finite and infinitival conditions alike. This concern arises from the fact that the NP following the *wh*-phrase (e.g., *plan, proposal, campaign*) has a different syntactic status in the finite and infinitival conditions. In the finite conditions it receives the agent thematic role in the relative clause. In the infinitival conditions it does not receive a thematic role in the infinitival clause. The 51 participants in the on-line study answered the plausibility questionnaire after completing the on-line study, so that the questionnaire materials would not affect on-line reading times. Note that the only way in which participants could have previously considered the verb-object combinations tested in the plausibility questionnaire would be if they had considered constructing a wh-dependency into an island in the on-line study.

The conditions for the plausibility questionnaire were derived from the materials for the on-line study, but were simplified in order to remove any effects of the *wh*-dependency or the subject island. The 24 sets of 4 items were distributed among 4 lists in a Latin Square design, and two randomizations of each list were generated to create a total of 8 versions of the questionnaire. Participants rated the plausibility of each sentence on a scale from 1 (implausible) to 5 (plausible). Table 4 shows a sample set of items, derived from the items shown in Table 3.

infinitive	*plausible*	The superintendent made the proposal to expand the schools.
	implausible	The superintendent made the proposal to expand the high school students.
finite	*plausible*	The superintendent made the proposal that expanded the schools.
	implausible	The superintendent made the proposal that expanded the high school students.

Table 4: Sample items for plausibility questionnaire.

Mean ratings for the plausibility questionnaire are shown in Table 5. A 2 x 2 repeated measures ANOVA showed that there was a main effect of plausibility ($F_1(1,50) = 215.6$, $p < .0001$; $F_2(1,23) = 160.5$, $p < .0001$), a marginally significant tendency for higher ratings in the infinitival conditions ($F_1(1,50) = 3.39$, $p = .07$; $F_2(1,23) = 2.46$, $p = .13$), and an interaction of plausibility and finiteness ($F_1(1,50) = 10.22$, $p < .01$; $F_2(1,23) = 4.74$, $p < .05$). However, planned comparisons showed that this interaction was due to a small but reliable difference in the ratings for the two plausible conditions ($F_1(1,50) = 14.37$, $p < .001$; $F_2(1,23) = 8.14$, $p < .01$), but that there was no difference in the plausibility ratings for the two implausible conditions ($Fs < 1$).

	infinitive	*finite*
plausible	4.55 (.87)	4.20 (1.10)
implausible	2.56 (1.55)	2.63 (1.38)

Table 5: Mean ratings in the plausibility questionnaire. Standard deviations are shown in parentheses.

6.4. PROCEDURE. The experiment was conducted on Macintosh computers running the `mw-run` software developed at MIT. Participants were timed in a word-by-word self-paced non-cumulative moving-window reading task (Just, Carpenter, & Woolley 1982). Stimulus segments initially appeared as a row of dashes, and participants pressed the space bar of the keyboard to reveal each subsequent region of the sentences. In order to ensure that participants attended to the stimuli, a yes-no comprehension question was presented after each trial, and visual feedback indicated whether the answer given was incorrect. All trials on which the comprehension question was answered incorrectly were excluded from further analysis. The experimental trials were preceded by two screens of instructions and a small number of practice trials. Analyses were conducted on comprehension task response accuracy and reading times.

6.5. RESULTS.

6.5.1. COMPREHENSION ACCURACY. Mean accuracy on the yes/no comprehension questions for the experimental items was 91.6%, with individual condition means of 89.2% and 89.9% for the infinitival and finite plausible conditions respectively, and 92.0% and 95.1% for the infinitival and finite implausible conditions respectively. A repeated measures ANOVA on accuracy scores with the factors finiteness and plausibility showed that there was no main effect of finiteness ($F1(1,47) = 1.73$, $p = .19$; $F2(1,23) = 1.97$, $p = .17$) and that there was a main effect of plausibility, although it was only marginally significant in the items analysis ($F1(1,47) = 4.80$, $p < .05$; $F2(1,23) = 2.95$, $p < .1$). There was no interaction of finiteness and plausibility ($Fs < 1$). Scheffé post-hoc tests showed that the accuracy on the implausible finite conditions was reliably higher than that on either of the plausible conditions, but that no other pairs were reliably different. It is possible that the high accuracy scores in the implausible finite condition reflect the fact that this condition is the least likely to involve a garden-path due to initially creating a gap inside the subject island. Three experimental items for which comprehension questions were answered incorrectly on more than one sixth of trials were excluded from further analyses.

6.5.2. SELF-PACED READING. A regression equation predicting reading time from word length was constructed for each subject, using all items (filler and experimental). At each word, the reading time predicted by the subject's regression equation was subtracted from the actual measured reading time, and all analyses were performed on these differences (residual reading times). This transformation removes extraneous variance by subtracting out a baseline for each subject, and by controlling for noise due to length effects (Ferreira & Clifton 1986, Trueswell & Tanenhaus 1991). All times reported here are based on the residual reading times for trials in which the comprehension question was answered correctly. Reading times greater than 1500ms were excluded, affecting 0.4% of trials, and residual reading times greater than 500ms were trimmed to 500ms, a value that lay approximately 3 standard deviations from the mean and affected an additional 1.5% of trials.

Residual reading times for each of the first 14 regions were entered into a 2 x 2 repeated measures ANOVA, with the factors finiteness (infinitival vs. finite) and plausibility (plausible vs. implausible). In cases where the results of the ANOVA showed significant or marginally significant effects, planned comparisons were run to test for the effect of plausibility at each level of the finiteness factor. Mean reading times for the infinitival and finite conditions are shown in Figures 2 and 3 respectively.

