An expectation-based account of subject islands and parasitism[1]

RUI P. CHAVES

Linguistics Department – University at Buffalo
The State University of New York

Subject phrases impose particularly strong constraints on extraction. Most research assumes a syntactic account (e.g. Kayne 1983, Chomsky 1986, Rizzi 1990, Lasnik & Saito 1992, Takahashi 1994, Uriagereka 1999), but there are also pragmatic accounts (Erteschik-Shir & Lappin 1979; Van Valin 1986, 1995; Erteschik-Shir 2006, 2007) as well as performance-based approaches (Kluender 2004). In this work I argue that none of these accounts captures the full range of empirical facts, and show that subject and adjunct phrases (phrasal or clausal, finite or otherwise) are by no means impermeable to non-parasitic extraction of nominal, prepositional and adverbial phrases. The present empirical reassessment indicates that the phenomena involving subject and adjunct islands defies the formulation of a general grammatical account. Drawing from insights by Engdahl (1983) and Kluender (2004), I argue that subject island effects have a functional explanation. Independently motivated pragmatic and processing limitations cause subject-internal gaps to be heavily dispreferred, and therefore, extremely infrequent. In turn, this has led to heuristic parsing expectations that preempt subject-internal gaps and therefore speed up processing by pruning the search space of filler–gap dependencies. Such expectations cause processing problems when violated, unless they are dampened by prosodic and pragmatic cues that boost the construction of the correct parse. This account predicts subject islands and their (non-)parasitic exceptions.

1. INTRODUCTION

Since the seminal work of Ross (1967), extraction islands have proved to be elusive for generative grammar. It remains unclear why islands exist in the first place, and there is yet no complete and precise characterization of how islands operate. The latter is challenging because virtually every island ever

---

proposed has known counterexamples. The result is a complex landscape of grammatical patterns, which are difficult to characterize in empirical terms, let alone capture in theoretical grammars.

There are three major schools of thought. The mainstream view holds that human grammars have structural – and sometimes semantic – conditions governing which kinds of extraction may or not occur (for overviews see e.g. Szabolcsi & den Dikken 1999, Hornstein, Lasnik & Uriagereka 2006, Abrúsan 2007, Boeckx 2008, Truswell 2011). A second view is that islands are a pragmatic epiphenomenon (Erteschik-Shir 1973, Erteschik-Shir & Lappin 1979, Kuno 1987, Van Valin 1995), and finally, a third view argues that at least some islands may be best explained as the cumulative effect of cognitive limitations (Givón 1979; Grosu 1981; Deane 1991; Kluender 1992, 1998, 2004; Kluender & Kutas 1993; Hofmeister 2007; Sag, Hofmeister & Snider 2007; Hofmeister & Sag 2010). The latter has recently regained considerable interest because it has psycholinguistic support. If successful, such an approach would explain the existence of (at least) some island effects and their circumvention in terms of independently motivated cognitive factors. The existence of graded or even acceptable island counterexamples would correspond to cases where the particular choice of words does not exhaust the cognitive resources available to the processor and does not mislead the processor in finding the correct parse (e.g. sentences composed of phrases that cohere particularly well and do not induce garden-paths). As a result, we would obtain a simpler grammatical theory of extraction in which at least some islands result from performance aspects. The matter of which account is on the right track remains unresolved, and the debate between these views remains lively, as seen in Erteschik-Shir (2007), Sprouse, Wagers & Phillips (2012a, b), and Hofmeister, Staum Casasanto & Sag (2012a, b, in press).[2]

The goal of the present work is twofold. First, it aims to draw attention to the fact that subject and adjunct islands can – in certain conditions – be robustly circumvented without the use of so-called 'parasitic' gaps. In particular, it will be shown that none of the syntactic environments that have been claimed to prohibit extraction are impermeable to extraction, including those that have been subject to experimental study such as Phillips (2006), and others cited therein. In the present work, graded and passable extractions from subject islands are not taken as uninteresting marginal cases. On the contrary, they reveal patterns which suggest that modern syntactic theory has overstated the role that configurational conditions plays in island effects.

---

[2] Different islands are probably best seen as resulting from different blends of syntactic, semantic, pragmatic, and cognitive factors. For example, whereas the ELEMENT CONSTRAINT and its exceptions can be seen as predominantly semantic-pragmatic phenomena (Kehler 2002, Chaves 2012), other conditions, like the CONJUNCT CONSTRAINT, can be explained in more syntactic terms (Bouma, Malouf & Sag 2001, Chaves 2012).

The second goal is to re-interpret the empirical data so that a coherent explanation can be offered for the full range of facts: subject island effects, parasitism, and the (sometimes graded) circumvention of non-parasitic subject islands.

The structure of this paper is as follows. Section 2 offers an empirical overview of strong islands in English, and a brief survey of the literature. The picture that emerges is one where subject and adjunct islands cannot be reduced to one and the same grammatical condition, and where there is no clear motivation to distinguish 'parasitic' gaps from 'real' gaps in the grammar. Section 3 focuses on subject islands, and argues that previous research has focused on an unrepresentative set of data. In fact, there seem to exist several classes of acceptable subject island violations which have been overlooked by mainstream research, and which suggest that such islands are not due to a uniform grammatical condition. Section 4 proposes a new interpretation of the data in which the extraction from subjects is grammatical, but hampered by heuristic parsing expectations. Such expectations can, however, be weakened by grammatical cues that aid the parser identify the correct location of the subject-internal gap. This work proposes a general and parsimonious explanation of subject islands and their (non-)parasitic circumvention.

2. STRONG ISLANDS

2.1 Background

The A-OVER-A CONSTRAINT proposed by Chomsky (1962) had the effect that any category *X* could not be extracted across a node of the same category. For example, this prevented NP extraction from NPs as in *What did he know* [*someone who has __*]? and in *Who did you hear* [*the rumor* [*that Mary kissed __*]]?[3] But it was soon observed that such a constraint was too strong as it ruled out perfectly acceptable extractions like *Who did you read* [*a book about __*]? Subsequent accounts like Chomsky (1973, 1986) were more permissive, allowing (1a) while ruling out (1b).

(1) (a) Who did you take a photograph of __?
    (b) *??Who did you take a photograph of a statue of __?

(Fodor 1983: 190)

In the Barriers framework (Chomsky 1986), island effects were considered additive in proportion to the number of 'cyclic blocking categories' that were crossed (in English, NP and IP nodes). This approach also

---

[3] The diacritic '*' is used to indicate that a string is unacceptable. Judgments are always about acceptability, not grammaticality, and the distinction between the two is not easily drawn. The main contention of this paper is that subject island violations are grammatical, although their acceptability is often graded.

blocked the extraction of NPs from subject phrases as in (2), because under transformational assumptions extraction crosses two blocking categories in such cases, which 'should yield a considerable decrement in acceptability' (Chomsky 1986: 86).

(2) (a) Your interest in him seemed to me rather strange.
    (b) *Whom did your interest in seem to me rather strange?

(Chomsky 1962: 437)

    (c) The hoods of these cars were damaged by the explosion.
    (d) *Which cars were the hoods of damaged by the explosion?

(Ross 1967, ex. (4.252))

There are a variety of problems with this kind of account. First, it has never been clear why such 'cyclic bounding nodes' should interfere with extraction in the first place. Second, this view assumes that there are various degrees of grammaticality. Not only is it unclear what this means for a derivational grammatical theory (where, presumably, a grammatical violation would simply abort the parse) but it is also empirically hard to justify since all speaker judgments are necessarily acceptability judgments rather than grammaticality judgments. And if (1b) is odd due to performance rather than grammar, then the grammar should remain silent about extraction constraints on bounding nodes. For Fodor (1983) and others, the fact that multiple embeddings often make object extraction more difficult, as in (1), suggests an explanation based on processing complexity.

   A third problem was noted by Ross (1967) and Deane (1991). The existence of deep extraction cases like (3) suggests that there is no syntactic depth limitation to NP extraction from object NPs.[4] To my knowledge, these facts have never been taken into consideration in the formulation of subjacency-based accounts of islands.

(3) (a) Nixon was one president that [they had [no trouble [getting votes for [the impeachment of __]]]].
    (b) Which committee did you have [aspirations for [an appointment to [the chairmanship of __]]]?

(Deane 1991: 11)

---

[4] Ross's original data are shown in (i) and (ii), and are less natural than (3). Still, Chomsky (1973) and others have acknowledged the acceptability of (i) and (ii).

  (i) Which reports does the government prescribe the height of the lettering on?
 (ii) Reports which the government prescribes the height of the lettering on the covers of are invariably boring.

Deaene (1991) and other authors (Chung & McCloskey 1983, Hegarty 1990, and Pollard & Sag 1994: 191, 206) noted further counterexamples like (4), which show that complex NPs are not impermeable to extraction.

(4) (a) How much money are you making [the claim [that the company squandered]]?
(b) Which rebel leader did you hear [rumors [that the CIA assassinated]]?
(c) Which Middle Eastern country did you hear [rumors [that we had infiltrated]]?
(d) Which diamond ring did you say there was [nobody in the world [who could buy]]?

The evidence above suggests that the oddness of (1b) may not result from purely configurational factors. A promising avenue of inquiry should consider semantic, pragmatic, and performance factors as well. As an analogy, it would be misguided to postulate a ban on NP extraction from multiple embedded clauses in order to capture the oddness of ?*_Who did you say Tom thinks Mary denied Tom believes Fred trusts?_. Although humans routinely process longer sentences, the oddness of such examples cannot be seen as a matter of competence.

Matters are further complicated by the fact that extraction out of subject phrases is generally held to be impossible and therefore much more restricted than extraction from objects, as the sample in (5) shows. The judgments reported below are the cited author's.[5]

(5) (a) *Who did stories about terrify John?   (Chomsky 1977: 106)
(b) ??Which tree did John see the leaves of turn color?

(Kuno 1973a: 381)
(c) *Who did pictures of lay on the table?   (Postal 1974: 189)
(d) *Who was a picture of lying there?   (Kayne 1981: 114)
(e) *Who do you think pictures of would please John?

(Huang 1982: 497)
(f) ?Who did John's talking to bother you most?   (Engdahl 1983: 14)
(g) *Which books did talking about become difficult?   (Cinque 1990: 1)
(h) *Who does the claim that Mary likes upset Bill?

(Lasnik & Saito 1992: 42)
(i) *Which politician did pictures of upset the voters?

(Nunes & Uriagereka 2000: 21)
(j) *Which book did a review of appear in the Times?

(Jackendoff 2002: 48)

---

[5] Note that the degree of acceptability varies, and that examples like (5i–k) are actually passable rather than unacceptable, which undermines the predominant view that holds that subjects are strong islands in English. Section 3 offers an explanation for the source of gradient acceptability.

(k) *Which candidate were posters of all over the town?

<div align="right">(Lasnik & Park 2003: 651)</div>

(l) ?*Who was a friend of arrested? (Stepanov 2007: 85)

The puzzle is not only to explain why extraction from objects is possible in some cases but not in others, but to explain why extraction from subjects is more restrictive. Standard accounts categorically prohibit movement from the subject phrase. In Takahashi (1994) such islands are the consequence of the requirement that movement must be as short as possible and the stipulation that nothing can adjoin to a chain (subjects move from [Spec,vP] to [Spec,IP], thus creating an A-chain with the in situ copy of the subject). Any extraction from a subject would be required to adjoin to the closest possible site, which is the subject DP, hence violating the ban on chain adjunction. Conversely, a longer movement is banned by the closest move condition. As a consequence, the subject island effect arises. In Uriagereka (1999) and Nunes & Uriagereka (2000), phrases are assumed to be linearized before they can be merged with other phrases. Thus, subjects must be linearized before combining with the V′. Once linearized, the DP is considered a morphological unit, and hence, no extraction from it can occur. Syntactic accounts like the ones above predict that extraction from an extracted phrase is banned in general, as a FREEZING effect (Wexler & Culicover 1980). A conclusion that may be excessive, as the acceptable FREEZING violations in (6) suggest.

(6) (a) This is the handout$_i$ that I can't remember [how many copies of __ $_i$]$_j$ we have to print __ $_j$.
   (b) [How many videos __ $_i$]$_j$ are there __ $_j$ on the web [of Mitt Romney getting booed]$_i$?

A competing line of research holds that subject islands are due to pragmatic factors, such as Erteschik-Shir (1973, 2006, 2007), Van Valin (1986, 1995), and Goldberg (2006) *inter alia*. Basically, the claim is that subjects (and their adnominal structures) are typically part of the common ground and therefore cannot be felicitously questioned. For example, Erteschik-Shir (2006: 325) states that extraction out of subjects is always ungrammatical and cannot be contextually ameliorated. This categorical prohibition follows from her assumption that subjects must be assigned TOPIC and that extraction never targets topics. Hence, Erteschik-Shir obtains the prediction that such extractions are always blocked.

Adjunct phrases are another type of phrase that is usually regarded as being particularly difficult to extract from, as (7) illustrates.

(7) (a) John met a lot of girls without going to the club.
   (b) *Which club did John meet a lot of girls [without going to __]? (Cattell 1976: 38)
   (c) Who cried after John hit Mary?
   (d) *Who did Mary cry [after John hit __]? (Huang 1982: 503)

Huang (1982) argued that subject and adjunct islands ought to be unified under a broader universal condition called the CONSTRAINT ON EXTRACTION DOMAIN (CED). This rule stipulates that a phrase *P* may be extracted out of domain *D*, iff *D* is properly governed. An α is said to be properly governed by β if α is c-commanded by a lexical head β and no major category or major category boundary appears between α and β. Many authors have sought to derive the CED from independent principles (Kayne 1983, Chomsky 1986, Rizzi 1990, Lasnik & Saito 1992, Uriagereka 1999, Sabel 2002, and Müller 2010).

2.2 *Counterindications*

There are good reasons to doubt that subject islands are so intimately related to adjunct islands, and more importantly, to doubt that such islands are simply due to grammatical conditions. The first problem for the grammatical status of strong islands comes from the observation made by Ross (1967), Taraldsen (1980) and Engdahl (1983) that extraction from subjects and adjuncts is usually more acceptable if the gap is 'parasitic' on another gap, located in a non-island environment, as in (8)–(10).

(8) (a) *Who did [the rivals of __] shoot Castro?
     (b) Who did [the rivals of __] shoot __?
     (c) Who did [the rivals of Castro] shoot __?

(9) (a) *What did [the attempt to repair __] ultimately damage the car?
     (b) What did [the attempt to repair __] ultimately damage __?
     (c) What did [the attempt to repair the car] ultimately damage __?

(10) (a) *Which paper did John [read his email [before filing __]]?
     (b) Which paper did John [read__[before filing __]]?
     (c) Which paper did John [read__[before filing a complaint]]?

Instead of taking (8b), (9b) and (10b) as counterexamples, these island-violating gaps have been seen as some kind of null resumptive pronoun. As such, they would pose no challenge to the view that subjects and adjuncts are strong extraction islands. For example, Cinque (1990) and Postal (1993, 1994, 1998) draw a distinction between 'A-extractions' and 'B-extractions'. While the former are canonical *wh*-extractions, the latter arise by insertion of a null resumptive pronoun at the extraction site. In both accounts, B-extractions are identified as parasitic extractions.

But, as Steedman (1996: 98, fn.41), Levine (2001), Levine, Hukari & Calcagno (2001), and Levine & Hukari (2006: 256) show, parasitic gaps can be non-nominal, as illustrated in (11).

(11) (a) [How harshly] do you think we can treat them __ without in turn
          being treated __ ourselves?

  (b) I wonder [how nasty] you can pretend to be __ without actually becoming __?
  (c) This is a cause [to which] many people are attracted __ without ever becoming seriously devoted __.
  (d) This is the table [on which] anyone who puts some books __ must subsequently put some magazines __ as well.

This casts doubts on the assumption that parasitic extractions can be explained away in terms of null resumptive pronouns. Furthermore, the notion that only one gap is 'real' and all others are null pronouns is refuted by cases where neither gap occurs in a parasitic environment, as in *Who did you inform __ that the police was about to arrest __?* or in *Who did you hug __ and kiss __?*.

   The view that adjuncts are islands and that the only gaps that such phrases allow are of the 'parasitic' type is problematic. There is robust evidence that adjuncts do in fact allow extraction, as illustrated in (12), from Ross (1967), Chomsky (1982), Engdahl (1983), Kayne (1983), Hegarty (1990), Cinque (1990), Pollard & Sag (1994), Borgonovo & Neeleman (2000).

(12) (a) That's the symphony that Schubert [died [without finishing _]].
     (b) Which report did Kim [go to lunch [without reading _]]?
     (c) What did he [fall asleep [complaining about _]]?
     (d) What did John [drive Mary crazy [trying to fix _]]?
     (e) What did John [come back [addicted to _]]?
     (f) Who did you [go to Girona [in order to meet _]]?
     (g) Who would you rather [sing [with _]]?
     (h) What temperature should I [wash my jeans [at _]]?

   See Szabolcsi (2006) and Truswell (2011) for overviews of adjunct islands and their exceptions. To my knowledge, there is no direct empirical evidence that can justify the assumption that the examples in (12) are grammatical because the adjuncts have some special syntactic status that does not lead to an island violation. The natural conclusion is that adjunct islands are probably not a syntactic phenomenon.

   Although untensed adjuncts do not always block extraction, it has been nearly universally held that tensed adjuncts, as in (13), are truly exceptionless islands.

(13) (a) John came back before I had a chance to talk to you.
     (b) *Who did John [come back [before I had a chance to talk to _]]?
                                                      (Huang 1982 : 491)

But a small number of authors have occasionally noted that it is in principle possible to extract from finite adjuncts. The first cases were noted

by Grosu (1981: 88). The examples in (14) are adapted from Grosu's original data.[6]

(14) (a) These are the pills that Mary died before she could take.
    (b) These is the house that Mary died before she sold.

Taylor (2007) claims that extraction from tensed conditional adjuncts is allowed as long as the adverbial clause is fronted, as in (15a) (her judgments).

(15) (a) Which book does John believe if Kim reads, she will understand linguistics better?
    (b) *Which book will Kim understand linguistics better if she reads?
<div align="right">(Taylor 2007: 192)</div>

Taylor (2007) argues that such fronted conditionals allow extraction because they are base-generated, but (14) and (16) refute the generality of this view.

(16) (a) This is the watch that I got upset when I lost.
<div align="right">(attributed to Ivan Sag (p.c.) by Truswell 2011: 175, fn. 1)</div>
    (b) Robin, Pat and Terry were the people who I hung around at home all day without realizing were coming for dinner.
<div align="right">(Levine & Hukari 2006: 287)</div>

Further counterexamples are provided in (17), drawn from Chaves (2012: 471).

(17) (a) Which email account would you be in trouble if someone broke into?
    (b) Which problem would you be impressed if someone had already solved?
    (c) Which costume would mom freak out the most if I wore on Halloween?
    (d) This is the type of problem that I would be relieved if someone had already solved.
    (e) Which toy did Tom throw a tantrum because somebody broke?

Some of these sentences may be complex and difficult to process, which in turn can lead speakers to prefer the insertion of an 'intrusive' resumptive pronoun at the gap site, but they are more acceptable than the classic examples in (7) and (13) above.

---

[6] Grosu's original datum was *There are a number of important people WHO Mary died before she could meet*. The possibility of tensed 'reason' adjuncts not being strong islands is also raised independently by Kluender (1998) as well as by Goldberg (2006: 144), given the graded acceptability of *Who did she travel to Memphis because she wanted to see?*.

The data discussed so far suggest that adjunct islands are not absolute and have acceptable non-parasitic violations, contrary to widespread belief. The problems for the standard notion of parasitism are underlined by the fact that the 'parasitic' gap need not be semantically linked to the 'real' gap. In (18) – attributed to Polly Jacobson (p.c.) in Pollard & Sag (1994: 199, fn. 35) – the subject-internal gap is not co-referenced with the object gap.

(18) (a) There are certain heroes that [long stories about __] are always very easy to listen to __.
  (b) There are certain heroes that [long stories about __] are too boring to listen to __.