In the first 6 regions, which lasted through the end of the *wh*-phrase, there were no significant main effects or interactions. The only effect that approached significance in this interval was at the subject noun in region 3, where there was a marginally significant main effect of plausibility, due to slower reading times in the plausible conditions $(F1(1,47) = 2.67, p = .11; F2(1,20) = 3.56, p = .074)$. However, pairwise comparisons showed only a non-significant trend in the infinitival conditions $(F1(1,47) = 1.83, p = .18; F2(1,20) = 2.51, p = .13)$, and no effect of plausibility in the finite conditions $(Fs < 1.4)$.
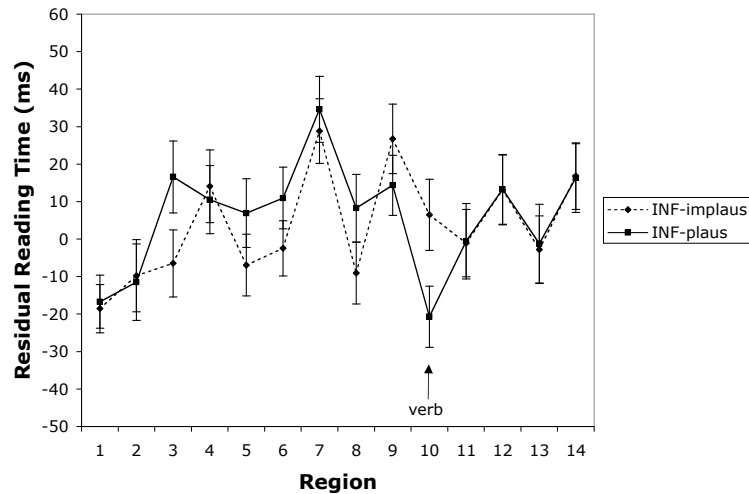
The$_1$ school$_2$ superintendent$_3$ learned$_4$ which$_5$ schools$_6$ the$_7$ proposal$_8$ to$_9$ expand$_{10}$ drastically$_{11}$ and$_{12}$ innovatively$_{13}$ upon$_{14}$ …

Figure 2: Experiment 2, mean residual reading times in infinitival conditions.
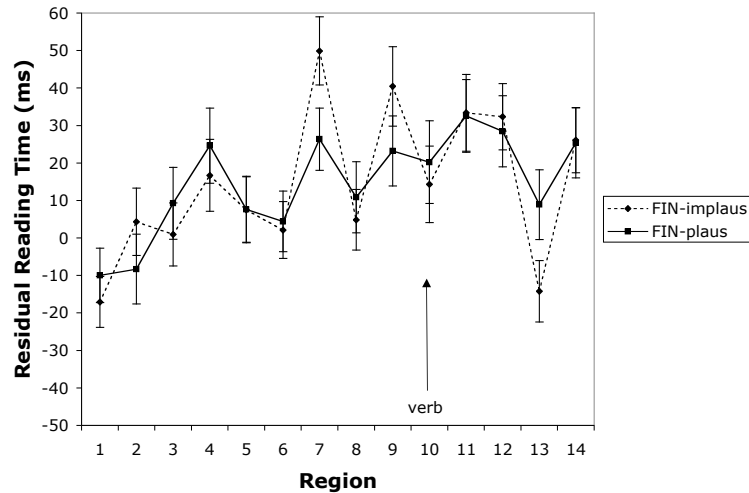
The$_1$ school$_2$ superintendant$_3$ learned$_4$ which$_5$ schools$_6$ the$_7$ proposal$_8$ that$_9$ expanded$_{10}$ drastically$_{11}$ and$_{12}$ innovatively$_{13}$ upon$_{14}$ …

Figure 3: Experiment 2, mean residual reading times in finite conditions.

At the determiner in region 7 the main effect of plausibility was marginally significant in the participants analysis and showed no effect in the items analysis ($F1(1,47) = 2.91$, $p = .09$; $F2(1,20) = 1.11$, $p = .30$). Pairwise comparisons showed that there was no effect of plausibility in the infinitival conditions (Fs < 1), and that in the finite conditions the effect of plausibility was significant in the participants analysis and marginally significant in the items analysis, due to slower reading times for the implausible condition ($F1(1,47) = 4.54$, $p < .05$; $F2(1,20) = 3.52$, $p = .08$).

At the noun in region 8 there were no significant effects or interactions. At region 9, which contained *to* in the infinitival conditions and *that* in the finite conditions, the main effect of plausibility was marginally significant in the participants analysis and significant in the items analysis ($F1(1,47) = 3.28$, $p = .077$; $F2(1,20) = 4.71$, $p < .05$), due to reading times that were on average 14.8ms slower in the implausible conditions. However, planned comparisons showed that the effect of plausibility was not significant or marginally significant in either the infinitival or the finite conditions (infinitival: $F1(1,47) = 0.98$, $p = .33$; $F2(1,20) = 1.04$, $p = .32$; finite: $F1(1,47) = 1.09$, $p = .30$; $F2(1,20) = 2.04$, $p = .17$).

At the critical embedded verb in region 10 there was a main effect of finiteness ($F1(1,47) = 6.54$, $p < .05$; $F2(1,20) = 4.96$, $p < .05$), no main effect of plausibility ($F1(1,47) = 1.53$, $p = .22$; $F2(1,20) = 2.05$, $p = .17$), and an interaction of finiteness and plausibility that was significant in the participants analysis and marginally significant in the items analysis ($F1(1,47) = 5.78$, $p < .05$; $F2(1,20) = 3.26$, $p = .086$). The primary test of the experimental hypothesis involves the effect of plausibility within each level of the finiteness factor. In the infinitival conditions there was a significant effect of plausibility ($F1(1,47) = 6.95$, $p = .011$; $F2(1,20) = 4.20$, $p = .054$), due to average reading times that were 27ms slower in the implausible conditions. In the finite conditions there was no effect of plausibility (Fs < 1).

At the adverb in region 11 there was a significant main effect of finiteness, due to slower reading times in the finite conditions ($F1(1,47) = 15.59$, $p < .001$; $F2(1,20) = 11.02$, $p < .01$), but there was no hint of a main effect or interaction involving plausibility (all Fs < .1). The effect of

finiteness was also weakly present at the conjunction in region 12 ($F1(1,47) = 3.37$, $p = .07$; $F2(1,20) = 2.73$, $p = .11$), where there was again no main effect or interaction involving the plausibility factor (all $Fs < 1$). At the adverb in region 13 there were slower reading times in the finite plausible condition than in the finite implausible condition. This led to a main effect of plausibility that was marginally significant in the participants analysis ($F1(1,47) = 3.13$, $p = .09$; $F2(1,20) = 1.83$, $p = .19$). Planned comparisons showed that there was no effect of plausibility in the infinitival conditions ($Fs < .1$), and that the effect of plausibility was significant in the participants analysis and marginally significant in the items analysis ($F1(1,47) = 4.50$, $p < .05$; $F2(1,20) = 3.61$, $p = .07$). At the preposition in region 14 there were no significant effects ($Fs < 1.4$). Subsequent regions were not considered since they are not directly relevant to the experimental hypotheses about the processing of the gap inside the subject island.

7. DISCUSSION.

7.1. PROCESSING PARASITIC GAPS. The immediate goal of this study was to test whether the parser incrementally constructs gaps inside subject islands that support parasitic gaps, using the same active gap creation mechanisms used in processing normal unbounded dependencies. The results from the infinitival conditions in Experiment 2 showed an immediate effect of the semantic fit between the wh-phrase and the verb inside a subject island. This indicates that comprehenders attempt to create a dependency between the wh-phrase and the verb in a sequence like *which schools the proposal to expand* as soon as they encounter the verb, without waiting for further confirmation of a gap. Thus, the same active gap creation mechanisms seem to operate here as in simple clauses.

Although the critical verbs in this study did not require a direct object NP, the plausibility effect indicates that comprehenders temporarily analyzed the wh-phrase as the direct object of the verb inside the island. If the verb inside the island were simply analyzed as intransitive and not related to the *wh*-phrase, no plausibility effect should have occurred.

The findings in the infinitival conditions are relevant to the suggestion by Goodluck and colleagues that on-line gap creation mechanisms might be restricted to apply only at a 'potential end of sentence' (Goodluck, Finney, & Sedivy 1991; Bourdages 1992). This generalization offered one possible account of why active gap creation was not observed in the subject islands in Stowe's (1986) study of islands. However, since active gap creation is observed inside the infinitival subjects in the current study, the domain of active gap creation needs to be more fine-grained than 'potential end of sentence'.

The plausibility effect observed in the infinitival conditions was not found in conditions where the subject island contained a finite relative clause, indicating that gap creation was not attempted in these islands. This finding is consistent with a previous study of the processing of subject islands containing relative clauses (Traxler & Pickering 1996). The contrast between the finite and infinitival conditions shows that gap creation does not apply indiscriminately inside islands, as is to be expected from previous experimental results. More interesting for current purposes is the convergence between the class of subject islands that show active gap creation (Experiment 2) and the class of subject islands that show acceptability of parasitic gap constructions (Experiment 1). This suggests that the parser's apparent violation of an island constraint in the infinitival conditions is directly related to the possibility of a parasitic gap construction. The difference between the finite and infinitival conditions is unlikely to be due to the overall acceptability of the subject islands themselves, since they received similarly low

ratings in Experiment 1. It is also unlikely to reflect plausibility differences between the finite and infinitival conditions, since the two implausible conditions were rated as equally implausible. Finally, the active construction of gaps inside the infinitival subject islands is unlikely to reflect an experimentally induced priming effect, since participants never saw a confirmed parasitic gap in the reading time experiment. Also, any effect of the experimental items should have affected finite and infinitival conditions similarly.

In sum, the results of the on-line study suggest that the parser avoids constructing gaps inside islands that cannot be licensed, but actively constructs gaps inside islands that may subsequently be licensed as parasitic gap constructions.

A logical next question is what kind of mechanism could implement the grammar of parasitic gaps so accurately without look-ahead or backtracking? The following is one way to achieve this, by implementing a left-to-right version of Kayne's (1983) Connectedness theory and its descendants (e.g. Manzini 1994). The leading idea in Kayne's approach to parasitic gaps is stated in tree-geometric terms, and thus is easily implemented in a parser. Kayne proposes that the basic property of a well-formed wh-dependency is that the path from the wh-phrase to the XP node that contains the gap traverses a sequence of XP nodes each of which is the complement of the next. The special property that allows parasitic gap constructions under Kayne's approach is that a pair of paths can be treated as 'connected' and hence well-formed when they include nodes that are sisters, as illustrated in Figure 4. Following this reasoning, the infinitival subject islands in the current study allow parasitic gaps because there is a well-formed path that extends from the licensing gap to the wh-phrase and another well-formed path that extends from the gap inside the subject NP as far as the top node of the subject, and because the top node of this second path is a sister to a node that participates in the first path, allowing the paths to be connected. In a left-to-right structure building mechanism this account of wh-dependencies could be implemented as a requirement that a path from a wh-phrase to a gap traverse only head-complement relations, except at nodes where the path splits into two separate paths, each of which subsequently traverses only head-complement relations. Thus, the parser could pursue a path for a wh-dependency into a subject NP in search of a gap, but only if it treats this as one of a pair of connected paths, where the second member of the pair is a path that continues to follow head-complement relations into the VP in search of an additional gap. Once inside the subject NP, the path can continue on to construct a direct object gap in the infinitival complement of the subject NP, as the rest of the path involves only head-complement relations. In the case of the finite conditions, on the other hand, the path is unable to continue into the relative clause modifying the subject NP, since the path from the subject NP into the relative clause involves a non-complement relation.