If the two gaps are not related, then in what grammatical sense is one licensing the other? Pollard & Sag (1994: 199) note other cases, such as (19), in which the subject-internal gap is not linked to the object-internal gap. Although the extraction of the phrase *which topics* causes low acceptability in (19a), the existence of a second gap in (19b) improves the sentence. Again, it is most unclear in what grammatical sense a 'parasitic' gap must be licensed by a 'real' gap.

(19) (a) ??I never know [which topics]$_j$ [jokes about __ $_j$] are likely to offend people.
  (b) [People that sensitive]$_i$, I never know [which topics]$_j$ [jokes about __ $_j$] are likely to offend __ $_i$.

This pattern also occurs in adjunct gaps as shown by Chaves (2012: 471). In (20) the gap located in the adjunct phrase is not semantically linked to the preceding gap, and hence would not be expected to be parasitic.

(20) (a) [A project this complex]$_i$, [how much time]$_j$ could they spend __ $_j$ before finishing __ $_i$?
  (b) This was [the kind of person]$_i$ that even [the simplest problem]$_j$ became difficult to solve __ $_j$ without shouting at __ $_i$.

The standard view on parasitic gap licensing is also challenged by acceptable cases noted by Horvath (1992:201), in which a parasitic gap is c-commanded by co-indexed non-parasitic gap (e.g. *Who do you expect __ to withdraw his candidacy* [*before the Committee has a chance to interview __]?*). See Culicover & Postal (2001) for more discussion.

Finally, Levine & Sag (2003) and Levine & Hukari (2006: 256) note what they call 'symbiotic gaps'. In such cases, one gap is located in the subject and the other in the adjunct, as shown in (21).

(21) (a) *What kinds of books do authors of __ argue about royalties after writing malicious pamphlets?
  (b) ??What kinds of books do authors of malicious pamphlets argue about royalties after writing __?

294

    (c) What kinds of books do the authors of ＿ argue about royalties after writing ＿ ?

The crucial point of these data is that (21a, b) are less acceptable than (21c). This is unexpected, given that the sentence arguably has no legitimate non-parasitic gap. Although (21c) is by no means impeccable, it should be at least as bad as (21a, b), contrary to fact.[7] Given the evidence discussed above, I follow Levine et al. (2001) and Levine & Hukari (2006: 292) in assuming that there is no grammatical distinction between 'parasitic' and 'real' gaps. Furthermore, I conclude that any account of parasitism based on stipulations about c-command, if any exists, would shed no light on the phenomena.

The claim that the CED is a universal syntactic constraint is also problematic. It is known since Ross (1967) that there is a variety of languages that seem to allow extraction from subjects. According to the survey by Stepanov (2007), these include Swedish, German, Russian, Spanish Palauan, Malagasy, Tagalog, Armenian, Hungarian, Turkish, Hindi, Japanese and Navajo. See also Chung (1991: 71) for acceptable sentential subject island violations in Chamorro. For illustration, consider the Swedish example in (22) from Sells (1984: 304ff.).

(22) den deckare   som de sista sidorna i＿ hade kommit bort
      that detective.novel that the last pages in had come   away
      'that detective novel that [the last pages in had come away] … '

Because there are many languages that allow extraction from subjects, but not from adjuncts, Stepanov (2007) concludes that the subject and adjunct islands are due to different mechanisms. In Stepanov's view, English just happens to be one of those languages where both subjects and adjuncts are opaque to extraction.

In spite of this robust typological evidence, there is still resistance in abandoning the universality of the CED, as illustrated in Hornstein et al. (2006), Boeckx (2008: 161), and Sheehan (2009). For instance, Jurka, Nakao & Omaki (2011) conducted acceptability studies which show that extraction out of specifiers in German, English, Japanese and Serbian are consistently less acceptable than extraction out of complements. The authors of the study take this to show the CED is a valid constraint and that syntactic accounts like Uriagereka (1999) are still tenable. I do not dispute the experimental findings, nor the oddness of the subject island violations reported in the standard literature. Rather, I dispute the conclusion that the CED is

----

[7] In the case of (21b), replacing the highly specific intervening NP *authors of malicious pamphlets* by *they*, and replacing *do* by *did* significantly improves acceptability. (21a) is particularly odd because it is not fully coherent: writing malicious pamphlets would not lead to book authors arguing about book royalties.

empirically viable as a grammatical phenomenon. When a more representative data are taken into consideration, it becomes highly unlikely that a parsimonious grammatical generalization can be formulated to capture the full range of facts.

2.3 *On a processing account*

Given the empirical problems with the notions of strong islands, parasitism and null resumption, it is possible that the configurational nature of subject and adjunct islands has been overstated. Perhaps such extractions are syntactically well-formed, but constrained by more cognitive and pragmatic general factors. In such a view, there would be no need to stipulate a grammatical distinction between parasitic and real gaps, nor the existence of null resumption in English.

Kluender (1992, 1998), Kluender & Kutas (1993), Hofmeister (2007), Sag et al. (2007), and Hofmeister & Sag (2010) argue that at least some object islands may be best viewed as cases where grammatical sentences obtain low acceptability because they are difficult to process. For example, infrequent expressions, semantically rich phrases, and discourse-new referents are known to consume more cognitive resources than highly frequent (or collocational) expressions, semantically light phrases and discourse-old referents. The processing cost is also inflated when sentences are long and ambiguous, contain phrases that do not cohere or are stylistically deviant, or include filler–gap dependencies. It is known that reading and response times to various kinds of tasks increase inside long-distance dependencies (e.g. Wanner & Maratsos 1978, Chen, Gibson & Wolf 2005), which shows that maintaining an extracted phrase in memory, searching for possible gap sites, and linking the gap to the filler significantly exhausts cognitive processing resources. Such an account would allow for the existence of apparent counterexamples because two isomorphic sentences can incur different processing costs depending on the particular words they contain.

There is independent evidence that processing difficulty can cause low acceptability. For example, doubly center-embedded sentences are known to become easier to process and more acceptable if there are certain cues that help the processor link the predicates to their respective arguments: overt relative pronouns, verbs that cohere well with their respective arguments, semantically light expressions (e.g. pronouns), and a prosodic break between the subject and the VP. This explains why the double center-embedding sentences in (23) obtain completely different degrees of acceptability: whereas (23a) is unparsable, and (23b–d) are graded, (23e) is easy to parse and fully acceptable.

(23) (a) *People people people left left left. (Rogers & Pullum 2011: 330)
  (b) ?*The boy the cat the dog bit scratched started crying.

(c) ?The guy whom the secretary we fired slept with is a real lucky dog.
(Kimball 1975)

(d) ?A syntax book that some Italian that I had never heard of wrote
was published by MIT Press. (Frank 1992)

(e) The movie everyone I know loved was *Inception*.

(Chaves 2012: 480)

This shows that the language processor can be influenced by various kinds of linguistic information, such as complexity, prosody, context and pragmatics. Many examples abound, as Sag (1992) notes. For example, the ambiguous string *I forgot how good beer tastes* is preferentially parsed as [I forgot [[*how good*] [*beer tastes*]]] if it is known that the speaker just returned from a place where beer is banned, and preferentially parsed as [*I forgot* [*how* [*good beer tastes*]]] if the speaker just returned from a place that only has poor quality beer. In this paper it is argued that the same factors that facilitate the processing of double center-embeddings like (23) also facilitate the processing of subject islands.

Kluender (2004) argues that complex subject phrases place particularly heavy demands on working memory, which may in turn lead to low acceptability when these demands are increased by processing a filler–gap dependency. This view receives experimental support from Clausen (2010, 2011), which shows that clausal subjects cause a measurable increase in processing load, with and without extraction. Kluender notes that there are several kinds of evidence in favor of his claim. First, the acquisition data from Bloom (1990) indicate that there is a trade-off between subject length and verb phrase length in early acquisition: the longer the subject, the shorter the verb phrase, and vice-versa. This trade-off appears to also exist in the grammar of adults, and remains constant with age in spontaneous written production as well (Kemper 1987). Bloom (1990) also found that children reduce subjects by whatever means possible. For example, lexical noun subjects are significantly shorter than lexical noun objects, and pronouns are used more often in subjects than in object position. Kemper (1986) also shows that adults between 70 and 90 years of age have far more difficulty repeating sentences with complex subjects than sentences with complex objects. In fact, in 97% (69/75) of the cases, the elderly eliminated the complex subjects from the repetition, whereas there was no such tendency to eliminate complex objects. A similar difficulty was found in timed reading comprehension tasks by Kynette & Kemper (1986). The same pattern is observed in disfluencies in non-elderly adults. For example, Clark & Wasow (1998) show that in spontaneous unmonitored speech, adult speakers of English are more likely to experience a disfluency with subjects than with objects. Interestingly, disfluency rates for both simple and complex NPs show a monotonic decrease across the course of the sentence, in the order topic > subject > prepositional object. Kluender argues that this trend

297

mirrors his results of adult comprehension studies using N400 amplitude as a measure of referential and lexical processing cost. Finally, Garnsey (1985), Kutas, Besson & Van Petten (1988), and Van Petten & Kutas (1991) show that the processing of open-class words, particularly at the beginning of sentences, cause a significantly greater processing effort than closed-class words.

I note that there are other sources of support for Kluender's claim that complex subjects impose a specially heavy processing burden. Ferreira (1991) and Tsiamtsiouris & Cairns (2009) found that speech initiation times for sentences with complex subjects were longer than for sentences with simple subjects, and Amy & Noziet (1978) and Eady & Fodor (1981) provide evidence that sentences with center-embedding in subjects are harder to process than sentences with center-embedding in objects.

It is intuitively plausible that a sentence that starts with a complex structure will consume more processing resources than one that starts with a simple structure and ends with a complex structure. This is because the speaker's memory resources are strained sooner in the sentence, and longer, since those resources are not available for processing the remainder of the sentence. And the fact that complex subjects are harder to process than complex objects is consistent with the fact that extraction from objects is easier than extraction from subjects.

Consequently, speakers will prefer to use sentences with complex structures toward the end of the sentence, thus creating a temporally shorter strain on memory resources (see Yngve 1960; Hawkins 1994; Wasow 1997, 2002; and Gibson 1998, among others). This account can shed light on subject island effects as well as on the existence of graded exceptions. Because subjects are inherently more difficult to process, filler–gap dependencies in them are harder to maintain in working memory without additional support. If the sentence in question happens to contain other elements that are independently harder to process (e.g. infrequent words, complex structures, semantically unexpected constructions, misleading prosody, or stylistically awkward phrasings), then the combined effect causes difficulty. Conversely, sentences with highly frequent words (e.g. collocations), discourse-old nominal phrases, and semantically simpler expressions (such as pronouns), and prosodic phrasings that cue the location of the gap will cause less strain on the cognitive resources needed to process the filler–gap dependency.

Another factor that may play a role in the asymmetry between subject-internal gaps and object-internal gaps is that the former gaps are sentence-medial and the latter tend to be sentence-final. This can make a difference. If there is no material after the gap then there is no risk of creating a garden-path. Conversely, if the gap is sentence-medial, there is a good chance that the gap is not recognized. This is illustrated by the contrasts in (24), where the brackets indicate prosodic boundaries.

(24) (a) ?*[Which cars did the explosion bulge the hoods of a good four inches]?
    (b) [Which cars] [did the explosion bulge the hoods of] [a good four inches]?
    (c) ?*[Who did you give to the canary that I brought home yesterday]?
    (d) ?[Who did you give to] [the canary that I brought home yesterday]?

In all of these data, the object-internal gap is sentence-medial and strands the preposition. Without the proper intonation, these sentences obtain low acceptability because there is a tendency for the preposition to combine with the following phrase. The oddness of such examples is related to the CLAUSE NON-FINAL INCOMPLETE CONSTITUENT CONSTRAINT (Kuno 1973a), which was originally intended to be a more general version of Ross's (1967) SENTENTIAL SUBJECT CONSTRAINT. Interestingly, there is evidence that Kuno's constraint is at least partially a performance effect, since it can be weakened with the proper intonation and a judicious choice of words. Consider the acceptable example in (25), where the brackets indicate prosodic phrasing.

(25) [Who$_i$ did you offer to __ $_i$] [the chance to win $1,000]?

Although this sentence is isomorphic to (24c, d), it is much easier to process. Hukari & Levine (1991) and Fodor (1992) note various counter-examples to Kuno's CLAUSE NON-FINAL INCOMPLETE CONSTITUENT CONSTRAINT, such as (26).

(26) (a) Who did Kim argue with about politics?
    (b) Who did Kim have an argument with about politics?
    (c) Who did you appeal to to get the requirement waived?
    (d) Which company did you persuade the director of to make an appearance?

What these data suggest is that there is a general difficulty in identifying the location of a gap when it is sentence-medial (be it subject-internal or not). This is a likely contributing factor for why it is usually harder to extract from subjects than from sentence-final objects.

Although the proposal in Kluender (2004) is plausible, there are two problems which indicate that there are additional factors at work. First, Kluender admits that nothing explains parasitic gap effects (e.g. in which case the existence of a second gap outside the subject causes the sentence to be acceptable). Second, it is not clear how simple sentences like (27) can lead to sufficient strain on memory and processing resources, causing the parser to crash.

(27) (a) *What did the owner of sneeze? (the owner of the cat sneezed)
    (b) *Who did that Bill married surprise you? (that Bill married Mia surprised you)   (Huang 1982 : 495)

It seems implausible that sentences that are so short can cause such a cognitive breakdown, given that speakers regularly process much larger sentences with much larger subject phrases. Although I agree with Kluender (2004) that the processing strain caused by complex subjects has a central role in subject island phenomena, it is in my view unlikely that the mere presence of a filler–gap dependency in the subject phrase exhausts memory resources so severely as to cause subject islands effects.

In what follows I concentrate on subject islands and hope to convince the reader that the range of acceptable subject island violations is substantially wider than previously held. This means that the strength of so-called 'strong islands' has been overstated in the literature, and that no previous account discussed so far can cope with the full range of facts. I then proceed to show how the processing account put forth by Kluender (2004) can be expanded so that all the subject island phenomena under discussion are predicted.

3. SUBJECT ISLANDS AND THEIR NON-PARASITIC CIRCUMVENTION

Though widely accepted, the status of English subject phrases as syntactic extraction islands is questionable. It is well-known since Ross's seminal work that the extraction of PPs from subjects is possible (although this too is sometimes disputed, as in Culicover (1999: 230–232) and Lasnik & Park (2003: 653), for example). A variety of evidence supports this view, shown in (28), which includes attested examples from Santorini (2007).

(28) (a) Of which cars were the hoods damaged by the explosion?
(Ross 1967, ex. (4.252))
(b) They have eight children of whom five are still living at home.
(Huddleston, Pullum & Peterson 2002: 1093)
(c) That is the lock to which the key has been lost.
(d) A house of which only the front has been painted will be on your left at the second light; you make a right turn there.
(Levine & Hukari 2006: 291)
(e) … a letter of which every line was an insult …
(f) Already Agassiz had become interested in the rich stores of the extinct fishes of Europe, especially those of Glarus in Switzerland and of Monte Bolca near Verona, of which, at that time, only a few had been critically studied. (Santorini 2007)
(g) It was the **car** (not the **truck**) of which [the driver was found].
(h) Of which car was the driver awarded a prize? (Chomsky 2008: 147)

Although it is generally assumed that such cases are the only robust class of exceptions to subject islands, I believe there are various other classes that deserve attention. The first class of exceptions involves adverbial extraction, as originally noted by Grosu (1981: 72).

(29) (a) The 'Hunan' Restaurant is a place where having dinner promises to be most enjoyable.
  (b) The pre-midnight hours are the time when sleeping soundly is most beneficial to one's health.

A small number of authors have argued that, rather than being impossible, some NP extractions from subject NPs are passable or fairly acceptable, as illustrated in (30).[8]

(30) (a) What were pictures of seen around the globe? (Kluender 1998: 268)
  (b) It's the kind of policy statement that jokes about are a dime a dozen. (Levine et al. 2001: 204)
  (c) There are certain topics that jokes about are completely unacceptable. (Levine & Sag 2003: 252, fn. 6)
  (d) Which car did some pictures of cause a scandal? (Fernández 2009: 111)
  (e) What did the attempt to find end in failure? (Hofmeister & Sag 2010: 370)

This matter deserves closer scrutiny than the one it has received so far. In an attempt to improve and expand this small data set, I have constructed various examples involving a variety of predicates, including passives, actives, unaccusatives, and unergatives. According to my informants, the sentences in (31) and (32) are acceptable. Crucially, the square brackets indicate prosodic phrasing: one break must be inserted after the extracted phrase, and another at the gap site.

(31) (a) [Which disease] [will the cure for] [never be discovered]?
  (b) [Which question] [will the answer to] [never be known]?
  (c) [Which problem] [will a solution to] [never be found]?
  (d) [Which transaction] [will the value of] [never be known]?
  (e) [This is the bill] [that an amendment to] [will never be accepted].
  (f) [This is the film] [that a sequel to] [will never be produced].

(32) (a) [Which president] [would the impeachment of] [cause outrage]?
  (b) [Which airline] [is the crew of] [currently on strike]?
  (c) [Which doctors] [have patients of] [filed malpractice suits in the last year]?
  (d) [Which schools] [did the students of] [earn the highest scores]?
  (e) [Which school] [has the principal of] [recently resigned]?

It is important to note that a very small number of syntactic accounts allows for some degree of subject island circumvention. Sauerland &

---

[8] See also Pollard & Sag (1994: 195, fn. 32), Postal (1998), Culicover (1999: 230), Sauerland & Elbourne (2002: 304), Levine & Hukari (2006: 71), and Chaves (2012: 470, 471).

Elbourne (2002: 304) propose that subject island circumvention only arises as a PF phenomenon, in the presence of some scope-taking element that interacts (non-spuriously) with the scope of the *wh*-phrase. Sauerland & Elbourne claim that this accounts for the contrast in *Which constraint are good examples of always sought/*provided?* (their judgments). Basically, *good examples* can take narrow or wide scope under *sought* but cannot do so under *provided*. Since the latter is not scopally ambiguous, the movement of *which constraint* cannot be delayed until PF, and the sentence is rejected. It is easy to see that this account fails to capture most of the data reported in this section. Chomsky (2008) proposes that extractions from subjects are only possible in passive structures since the subject is an internal argument. But this view is problematic in the light of examples like (32). Second, according to the Subject Internal Hypothesis that is widely accepted in transformational grammar, all subjects are generated VP-internally. Consequently, Chomsky's proposal becomes vacuous. Finally, Fernández (2009) proposes a rather stipulative account: if a subject is definite and non-D-LINKED then it is a strong phase, and sub-extraction is blocked. A D-linked *wh*- phrase is typically viewed as one that limits the range of felicitous answers to the members of a contextually defined set. However, there are no known non-circular criteria for determining if something is D-linked or not and it remains unclear why D-linking should have any effect on extraction in the first place (Pesetsky 2000: 16).

It is even more widely held that English clausal subjects are insuperable extraction islands. This goes back to Ross (1967, ex. (4.254)) and to Huang's (1982: 495–497) examples, shown in (33).

(33) (a) *Who did that Bill married surprise you?   (=ex. (27b) above)
     (b) *Who did he say that for Bill to marry was a surprise?

Ross (1967, ex. (4.254)) described this condition as the SENTENTIAL SUBJECT CONSTRAINT (SSC), which stipulated that no element dominated by an S may be moved out of that S if that node S is dominated by an NP which itself is immediately dominated by S. This conclusion is also based on an unrepresentative data set, since there are various kinds of counterexamples involving a wide range of clausal subjects. The first case was first noted by Huddleston et al. (2002), who point out that extraction from infinitival VP subjects is attested in (34).

(34) The eight dancers and their caller, Laurie Schmidt, make up the Farmall Promenade of nearby Nemaha, a town that [[to describe __ as tiny] would be to overstate its size].

<div align="right">(Huddleston et al. 2002: 1094, fn. 27)</div>

There is nothing exceptional about this particular data point. I have found many other attestations, a subset of which is given in (35).

(35) (a) In his bedroom, which to describe as small would be a gross understatement, he has an audio studio setup.