Note that this is only one way in which the experimental findings might be implemented, and that the experimental results do not provide a reason to favor any individual theoretical account of parasitic gap constructions.
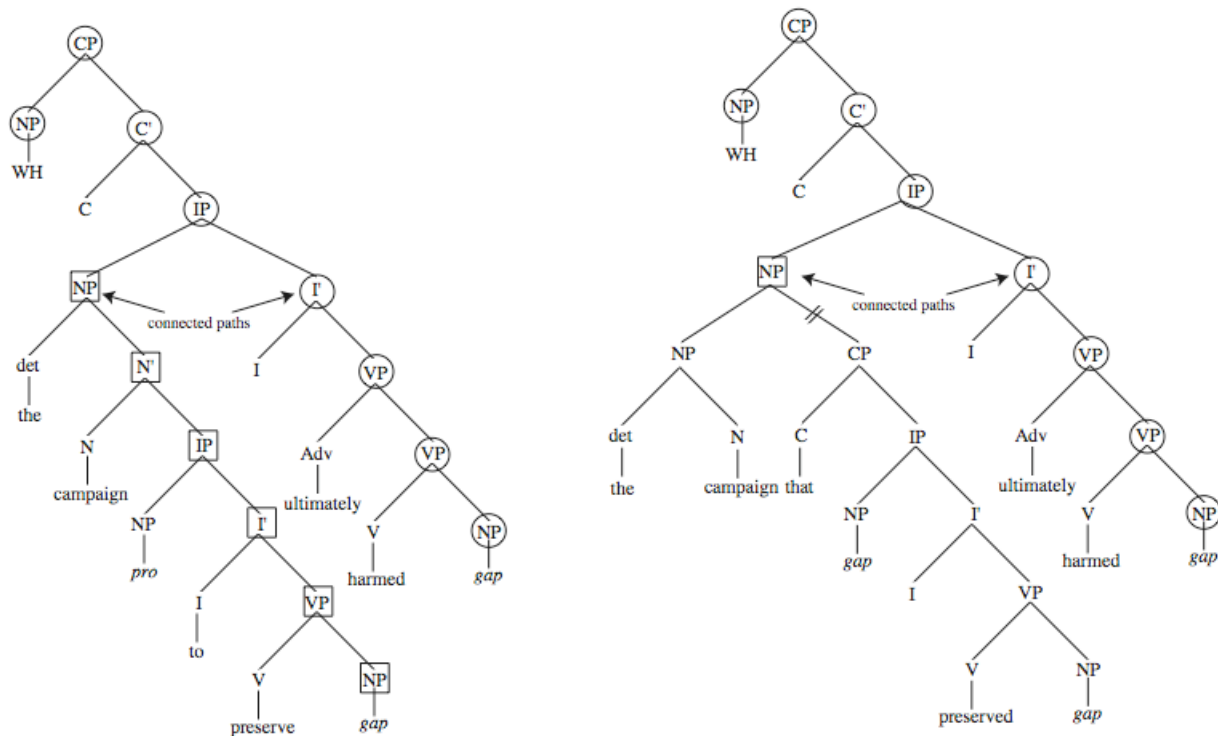
Figure 4: Illustration of path-based approach to on-line parasitic gap licensing, for infinitival (left) and finite (right) clauses attached to subject NPs.. Circled nodes indicate the primary path from the wh-phrase to a licit gap. Squares indicate a second path that is connected to the first path by virtue of sisterhood. Paths terminate upon encountering an XP node that is not the complement of the head of the next higher phrasal node in the path. In the finite relative clause example, licensing of the parasitic gap fails because the CP of the relative clause is neither the complement of the immediately dominating NP nor connected via sisterhood to a well-formed path.

7.2. ISLAND CONSTRAINTS IN REAL-TIME. Previous studies of island constraints in parsing have arrived at apparently conflicting conclusions. Some studies report immediate effects of island constraints, whereas others do not. Although it is possible to attribute these differences to artifacts of different participants, different materials or different experimental measures, this is not possible for the current results. Experiment 2 showed that comprehenders posit gaps inside some islands but not others, using the same participants and the same experimental manipulation. This shows that differences among islands matter for real-time processing, and leads to the question of whether the variability in previous findings is also due to differences in the structures tested.

In order to address this question, sample items from previous studies of islands in parsing were included among the filler items in the acceptability rating study in Experiment 1. Examples of these different constructions are shown in Table 6 and mean ratings are shown in Figure 5.

Condition	Example (study)
WH-island – Argument	The flower girl knew who the beautiful bride admired what she had gotten __ from __ at the reception. (Kluender & Kutas 1993)[5]
PG-FIN-bad	The outspoken environmentalist worked to investigate what the local campaign that preserved __ had harmed the annual migration. (this study)
PG-INF-bad	The outspoken environmentalist worked to investigate what the local campaign to preserve __ had harmed the annual migration. (this study)
WH-island - whether	The prosecutor tried to discreetly find out what the defense attorney knew whether the defendant had done __.
Subject + PP	The bookstore clerk knew who the new book about __ aimed to make __ respectable once again. (Stowe 1986; Pickering et al. 1994, expt. 2)[6]
PG-FIN-both	The outspoken environmentalist worked to investigate what the local campaign that preserved __ had harmed __. (this study; Clifton & Frazier, 1989; Pickering et al. 1994, expt 1; Traxler & Pickering 1996)
Complex NP	The publicity manager knew what the casting agency had seen John's picture of __. (Freedman & Forster 1985)
PG-INF-both	The outspoken environmentalist worked to investigate what the local campaign to preserve __ had harmed __. (this study)
PG-FIN-good	The outspoken environmentalist worked to investigate what the local campaign that preserved the important habitats had harmed __. (this study)
PG-INF-good	The outspoken environmentalist worked to investigate what the local campaign to preserve the important habitats had harmed ___. (this study)
2 clause WH	The cynical skeptic knew what the spiritual woman thought the psychic advisor saw __.
1 clause WH	The observant resident saw who the local policeman arrested __.

Table 6: Sample items used for acceptability ratings in Experiment 1. PG conditions are the same as those shown in Table 2. Other types of islands are based upon materials used in previous studies of island constraints in parsing.

Unsurprisingly, the highest ratings were found in the four conditions that contained no gaps inside islands. This included the two 'good gap only' conditions derived from the items used in the current study, plus two conditions that contained simple one-clause or two-clause object wh-extraction. Among the conditions that contained gaps inside islands, the two conditions that received the highest ratings corresponded to the two environments that have provided evidence of gaps inside islands in the existing literature. This included the infinitival parasitic gap examples from the current study and examples of Complex NP Constraint violations derived from Freedman & Forster's (1985) sentence matching study. Examples of islands created by finite relative clauses, wh-islands and subject islands containing PPs all received lower ratings, and correspond to islands that previous studies have shown to resist active gap creation mechanisms. Interestingly, parasitic gap constructions derived from the subject islands in Stowe's (1986) study received almost identical ratings to parasitic gap constructions derived

---

[5] The examples of WH-island Constraint violations based on Kluender & Kutas (1993) and the examples of Complex NP Constraint violations based on Freedman & Forster (1985) were converted from direct to indirect questions, in order to allow a fairer comparison with the other items in this study.

[6] The examples of subject islands based on Stowe (1986) contained two gaps, one in the spurious gap position inside the subject NP, and another in the direct object position of a higher verb. Stowe's examples were not presented as parasitic gap constructions, but this is the most appropriate way to evaluate the acceptability of her subject islands, in light of the hypothesis that active gap creation mechanisms are sensitive to the availability of a parasitic gap construction. Following the same reasoning, the rating for the finite parasitic gap condition ('PG-FIN-both') is the most appropriate measure of the acceptability of the subject islands in Traxler & Pickering's (1996) study of subject islands.

from the finite subject islands in the current study. Both studies found no evidence for active gap creation inside those islands.
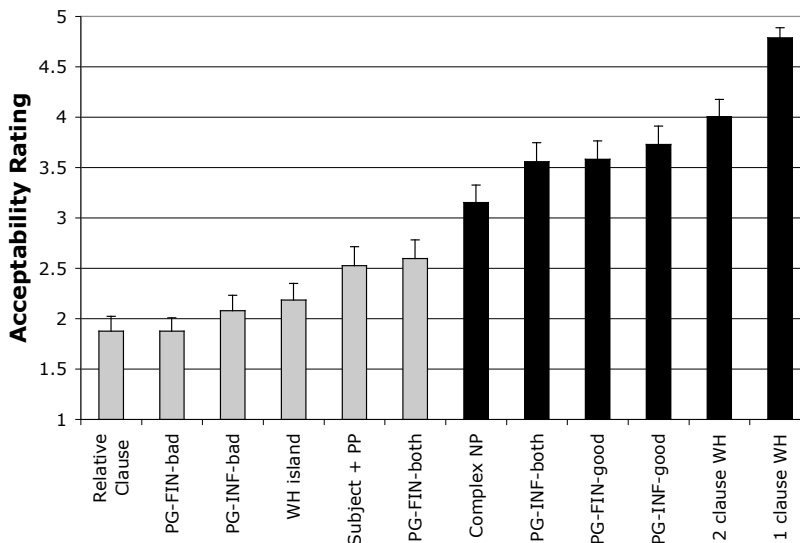

Figure 5: Acceptability ratings for different types of wh-extraction used in current and previous studies of islands in processing. Black bars correspond to syntactic contexts where active gap creation mechanisms have been argued to operate.

Although these ratings do not provide exhaustive coverage of the structures used in previous studies of island constraints in processing, the results suggest a connection between the degree of acceptability of an island environment and the use of active gap filling mechanisms in parsing. Importantly, for subject islands the best predictor of parsing behavior is the rating of the subject island as a part of a parasitic gap construction, and not the rating of the subject island violation when no licensing gap is available.