(http://pipl.com/directory/name/Frohwein/Kym, retrieved 21 February 2012)

(b) They amounted to near twenty thousand pounds, which to pay would have ruined me.

(Bejnamin Franklin, William Temple Franklin & William Duane. 1834. *Memoirs of Benjamin Franklin*, vol. 1. p. 58; http://www.archive.org/details/membenfrank01frankrich, retrieved 21 February 2012)

(c) The … brand has just released their S/S 2009 collection, which to describe as noticeable would be a sore understatement.

(http://www.missomnimedia.com/2009/page/2/?s=art + radar &x=0&y=0, retrieved 21 February 2012)

(d) Because this does purport to be a food blog, I will move from the TV topic to the food court itself, which, to describe as impressive, would be an understatement.

(http://phillyfoodanddrink.blogspot.com/2008/06/ foodies-food-court.html, retrieved 21 February 2012)

Stowe (1986), Ellis (1991) and Pickering, Barton & Shillcock (1994) show that comprehenders tend to postulate gaps in objects whenever possible (the so-called 'filled-gap' effect) except in subject phrases. In other words, speakers do not expect to find subject-internal gaps. In an online study, Phillips (2006) shows that comprehenders postulate the existence of a subject-internal gap for infinitival subjects, as in (36a), but not for finite subjects, as in (36b).

(36) (a) *The outspoken environmentalist worked to investigate what the local campaign to preserve had harmed the annual migration.

(b) *The outspoken environmentalist worked to investigate what the local campaign that preserved had harmed the annual migration.

This would be expected if subjects are not configurational islands and if finite clauses are independently harder to process than non-finite clauses. This is independently supported by the experimental results of Gibson (2000), which suggest that referential structures – including finite phrases – add a measurable cognitive load to processing. Moreover, Kluender (1992) plausibly argues that eventive finite clauses are the hardest to process, modal finites are intermediate, and infinitival clauses are easiest. If finiteness adds a significant cognitive load to sentence processing because of referentiality, as Kluender and Gibson argue, then it is reasonable that additional computational difficulty arises when it is compounded with the processing of complex structures containing filler–gap dependencies. Indeed, Phillips

(2006) observes a main effect of finiteness in his study, which further supports my interpretation of the results.

However, Phillips (2006) interprets the experimental findings differently, by arguing that comprehenders only postulate subject-internal gaps in contexts where the gap can be parasitic, which in turn is taken to be strong evidence that subject islands are part of grammar. This argument is problematic on two grounds. First, it assumes that finite subjects cannot contain parasitic gaps. This is false, as Phillips (2006: 803, fn. 6) admits, given well-known examples like (37).

(37) She is the kind of person that everyone who meets __ ends up falling in love with __.   (Kayne 1983)

Second, the evidence in (34) and (35) above undermines the assumption that gaps in non-finite subjects must be parasitic. A more likely explanation is that comprehenders postulate a subject-internal gap in such environments not because of parasitism, but simply because such gaps are syntactically legal. The results obtained by Phillips (2006) may simply reflect the fact that finite subjects are harder to extract from than infinitival subjects, because of the fact that finite clauses are harder to process.

Moreover, note that extractions from clausal subjects are not by any means limited to infinitival VPs. The subject clause can be a full infinitival CP, as seen in (38). A parenthetical intonation is optimal for the acceptability of these data points.

(38) (a) This is something which – for you to try to understand – would be futile.   (Kuno & Takami 1993: 49)
     (b) I just met Terry's eager-beaver research assistant – who for use to talk to about any subject other than linguistics – would be absolutely pointless.   (Leine & Hukari 2006: 265)
     (c) There are people in this world that – for me to describe as despicable – would be an understatement.   (Chaves 2012: 471)

We can take this further and seek acceptable extractions from finite subjects. Consider the example in (39), where brackets indicate prosodic phrasing boundaries.

(39) [Which actress does whether Tom Cruise marries] [make any difference to you]?

The above data are much more acceptable than the classic examples in (33) and (36). It is possible that the latter have low acceptability because they are more complex and less natural than (35) and (38). The sentences in (36) involve an embedded clause with multiple extractions and contain a null object structure (compare with *We investigated if the campaign to preserve the frogs had harmed the annual migration (of storks)*). This can cause additional processing difficulty because, out-of-the-blue, it is very unclear if the missing

objects in (36) are supposed to be co-referential with the extracted phrase or not. Also, if Kluender is correct, tense adds an additional processing load.

The speakers that I have consulted also do not find the datum in (40) below as bad as Ross (1967) claimed. With a parenthetical-like intonation on the relative clause, (40) is passable, when in fact it should be impossible outright. The square brackets indicate prosodic phrasing breaks. The ameliorative role of prosody is unexpected if subject islands are grammatical phenomena, but expected if they are an extra-grammatical epiphenomenon.

(40) ?[The hat which] – [**that I brought**] [seemed strange to the nurse] – [was a fedora]. (Ross 1967, ex. (4.260))

Finally, the existence of a syntactic constraint on deep extraction from subjects is highly dubious, as the acceptability of (41) shows. Again, square brackets indicate the required prosodic phrasing.

(41) [I have a question] [that the probability of you knowing the **answer to** __] [is zero].

In this datum, a nominal gap resides inside the object NP *the answer to*, which is inside an object gerundial clause *you knowing the answer to*, which is inside the subject NP *the probability of you knowing the answer to*. Crucially, all the heads semantically cohere well with the extracted nominal. In contrast, examples like (36) above are fairly unnatural because the components of the phrase do not cohere as well (for more discussion see Section 3.1.2 below).

In sum, the set of data that has been used to inform syntactic theories of extraction in the last half a century is not representative of the full spectrum of possibilities. As a consequence, the role of configurational limitations in the grammar of subject-internal extraction has been overstated, and the correct generalizations have been missed.

3.1 *Factors that facilitate the processing of subject island violations*

In a performance-based account of subject islands like Kluender (2004), it is expected that subject island violations involving non-finite subjects are easier to process than those with finite or nominal subjects. This is because the former are the weakest, semantically, as they carry no referentiality (Kluender 1992, 1998). This is independently supported. First, extraction from NPs is known to often be difficult even in object position (see Section 2 above). Second, finite object relative clauses are also well-known to impose stronger extraction limitations than non-finite object relatives (Ross 1967, Engdahl 1983), and are harder to process than non-finite clauses (Kluender 1992, Gibson 2000). Third, non-finite subjects are the only environment where 'filled-gap' effects have been detected (Stowe 1986, Ellis 1991, Pickering et al. 1994, Phillips 2006).

In what follows, I focus on three independently motivated factors that seem to contribute to the relative acceptability of subject island violations. These factors are particularly evident in NP extractions from subject NPs like the ones in (31) and (32): (i) the specificity of the *wh*-phrase; (ii) the highly coherent pragmatic relevance that the extracted phrase bears to the subject and the predicate; and (iii) the syntactic cues introduced by prosodic phrasing. To be clear, none of these factors is directly connected to any grammatical condition on subject islands, nor is it by itself sufficient to completely mitigate a subject island violation. Collectively, however, they can cue the correct parse and aid the processor in roughly the same way that double center-embeddings can be made easier to process with similar cues, as discussed in Section 2.3 above. Crucially, the ameliorative role of specificity, relevance and prosody is independently motivated, since it is known to aid the processing of a wide range of other complex constructions.

3.1.1 *Specificity*

Erteschik-Shir (1973), Kluender (1992) and others note that filler–gap dependencies with less specific filler phrases (i.e. less informative by simple subset–superset relationship) tend to be less acceptable than filler–gap dependencies with more specific ones, as (42) illustrates.

(42) (a) This is the car that I don't know how to fix __.
    (b) ?Which car don't you know how to fix __?
    (c) ??What don't you know how to fix __?

The ameliorative effect of specific fillers has been experimentally observed for a variety of different extraction phenomena, by several different psycho-linguistic methodologies (Kluender 1992, 2004; Kluender & Kutas 1993; Hofmeister 2007; Sag et al. 2007; Hofmeister et al. in press). For example, Hofmeister and colleagues show that sentences with more informative in-definite NP like in (43a) allow faster retrieval at the gap site than its less informative counterpart in (43b).

(43) (a) It was [an influential communist-leaning dictator] that Sandy said she liked.
    (b) It was [a dictator] that Sandy said she liked.

Focusing on the object position, Hofmeister & Sag also show that Complex NP Constraint violations with more informative fillers are more acceptable and are processed faster at the gap site than violations with less informative fillers. Take for instance the sentences in (44).

(44) (a) Which military dictator did you say that nobody in the world could ever depose?
    (b) ?Who did you say that nobody in the world could ever depose?

The same difference in reading times is found in sentences without extraction. Thus, (45b) was found to be read faster at *encouraged* than (45a).

(45) (a) The diplomat contacted the dictator who the activist looking for more contributions encouraged to preserve natural habitats and resources.

(b) The diplomat contacted the ruthless military dictator who the activist looking for more contributions encouraged to preserve natural habitats and resources.

Kluender (1998) and Hofmeister & Sag (2010) argue that memory retrieval and memory decay straightforwardly predict this ameliorative effect: more specific *wh*-phrases are less prone to memory decay and therefore are more easily retrieved downstream than less informative (by simple subset–superset relationship) *wh* -phrases like *who* or *what*.

Another reason for specific phrases facilitating processing may come from the fact that non-specific phrases like *who* are semantically compatible with a wider range of thematic roles than *which military dictator*. Hence, there are more candidate gap sites along the way to the actual gap site (Garnsey, Tanenhaus & Chapman 1989, Traxler & Pickering 1996). This tendency for the processor to link a filler to a gap as soon as possible (Stowe 1986, Frazier 1987, Ellis 1991) may cause a cascade of minor processing disruptions as the parser must attempt several plausible gap sites before finding one that allows the sentence to be coherently parsed.

The same specificity amelioration effect arises in subject islands. Clausen (2010) shows that more specific *wh*-phrases have an ameliorative effect in sentences with gerundial subject islands, although this manipulation is not by itself able to cause unacceptable sentences to become fully acceptable. This specificity effect has also been replicated by Chaves & Dery (in press) for interrogative sentences like those in (31) and (32) above.[9] Thus, sentences like (46a) are measurably more acceptable than (46b).