Overall, it seems clear that the parser avoids positing gaps in environments that are rated as highly unacceptable. On the other hand, an answer to the question of where the parser *does* actively posit gap positions must await more detailed research on Complex NP Constraint violations (also described as Specificity Constraint violations) as studied by Freedman & Forster (1985). Even if we set aside concerns about the sentence matching task raised by Crain & Fodor (1987) and Stowe (1992), this technique provides limited information about the time-course of structure building operations, and does not indicate whether active gap creation is operative in Complex NP islands.[7] Two possible generalizations are available, depending on the status of active gap creation in Complex NPs. If Complex NP islands do show active gap creation effects, then we are forced to concede that active gap creation does operate in some environments that are judged to be at least mildly unacceptable in off-line ratings. On the other hand, if Complex NP islands fail to show active gap creation effects, and if we take the results of Experiment 1 to show that infinitival parasitic gap constructions are fully well-formed, then we may conclude that

---

[7] Bourdages (1992) presents her Filled Gap Effect experiment as a study of Complex NP islands, but in fact the materials used in the experiment focus on a specific sub-class of complex NPs involving finite relative clauses.

the parser is more restrictive, and engages active gap creation mechanisms only in contexts that can yield fully acceptable wh-dependencies.

7.3. 'PROCESSING' ACCOUNTS OF ISLANDS. The current results present a challenge for attempts to derive island constraints from limitations of the language processing mechanisms (Berwick & Weinberg 1984, Deane 1991, Pritchett 1991, Kluender & Kutas 1993, Hawkins 1999). Experiment 1 showed that gaps inside infinitival subjects are rated as highly unacceptable when they are not rescued by an additional gap, and thus they qualify as genuine islands. Processing-based accounts of why subjects might be islands differ in a number of details (see §3 above), but share the straightforward prediction that the parser should be unable to construct a gap inside this type of island during real-time processing. However, the results of the on-line study in Experiment 2 showed that the parser is able to construct gaps inside infinitival subject islands fully incrementally. This casts doubt upon the notion that gaps are unacceptable inside subjects because of limitations in the parser's gap construction mechanisms. The unacceptability of gaps inside subjects must reflect some additional representational constraint, although the current results provide no reason to choose among the many available formal accounts of subject islands.

Note that this argument only applies to the subclass of islands that supports parasitic gaps, and thus does not exclude the possibility that other types of islands are appropriately explained in terms of processing limitations. For example, the current results are consistent with the possibility that the relative clause islands in the finite conditions of Experiment 2 are explained by limitations on real-time parsing mechanisms, and this may also be true for other islands that are consistently respected during parsing. However, taken to its logical conclusion, this approach implies that milder islands reflect formal or representational constraints but more severe islands are epiphenomena of parsing. This is a surprising conclusion, since it would suggest that the most severe instances of island violations are grammatically well-formed and only certain milder violations are formally illicit. An alternative approach that I consider to be more promising is one in which processing difficulty is not in itself sufficient to create a syntactic island, but may nevertheless be responsible for explaining some of the variation in the acceptability of different island types.

7.4. ARCHITECTURAL IMPLICATIONS. The parasitic gap constructions discussed here are also of interest for attempts to bring together theories of grammatical knowledge and theories of real-time structure building (e.g. Bresnan 1978, Phillips 1996, 2003, Steedman 2000, Kempson, Meyer-Viol & Gabbay 2001, Kempen & Harbusch 2002). These approaches assume a close link between what is acceptable and what is constructed in real time, and only allow for inaccuracy in real-time processes in cases of ambiguity or memory overload. Parasitic gaps inside subject islands present a challenge for such approaches, since the ordering of the illicit gap and its licensor appears to create a look-ahead problem for real-time structure building. As we have seen, however, even in this case the parser is able to preserve incrementality and accuracy. Such cases do not, of course, show that it is necessary to view the grammar as a real-time structure-building system. It is always possible to maintain the traditional distinction between a parser and a producer that operate in real-time and a separate grammatical system that does not. However, the more we find that real-time processes capture fine-grained grammatical distinctions, the less there is a need for an additional system that recapitulates such distinctions in a time-independent fashion.

8. CONCLUSIONS. Parasitic gap constructions are probably rather infrequent in normal discourse, but they do not deserve their reputation as linguistically 'marginal' phenomena.[8] This article showed two ways in which these constructions are treated quite normally. The off-line study showed that parasitic gap constructions are genuinely acceptable: normal speakers of English rate parasitic gap constructions just as highly as closely matched sentences that lack a parasitic gap. The on-line study showed that the normal mechanisms of active gap creation apply in parasitic gap environments, even when the illicit gap appears before a licensing gap has been confirmed. A comparison of different types of subject islands showed that the parser actively creates gaps inside only those subject islands that can support parasitic gaps. This finding has a number of implications. First, for theories of language processing it indicates that the parser accurately and incrementally implements the grammar of parasitic gaps, despite the rarity of these constructions. The contrast between the processing of two types of subject islands may help to explain variability in previous findings about island constraints in parsing. Second, for theories of island constraints the parser's ability to posit gaps inside some types of islands makes it unlikely that those island phenomena are reducible to limitations in language processing, and suggest the need for a representational account of those islands. More generally, the current findings provide further evidence that real-time structure building mechanisms show substantial grammatical precision, even in cases where surface word order would appear to make this rather difficult.

References

AOSHIMA, SACHIKO; COLIN PHILLIPS; and AMY S. WEINBERG. 2004. Processing filler-gap dependencies in a head-final language. Journal of Memory and Language 51.23-54.

BERWICK, ROBERT, and AMY S. WEINBERG. 1984. The Grammatical Basis of Linguistic Performance. Cambridge, Mass.: MIT Press.