---

[9] In one of the Chaves & Dery (in press) experiments, 43 native speakers of English were recruited to judge 40 quadruples of sentences, using a Magnitude Estimation methodology. The experiment manipulated two 2-level factors: specific *wh*-phrase and definite subject (*Which politician did the opponents of organize a protest?*), specific *wh*-phrase and indefinite subject (*Which politician did the opponents of organize a protest?*), non-specific *wh*-phrase and definite subject (*Who did the opponents of organize a protest?*), and non-specific *wh*-phrase and indefinite subject (*Who did opponents of organize a protest?*). There were three types of distractor items, 80 in total, half of which were ungrammatical. A linear mixed-effects regression was used with the following random-effect factors: a by-subject adjustment to the intercept ($SD = 0.32$); a by-item adjustment to the intercept ($SD = 0.054$); a by-list adjustment to the intercept ($SD < 0.001$); and the residual error ($SD = 0.28$). *Wh*-phrase specificity was significant, with less specific *wh*-phrases resulting in lower acceptability scores ($\beta = -0.086$, $t = -4.369$, $p < .0001$). Definiteness of the intervening noun was not significant in predicting acceptability ($\beta = -0.003$, $t = 0.162$, $p = .871$). An

307

(46) (a) Which musician will the full discography of never be released?
    (b) Who will the full discography of never be released?

Note that this account makes no appeal to the ill-understood notion of D-linking (see the passage earlier in this section, following examples (31) and (32)). As discussed by Kroch (1989), Chung (1994: 39), Ginzburg & Sag (2000: 247–250), Pesetsky (2000: 16), and Levine & Hukari (2006), the causal nexus between islands and D-linking remains obscure.

3.1.2 *Relevance*

Kuno (1987), Erteschik-Shir (1981) and Deane (1992) argue that the contrast between sentences like (47) is pragmatic in nature. In Kuno's terms, only those constituents in a sentence that qualify the topic of the sentence can undergo extraction (the term TOPIC refers to the entity that the speaker is talking about; it is usually also referred to as THEME). In other words, a filler must always be relevant for the assertion, otherwise the filler–gap dependency is not pragmatically coherent.

(47) (a) Who did John write a book about?
    (b) ?Who did John destroy a book about?
    (c) ?Who did John lose a book about?

Sentence (47a) is about writing a book. Since books have topics, the writing action is connected to the book's topic. In theories of lexical semantics like Pustejovsky (1995: Chapter 5), for example, a book's topic is a 'shadow argument' in the argument structure of *write*.

Sentences (47b, c), however, are about losing and destroying a book, which has no immediate relevance to the book's topic. Hence, *write a book about Nixon* is more coherent than *destroy a book about Nixon*. One major advantage of this approach over a syntactic account is that it explains why context can cause (47b, c) to be more felicitous. For example, if John is known to usually destroy books, then (47b) becomes fully acceptable.

---

interaction of *wh*-phrase specificity and definiteness of the intervening subject NP was also non-significant ($\beta = -0.019$, $t = -0.712$, $p = .476$). Satiation was found to be highly significant, with a *t*-value of 4.415 and a *p*-value of $p < .0001$. These results were replicated by a follow-up seven-point Likert scale experiment with 93 self-reported native English speakers with US IP addresses recruited via Amazon's Mechanical Turk. The resulting satiation results are consistent with other's (Hiramatsu 1999, Snyder 2000, Francom 2009), but not with Sprouse (2009) and Crawford (2011a, b). Chaves & Dery (in press) suggest that the latter null effects are a consequence of using items that are too difficult to process (e.g. contain finite-clause embedding and non-specific non-relevant *wh*-phrases). That is not to say that unacceptability necessarily ensues by clause-embedding a sentence with a subject island or by using a non-specific *wh*-phrase. This depends on how coherent and how difficult the particular sentence is, as is shown by the relative acceptability of examples like (30), (41), and (60) in the main text.

If relevance plays a role in the extractions in (47), then it is possible that it also plays a role in extractions from within subject phrases.

This same relevance condition seems to be important for subject islands. In his concluding remarks, Kluender (2004: 495) insightfully suggests that fillers maintain an association not only with their gaps, but also with the main clause predicate, such that the filler–gap dependency into the subject position is construed as of some relevance to the main assertion of the sentence. Kluender does not define what he means by relevance, but I will assume the following. An extracted phrase is considered relevant for the subject phrase to the degree that the concept described by the subject presupposes the concept described by the extracted phrase. For example, a solution necessarily presupposes the existence of a corresponding problem. Moreover, an extracted phrase is relevant for the main assertion to the degree that the referent it describes influences the truth conditions of the predication (for example, as in the case of (47) above). To illustrate, consider the pair in (48). Brackets indicate the ideal prosodic phrasing.

(48) (a) [Which problem] [will the solution to] [impress everyone]?
     (b) ?[Which city] [will the train to] [impress everyone]?

A solution is a concept that intrinsically depends on the existence of a corresponding problem. If there is a solution, then it follows that there must be a corresponding problem. Hence, any statement about a given solution depends on the problem under discussion. The problem under discussion is necessarily relevant for any claim about solving that problem. Thus, the *wh*-phrase in (48a) is relevant for the subject and for the predicate. Conversely, trains do not intrinsically presuppose the existence of a city, and therefore the relevance between the filler and the subject is weaker in (48b). Moreover, it is also harder to construe a way in which a city is relevant for the event of a train impressing someone.

The oddness of examples like (49) below is also correctly predicted.

(49) (a) *What did the owner of sneeze? (the owner of the cat sneezed)   (= ex. (27a) above)
     (b) *Which table did the cat under yawn? (the cat under the table yawned)

In (49a) the *wh*-phrase is maximally unspecific and therefore a relevance dependency between it and the verb and the subject noun is not easy to construe. Similarly, in (49b) the phrase *which table* is not relevant for *the cat* nor for the situation described by the main predicate *yawn*. To be clear, the filler must be relevant for both the subject and the VP. Sentences like *Who did your book about mention Clinton?* (compare with *Your book about Bush mentioned Clinton*) are odd because *who* is not relevant for the VP *mention me*, even though *who* (if construed as a topic) is relevant for the subject.

309

Note that all the sentences in (31) and (32) conform to the subject/assertion relevance conditions discussed above.

Further support for the importance of relevance in subject islands comes from subject islands in other languages. For example, extraction out of sentential subjects in Japanese is known to be acceptable, as (50) illustrates.

(50) ano hito$_i$ wa watakusi ga __$_i$ au   no   ga   muzukasii
     that person TOP I   NOM   meet NMZ NOM difficult
     'the person whom that I see (him) is difficult'   (Kuno 1973b: 241)

This is not always the case, however. Shimojo (2002: 111) notes the contrast in (51), and argues that the oddness of (51a) results from the fact that there is no obvious connection between the filler 'clothes' and the bridging clause 'the gentleman is missing'. Conversely, in (51b) the filler 'gentleman' is more relevant for both propositions: he is wearing the clothes and his clothes are dirty. The same is true for topicalizations.

(51) (a) *[[__$_i$ __$_j$ kiteita]   sinsi$_i$ ga   yukuehumeida] yoohuku$_j$.
                was.wearing gentleman NOM missing   clothes
          'The clothes which the gentleman who was wearing (them) is missing'
     (b) [[__$_i$ __$_j$ kiteiru] yoohuku$_j$ ga   yogoreteiru] sinsi $_i$.
                wearing clothes   NOM dirty   gentleman
          'The gentleman who the clothes that (he) is wearing are dirty.'

Sells (1984: 304ff.) also notes that subject island violations in Swedish require certain pragmatic assumptions in order to be acceptable. These facts suggest that pragmatics plays a role not only in English subject-internal extractions, but also in languages that more readily allow extraction from subjects. I am not proposing that there is a special pragmatic condition that specifically targets subject phrases. Rather, I am simply assuming, with Kuno (1987), that all extracted phrases must bear some relevance to the phrase from which they are extracted from. As discussed above, the filler–gap relevance is in some cases fairly direct and obvious (as in the case of *finding* and *a solution to a problem*), in other cases it can be easily inferred (as in *writing* and *a book about a person*), and in other cases it cannot be easily inferred but it can be explicitly provided by context (as in burning a book about a person). Otherwise, the filler–gap dependency is not coherent and the sentence is less than felicitous.

I conjecture that the subjects of passive and unaccusative predicates may be more transparent to extraction than subjects with unergative predicates (as suggested by Kravtchenko et al. (2009), for example) because in the latter case it is much harder for any subject-internal phrase to be relevant for the main assertion. Since the subject is an agent or an actor, it initiates or controls the event, and is therefore the most relevant participant for the

assertion. But if the subject is not an agent or actor, then it is easier for a phrase other than the subject to be construed as relevant.

3.1.3 *Prosody*

Prosody can influence sentence parsing and create or mitigate processing difficulty. For example, the garden-path effect in sentences like (52a) is reduced if a pause is inserted before the main verb as in (52b).

(52) (a) *[The government plans to raise taxes were defeated].
  (b) [The government plans to raise taxes] [were defeated].

Other examples abound. For instance, Morgan (1973) noted that (53) is odd with neutral intonation, but perfectly acceptable with a pause between the verb and the PP, as a reply to *How does Nixon eat his tapioca?*. Otherwise, the semantically odd parse where the PP modifies the VP is highly preferential and preempts any other alternative.

(53) I think with a fork.

Fodor (2002a, b), Kitagawa & Fodor (2006), Ackerman, Yoshida & Pierrehumbert (2011), Zahn & Scheepers (2011), and others have experimentally shown that prosody can have a measurable impact in syntactic judgments. For example, Zahn & Scheepers (2011) provides online evidence that the presence or absence of a strong phrase boundary before the relative clause in (54) can function as a cue to disambiguate the level of attachment of the relative clause to the complement phrase.

(54) The vet examined the leg of the horse that was badly injured.

Ackerman et al. (2011) show that prosody can have a significant effect in the acceptability of center-embedded sentences. For example, inserting a pause immediately before the matrix verb improves the acceptability of (55).