BOLAND, JULIE E., MICHAEL K. TANENHAUS, SUSAN M. GARNSEY, & GREGORY N. CARLSON. 1995. Verb-argument structure in parsing and interpretation: Evidence from wh-questions. Journal of Memory and Language 34.774-806.

BOURDAGES, JOHANNE S. 1992. Parsing complex NPs in French. Island Constraints: Theory, Acquisition and Processing, ed. By Helen Goodluck and Michael S. Rochemont, 61-87. The Netherlands: Kluwer Academic Publishers.

BRESNAN, JOAN. 1978. A realistic transformational grammar. Linguistic Theory and Psychological Reality, ed. by Joan Bresnan, Morris Halle, and George Miller, 1-59. Cambridge, Mass.: MIT Press.

CHOMSKY, NOAM. 1973. Conditions on transformations. A Festschrift for Morris Halle, ed. by Stephen Anderson and Paul Kiparsky, 232-286. New York: Holt, Rinehart & Winston.

CINQUE, GUGLIELMO. 1990. Types of A' Dependencies. Cambridge, Mass.; MIT Press.

CLIFTON, CHARLES E., JR., and LYN FRAZIER. 1989. Comprehending sentences with long-distance dependencies. Linguistic structure in language processing, ed. by Michael K. Tanenhaus and Gregory N. Carlson, 273-317. Dordrecht, The Netherlands: Kluwer Academic Publishers.

CRAIN, STEPHEN & JANET DEAN FODOR. 1985. How can grammars help parsers? Natural language parsing: psycholinguistic, computational, and theoretical perspectives, ed. by David Dowty,

---

[8] An internet search for the expressions 'marginal' and 'parasitic gaps' yields a surprisingly large set of matches.

Lauri Kartunnen, and Arnold M. Zwicky, 94-128. Cambridge, UK: Cambridge University Press.

CRAIN, STEPHEN, and JANET DEAN FODOR. 1987. Sentence matching and overgeneration. Cognition 26.123-169.

CULICOVER, PETER. 2001. Parasitic gaps: A history. Parasitic Gaps, ed. by Peter S. Culicover and Paul Postal, 3-68. Cambridge, Mass.: MIT Press.

CULICOVER, PETER W., and PAUL POSTAL, editors. 2001. Parasitic Gaps. Cambridge, Mass.: MIT Press.

DEANE, PAUL. 1991. Limits to attention: A cognitive theory of island constraints. Cognitive Linguistics 2.1-63.

FELSER, CLAUDIA; HARALD CLAHSEN; and THOMAS F. MÜNTE. 2003. Storage and integration in the processing of filler-gap dependencies: An ERP study of topicalization and wh-movement in German. Brain and Language 87.345-354.

FERREIRA, FERNANDA, and CHARLES E. CLIFTON, JR. 1986. The independence of syntactic processing. Journal of Memory and Language 25.348-368.

FODOR, JANET DEAN. 1978. Parsing strategies and constraints on transformations. Linguistic Inquiry 9.427-473.

FODOR, JANET DEAN. 1983. Phrase structure parsing and the island constraints. Linguistics and Philosophy 6.163-223.

FRAZIER, LYN, and GIOVANNI B. FLORES D'ARCAIS. 1989. Filler-driven parsing: a study of gap filling in Dutch. Journal of Memory and Language 28.331-344.

FREEDMAN, SANDRA E., and KENNETH I. FORSTER. 1985. The psychological status of overgenerated sentences. Cognition 19.101-131.

GARNSEY, SUSAN M.; MICHAEL K. TANENHAUS; and ROBERT M. CHAPMAN. 1989. Evoked potentials and the study of sentence comprehension. Journal of Psycholinguistic Research 18.51-60.

GIBSON, EDWARD, and GREGORY HICKOK. 1993. Sentence processing with empty categories. Language and Cognitive Processes 8.147-161.

GOODLUCK, HELEN; MALCOLM FINNEY; and JULIE SEDIVY. 1991. Sentence completeness and filler-gap dependency parsing. OTS Yearbook 1991, ed. by Peter Coopmans, Bert Schouten, and Wim Zonneveld, 19-31. Utrecht: University of Utrecht Press.

HAWKINS, JOHN. 1999. Processing complexity and filler-gap dependencies across languages. Language 75.224-285.

JACKENDOFF, RAY, and PETER W. CULICOVER. 1971. A reconsideration of dative movement. Foundations of Language 7.392-412.

JUST, MARCEL ADAM; PATRICIA A. CARPENTER; and JACQUELINE D. WOOLLEY. 1982. Paradigms and processes in reading comprehension. Journal of Experimental Psychology: General 11.228-238.

KAAN, EDITH; ANTHONY HARRIS; EDWARD GIBSON; and PHILIP HOLCOMB. 2000. The P600 as an index of syntactic integration difficulty. Language and Cognitive Processes 15.159-201.

KAYNE, RICHARD S. 1983. Connectedness. Linguistic Inquiry 14.223-249.

KEMPEN, GERARD, and KARIN HARBUSCH. 2002. Performance grammar: A declarative definition. Computational Linguistics in the Netherlands 2001, ed. by Anton Nijholt, Mariët Theune, and Hendri Hondorp, 148-162. Amsterdam: Rodopi.

KEMPSON, RUTH; WILFRIED MEYER-VIOL; and DOV GABBAY. 2001. Dynamic Syntax. Oxford: Blackwell.

KLUENDER, ROBERT, and MARTA KUTAS. 1993. Subjacency as a processing phenomenon. Language and Cognitive Processes 8.573-633.