(55) The man that everyone who I know raved about turned out to be boring.

Prosodic cues play a role in the acceptability of subject island violations as well, a factor that has never been controlled for in past research. In fact, many of the starred data in the standard literature such as (5) from Section 2.1 above improve in acceptability if they are produced with a prosody that cues where the extraction occurs. Take (5b), for example, *Which book did a review of appear in the Times?*, which is marked with an asterisk by Jackendoff (2002: 48). Its acceptability improves if the prosodic phrasing shown in (56a) is used instead of a neutral realization like (56b). Here, brackets signal the presence of prosodic breaks.

(56) (a) [Which book] [did a review of] [appear in the Times]?
    (b) ??[Which book did a review of appear in the Times]?

In fact, the subject island ameliorative effect of prosody was first noted by Ross (1967, ex. (4.267)), with regard to cases like (57). Ross deemed this datum passable in English with the proper intonation. Hence, it is clear that Ross did not believe that NPs could not be extracted from phrasal subjects, a fact that is for some reason never mentioned in the literature on strong islands, as far as I can tell.

(57) That piano, which the boy's loud playing of drove everyone crazy, was adly out of tune.

To be clear, I am not suggesting that prosody has anything to do with subject island effects. Rather, the claim is that prosody can cue the presence of the (rather unconventional) location of the subject-internal gap, and aid processing the sentence.

4. AN EXPECTATION-BASED ACCOUNT

It is independently known that the more committed the parser becomes to a syntactic parse, the harder it is to reanalyze the string (Ferreira & Henderson 1991, 1993; Tabor & Hutchins 2004). This is usually referred to as a 'digging-in' effect. For example, unless prosodic or contextual cues are employed to boost the activation of the correct parse, (58) will be preferentially mis-analysed as having the structure [NP [V [NP]].

(58) Fat people eat accumulates.

The garden-path caused by the digging-in effect in (58) serves as an analogy for what may be happening in unacceptable sentences with subject island violations. In both cases, the sentences have exactly one grammatical analysis, but that parse is preempted by a highly preferential alternative which ultimately cannot yield a complete analysis of the sentence. Thus, without prosodic cues indicating the extraction site, sentences like (59) induce a significant digging-in effect as well.

(59) (a) *[Which problem will a solution to be found by you]?
    (b) *[Which disease will a cure for be found by you]?

Given that processing complex subjects is cognitively more strenuous than processing complex objects, and that certain pragmatic conditions restrict the use of filler–gap dependencies in general, it is plausible that speakers avoid the use of sentences with subject-internal gaps. This leads to extremely low (near zero) frequency. In turn, extremely low frequency may cause the language processor to develop a conventionalized processing

heuristic: expect gaps to be in the verbal structure, not in the subject phrase. In this view, the subject-internal gap parse is either very weakly activated or not even attempted by the parser not because of a cognitive breakdown or a gramaticized constraint, but because of the strong expectation that the gap is situated in the VP instead. Given the extremely low frequency of subject-internal gaps, these heuristic parsing expectations are usually met. However, if a sentence contains a subject-internal gap then the expectations about the location of the gap mislead the parser and cause a digging-in effect. The latter hampers backtracking, especially if the sentence is difficult to process for independent reasons (e.g. the filler–gap dependency is very long, semantically complex, or lacks coherence).

This gapless subject expectation cannot be seen as a grammatical condition. Precisely because it is a parsing expectation rather than a grammatical rule, it can be dampened by the presence of prosodic, pragmatic and contextual cues that signal the correct parse, as discussed in Section 3.1 above. In the same way that context and prosody can influence how an ambiguous relative clause is attached or how a quantifier scope ambiguity is resolved, so can the processing of subject islands be sensitive to prosodic and pragmatic information. In what follows it is argued that this parsing expectation can explain subject island phenomena as well as (non-)parasitic subject island circumvention.

4.1 *Cognitive and pragmatic reasons for avoiding subject-internal gaps*

As discussed in Section 2.3 above, there is evidence that sentences with complex subjects are significantly harder to process than sentences with complex objects, and that sentences with medial gaps are more difficult to process than sentences with sentence-final gaps. Hence, there are cognitive reasons for such structures to be avoided. Below, I note other factors that may further discourage speakers from using complex sentences with subject-internal gaps, one cognitive and one pragmatic.

As (60) illustrates, sentences with subject-internal gaps require discourse contexts which are fairly complex as well. This added complexity is bound to strain the processor further.

(60) [Biology exam question]
　　　For quite some time now, medical science has been seeking cures for the flu and for tuberculosis. Current research indicates that the cure for one of these diseases may be possible to find in the near future, but not for the other.
　　　Which disease do you think the cure for will not be available in the near future?

As a consequence, there is an additional reason to avoid complex sentences that have subject-internal gaps. Hence, speakers will plan and package their

discourse differently, using paraphrases that are simpler to plan and produce, and avoid discourse like (60).

Moreover, it is often the case that a subject-internal gap serves no special communicative purpose. Consider, for example, the possible alternative continuations in (61).

(61) (a) I asked Fred to write me two lines of poetry. But instead of that...
   (b) … he wrote me a letter, and **every single line of it** was a piece of poetry.
   (c) (?) … he wrote me a letter of which **every single line** was a piece of poetry.
   (d) ?* … he wrote me a letter which **every single line of** was a piece of poetry.

Although both (61c, d) are suitable continuations for (61a), they compete with the much simpler paraphrase in (61b), which involves a discourse-old pronominal and no extraction. Consequently, (61a) is not only easier to process, but also pragmatically more felicitous than (61c, d). Note also that the PP extraction in (61c) is more helpful for the processor than the NP extraction in (61d): in the latter the extracted phrase is maximally uninformative about its grammatical role.

There is no semantic or pragmatic motivation to use (61c) instead of (61b), nor to use (61d) instead of (61c). To be clear, my claim is not that the existence of a simpler paraphrase causes the more complex paraphrase to be unacceptable. Rather, the point is that simpler paraphrases are preferred over more complex ones, which in this case adds further pressure for the lower frequency of sentences with subject-internal gaps. In sum, the evidence above suggests that there may be Gricean and stylistic reasons for speakers to avoid continuation paraphrases like (61c, d).

Let us take stock. There are cognitive and pragmatic reasons for speakers to avoid complex subjects that contain filler–gap dependencies. From this it follows that under real-time communicative pressure speakers will package discourse in a simpler way and resort to sentences with simpler (gapless) subject phrases. This leads to extremely low (near zero) frequency of sentences with subject-internal gaps. As a consequence of this, it is possible that the language processor develops a parsing heuristic in which subjects are not expected to contain gaps. Such a heuristic is useful because it efficiently prunes the search space of filler–gap dependencies and speeds up processing. I now turn to this hypothesis in more detail.

4.2 Extremely low frequency leads to parsing expectations

Fodor (1978, 1983), Berwick & Weinberg (1984), and Hawkins (1999) propose that processing difficulties might lead to the grammaticization of islands.

This position is excessive for subject islands given that there are various grammatical exceptions (both parasitic and non-parasitic) as shown in Sections 2 and 3 above. An alternative is to assume that islands are extremely heterogenous phenomena that depend on the particular lexical and constructional types, as in Postal (1998), but such an account would lack explanatory power since it would merely stipulate which specific words and constructions block extraction.

I propose that the extremely low frequency of extractions from subjects causes speakers to create a processing heuristic about the absence of gaps in subject phrases. The rationale is the following. If a given construction is hard to process for cognitive reasons, this makes it less frequent, and therefore less expected, which in turn may cause it to be harder to process. This in turn makes the construction even less frequent, less expected, and so on. The acceptability of subject-internal extractions is possible, however, in the presence of cues that identify the correct parse and aid the filler-gap processing mechanism in coping with the unexpected location of the gap. I conjecture that different languages can create different parsing heuristics for filler–gap dependencies, informed by processing considerations, frequency, and pragmatic conditions, in the spirit of Hawkins (2004). For example, Engdahl (1983: 9–11, 29) notes that there is a hierarchy for the acceptability of parasitic gaps, and that the break between acceptable and unacceptable extraction domains seems to be language-specific. Engdahl (1983: 29) plausibly conjectures that these facts reflect parsing principles.

Given the view that I am advocating here, subject island violations are grammatical structures that have different degrees of acceptability in proportion to the amount of reprocessing needed for the correct parse to recover from the expectations that it has about the location of the gap. Suppose that the parser has encountered a filler, a *wh*-phrase of some kind. This causes it to actively seek a gap in the remainder of the sentence. When a subject phrase is encountered, a problem presents itself: either the subject contains the gap for that filler or it does not. An efficient decision based on the frequency and distribution of gaps in English is to assume the latter. Subject-internal gaps are so rare that this processing strategy will reliably speed up processing in the vast majority of cases, and therefore help the processor operate efficiently under real-time communicative constraints. The proposed processing strategy is formulated in (62).

(62) *Gapless subject heuristic*
      Expect subject phrases to be gapless.

It is important to stress that this is not part of the grammar. In other words, this is not a constraint on well-formed English syntactic structures, but rather, the preferential route taken by the parser when seeking a gap for a filler stored in memory. Crucially, (62) has the corollary in (63) given that the

315

absence of a subject-internal gap logically leads to the expectation that the gap is located in the subject's sister phrase.

(63) *Corollary*

  Expect any gap to reside in the sister of the subject phrase (i.e. the VP).

Violating any or both of these expectations does not lead to un-grammaticality, since these are parsing rules of thumb rather than grammar rules. The combined product of (62) and (63) is a digging-in effect: the parser is committed to the prediction that no gap will be found in the subject and that a gap will be found in the VP. The longer the parser remains committed to these expectations the stronger the digging-in. In turn, speakers will experience a strong garden-path when the two expectations about the gap location turn out to be incorrect. When there are prosodic, pragmatic and contextual cues which boost the grammatical parse – as discussed in Section 3 – the digging-in effect is naturally weaker and therefore it is easier for the parser to backtrack. But in the absence of such cues, the parser experiences a strong digging-in effect which hampers backtracking.