KURTZMAN, HOWARD S., and LOREN FROST CRAWFORD. 1991. Processing parasitic gaps. Proceedings of the 21ˢᵗ Annual Meeting of the North East Linguistics Society, ed. by Tim Sherer. 217-231. Amherst, Mass.: GLSA Publications.

LEE, MING-WEI. 2004. Another look at the role of empty categories in sentence processing (and grammar). Journal of Psycholinguistic Research 33.51-73.

MANZINI, MARIA RITA. 1992. Locality. Cambridge, Mass.: MIT Press.

MANZINI, MARIA RITA. 1994. Locality, minimalism, and parasitic gaps. Linguistic Inquiry 25.481-508.

MCKOON, GAIL; ROGER RATCLIFF; and GREGORY WARD. 1994. Testing theories of language comprehension: An investigation on the on-line lexical decision task. Journal of Experimental Psychology: Learning, Memory, and Cognition 20.1219-1228.

MCELREE, BRIAN, and TERESA GRIFFITH. 1998. Structural and lexical constraints on filling gaps during sentence comprehension: A time-course analysis. Journal of Experimental Psychology: Learning, Memory, and Cognition 24.432-460.

MCKINNON, RICHARD, and LEE OSTERHOUT. 1996. Constraints on movement phenomena in sentence processing: Evidence from event-related potentials. Language and Cognitive Processes 11.495-523.

NEVILLE, HELEN; JANET NICOL; ANDREW BARSS; KENNETH FORSTER; and MERRILL GARRETT. 1991. Syntactically-based sentence processing classes: Evidence from event-related brain potentials. Journal of Cognitive Neuroscience 3.151-165.

NICOL, JANET L.; JANET DEAN FODOR; and DAVID SWINNEY. 1994. Using cross-modal lexical decision tasks to investigate sentence processing. Journal of Experimental Psychology: Learning, Memory, and Cognition 20.1229-1238.

NICOL, JANET L., & DAVID SWINNEY. 1989. The role of structure in coreference assignment during sentence comprehension. Journal of Psycholinguistic Research 18.5-20.

PHILLIPS, COLIN. 1996. Order and Structure. Unpublished PhD dissertation, Massachusetts Institute of Technology.

PHILLIPS, COLIN. 2003. Linear order and constituency. Linguistic Inquiry 34.37-90.

PHILLIPS, COLIN; NINA KAZANINA; and SHANI H. ABADA. 2004. ERP effects of the processing of syntactic long-distance dependencies. In press, Cognitive Brain Research.

PICKERING, MARTIN J., and GUY D. BARRY. 1991. Sentence processing without empty categories. Language and Cognitive Processes 6.229-259.

PICKERING, MARTIN J.; STEPHEN BARTON; and RICHARD SHILLCOCK. 1994. Unbounded dependencies, island constraints and processing complexity. Perspectives on Sentence Processing, ed. by Charles Clifton, Jr., Lyn Frazier, and Keith Rayner, 199-224. London: Lawrence Erlbaum.

PRITCHETT, BRADLEY L. 1991. Subjacency in a principle-based parser. Principle-based parsing: computation and psycholinguistics, ed. by Robert C. Berwick, 301-345. Dordrecht, The Netherlands: Kluwer Academic Publishers.

ROSS, JOHN R. 1967. Constraints on Variables in Syntax. Ph.D. dissertation, MIT.

SAG, IVAN A., and JANET DEAN FODOR. 1994. Extraction without traces. Proceedings of the 13th Annual Meeting of the West Coast Conference on Formal Linguistics, ed. by Raúl Aranovich, William Byrne, Susanne Preuss, and Martha Senturia, 365-384. Stanford: CSLI Publications.

STEEDMAN, MARK. 2000. The Syntactic Process. Cambridge, Mass.: MIT Press.

STOWE, LAURIE A. 1986. Evidence for on-line gap-location. Language and Cognitive Processes 1.227-245.

STOWE, LAURIE A. 1992. The processing implementation of syntactic constraints: The sentence matching debate. Island Constraints: Theory, Acquisition and Processing, ed. by Helen Goodluck and Michael Rochemont, 419-443. Dordrecht: Kluwer Academic Publishers.

SUSSMAN, RACHEL S., and JULIE C. SEDIVY. 2003. The time-course of processing syntactic dependencies: evidence from eye-movements during spoken wh-questions. Language and Cognitive Processes 18.143-163.

TANENHAUS, MICHAEL K.; GREGORY CARLSON; and JOHN C. TRUESWELL. 1989. The role of thematic structures in interpretation and parsing. Language and Cognitive Processes 4.211-234.

TRAXLER, MATTHEW J., and MARTIN J. PICKERING. 1996. Plausibility and the processing of unbounded dependencies: An eye-tracking study. Journal of Memory and Language 35.454-475.

TRUESWELL, JOHN C., and MICHAEL K. TANENHAUS. 1991. Tense, temporal context, and syntactic ambiguity resolution. Language and Cognitive Processes 6.303-338.

WANNER, ERIC, and MICHAEL MARATSOS. 1978. An ATN approach to comprehension. Linguistic Theory and Psychological Reality, ed. By Morris Halle, Joan Bresnan and George A. Miller, 119-161. Cambridge, Mass.: MIT Press.

YOSHIDA, MASAYA; SACHIKO AOSHIMA; and COLIN PHILLIPS. 2004. Relative clause prediction in Japanese. Talk presented at the 17th Annual CUNY Conference on Human Sentence Processing, College Park, Maryland.

Colin Phillips
Department of Linguistics
1401 Marie Mount Hall
University of Maryland

colin@umd.edu
http://www.ling.umd.edu/colin