  Let us consider some examples. Uttered with neutral prosody, (64a) violates two expectations: the expectation that the subject is gapless and the expectation that the verb phrase contains a gap. As a consequence of the expectations, the *wh*-phrase is retrieved in the VP but the parse crashes because the legal gap is located in the subject instead.

(64) (a) *Who did the rivals of __ shoot Castro?
     (b) Who did the rivals of __ shoot?
     (c) Who did the rivals of Castro shoot __?

Violating two expectations leads to a stronger digging-in effect than violating just one, and therefore to a stronger garden-path effect. Consequently, (64a) is harder to process than (64b). The relative acceptability of (64b) follows from the fact that the two gaps are in very close proximity (which causes less digging-in) and are referentially linked to the same filler. It is known that referents that have been recently accessed in memory are easier to reactivate and process (see e.g. Vasishth & Lewis 2006). Hence, the filler is reactivated as the parser finds each gap, and processing the filler–gap dependency is facilitated. The overall effect is the illusion that the subject gap in (64a) is somehow grammatically different from the subject gap in (64b).

  Cases like (19b), repeated as (65), are harder than (64b) because the two gaps are not referentially linked, but are still easier than (64a) because only one expectation is violated.

(65) [People that sensitive]$_i$, I never know [which topics]$_j$ jokes about __$_j$ are likely to offend __ $_i$.   (=ex. (19b) above)

  The same explanation applies to pairs of sentences like (66).

(66) (a) What did the attempt to repair __ ultimately damage __?
(b) What did the attempt to repair the car ultimately damage __?

In (66a) only one expectation is violated and the two co-referential gaps are in very close succession. As a consequence, this structure cannot yield a very strong digging-in effect, and therefore it is easy for the parser to overcome the gapless subject expectation and consider the possibility that there is a gap in the non-finite phrase. In (66b) there is more material between the filler and the gap but since the structure is not excessively complex and both gap expectations are met, no major difficulty arises.

It should be clear that the function of the rule in (62) is to aid processing. This heuristic is successful in the vast majority of cases, since subject-internal gaps are virtually absent from normal discourse. In the presence of a subject-internal gap, the heuristic backfires and misleads the parser. Digging-in effects and the extra processing cost incurred by reprocessing cause a complex sentence to be even harder to process, which leads to lower acceptability. If there are no strong cues for the existence of a subject gap as discussed above, then the expectations that the subject is gapless and that the gap is located in the VP are too strong for the parser to overcome.

A heuristic-based analysis might also offer an account of adjunct islands and their parasitic and non-parasitic exceptions. If gaps are expected to reside in verb phrases rather than in their adverbial adjuncts, then we predict that non-parasitic extraction from adjuncts is rather difficult (but not impossible, under ideal conditions, which may depend on how well the structure coheres pragmatically). As in the case of complex subject phrases, the reprocessing costs caused by violating such a heuristic compound with the overall processing costs of the sentence, and create a complex landscape of graded acceptability. This is left for future study.

There is a wide range of independent psycholinguistic motivation for the existence of frequency-based parsing preferences of the kind that I advocate here. Direct evidence is provided by Vasishth & Lewis (2006), who show that sentence processing can speed up when the parser encounters certain expressions that reinforce already existing expectations. Language-specific parsing preferences also seem to be at work in relative clause attachment biases: in languages like Spanish, Dutch, French, German, and Japanese there is a preference of 'high attachment' over 'low attachment' for relative clauses (Hemforth et al. 1985, Cuetos & Mitchell 1988, Brysbaert & Mitchell 1996, Kamide & Mitchell 1997, Zagar, Pynte & Rativeau 1997). The opposite preference is observed in Arabic, English, Norwegian, Romanian and Swedish (Frazier & Clifton 1996, Ehrlich et al. 1999, Quinn, Abdelghany & Fodor 2000, Fodor 2002a). The strength of this preference also seems to vary across languages. See Fodor (2002a) for an account of these phenomena in terms of prosodic phrasing preferences.

317

Further experimental evidence for parsing preferences in language is provided by Norcliffe (2009), who shows that the asymmetric frequency distributions of gaps and resumptive pronouns within and across various languages can be explained by the conventionalization of processing preferences, along the lines of Hawkins (1999). Another example of parsing preferences and conventionalized probabilistic information in language can be observed in the ordering of NP and PP in English VPs. The canonical ordering is NP–PP, but it is often reversed if the NP is significantly longer than the PP. Wasow (2007) reasons that this canonical but violable ordering tendency should be seen as part of the grammar of English. Crucially, this is gradient information: the preference for NP–PP ordering has a magnitude. By adopting the NP–PP ordering as a default, the grammar usually adopts the ordering that allows the earliest possible identification of the categories of the daughters of the node being constructed. Conversely, if the grammar were categorical rather than gradient, then it either would have to stipulate one ordering (incorrectly ruling out the other) or say nothing about the ordering (incorrectly predicting that shorter constituents will consistently precede longer ones). For further evidence of conventionalized probabilistic information in English dative alternations see Bresnan et al. (2007) and Bresnan & Ford (2010). See also Jurafsky (1996) for an overview of the role of probabilistically-enriched grammars in models of human language and in models of language processing.

4.3 *Formalization*

There are at least two ways to implement the heuristic proposed in (62). One possibility is to extend grammar rules with frequency information that corresponds to predictive information. In other words, the grammar rules that introduce subject phrases would be augmented with a convention specifying that gapped subject phrases have virtually zero frequency (or that gapless-subject phrases are the norm). This can be stated in a straightforward way in non-transformational frameworks such as Head-driven Phrase Structure Grammar (HPSG; Pollard & Sag 1994). For example, one can augment the phrasal rules that introduce subject phrases and state that the likelihood of such phrases being [GAP {}] (i.e. gapless) is close to 1. Or, alternatively, that the likelihood of such phrases being specified as [GAP {… X …}] (i.e. containing gaps) is close to 0. See Ginzburg & Sag (2000) and Sag (2010a) for discussion about how filler–gap dependencies are modeled in HPSG. As discussed above, such a convention follows from the extremely low distributional frequency of subject-internal gaps, which in turn follows from processing and pragmatic factors. This account would allow for constructionally-specific differential biases (i.e. stronger subject island violations in some constructions and weaker in others).

Another possibility is that (62) is a parsing rule completely separate from the grammar. This approach seems to be more in line with Engdahl (1983: 29), and would not allow constructionally-specific differential biases. Again, this could be formalized straightforwardly in theories like HPSG by stating a rule like (67), where the slash '/' operator is a PERSISTENT DEFAULT constraint in the sense of Lascarides & Copestake (1999).

(67) verb-lxm⇒[VALENCE ⟨XP[GAP/{ }] …⟩]

Basically, this rule states that for every verbal lexeme, the subject valent is by default expected to be [GAP {}]. By stating the constraint over the VALENCE feature (Sag 2010b) rather than over the ARG-ST feature, we allow extraction from 'subjects' of *there*-insertion, as in examples like (68).

(68) Who was there a picture of __ on the wall?   (Stepanov 2007: 32)

As before, the default parsing constraint in (67) could be associated with frequency information, if needed. I illustrate the effect of either account in Figure 1. The filler–gap expectations are highlighted, for perspicuity.
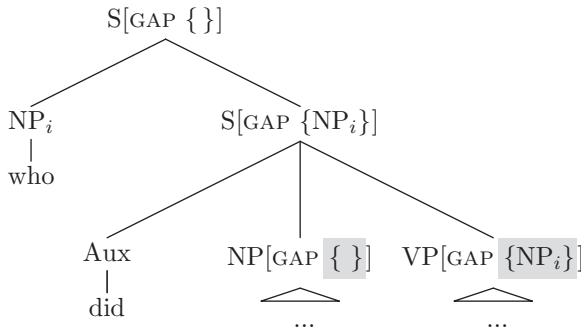

Figure 1
Application of parsing heuristics during sentence processing.

As the parser encounters the filler *who*, it creates the unbounded dependency recorded in GAP. Informally, this states that there is an NP_i gap embedded in the sentence. The information about the missing phrase percolates down in the structure as specified in Ginzburg & Sag (2000) and Sag (2010a). As the parser encounters the auxiliary verb, the (top–down) expectation that the gap is not located in the subject phrase is created. This in turn leads to the expectation that the gap must be located in the verb phrase, since there is no other logical possibility. If the subject NP turns out to not contain a gap, and there is a gap in the VP, then the expectations lead to the correct parse. This is by far the most frequent situation, which is why (62) is a useful rule-of-thumb for sentence processing. If both the NP and the VP each contain a gap, then some extra processing is needed in order to overcome the digging-in

effect and backtrack enough to link the filler to the subject- internal gap. If two gaps are co-referential and in close proximity then this backtracking is facilitated. Finally, if the subject contains a gap but the VP does not, then stronger digging-in effects arise. However, by employing prosodic cues and fillers that are highly relevant for the predicate and the subject, it is possible to boost the subject-internal gap parse and to reduce the digging-in effects, in which case backtracking is not so heavily preempted. As a consequence, the island violation can be relatively acceptable.

5. CONCLUSION

This work argues that there are robust counterexamples to the subject island constraint in English, contrary to widespread assumption. These involve the extraction of nominal, prepositional and adverbial phrases from nominal as well as verbal subjects, and include attested data. Adjunct islands are similarly prone to acceptable exceptions, even in the case of finite tensed adjuncts. These data and the graded extractions from subject islands that have occasionally been discussed in the literature by Ross (1967), Grosu (1981), Pollard & Sag (1994), Levine et al. (2001), Kluender (2004) and others should not be taken as uninteresting marginal cases. On the contrary, they suggest that modern syntactic theory has overstated the role that configurational syntax plays in subject island effects and parasitic extractions, and undermine the CED as a grammatical condition. The observed extraction patterns are problematic for all existing syntactic, pragmatic, or processing-based accounts.

The data indicate that subject islands may be an extra-grammatical phenomenon. Their processing can be facilitated by linguistic factors that are independently known to aid processing in a variety of different constructions: prosodic boundaries that draw attention to the gap site, high pragmatic coherence between the wh-phrase and the head that governs the gap (the relevance of which cues and pragmatically justifies the use of the extraction in the first place), the presence of semantically specific wh-phrases (which resist memory decay better than less specific wh-phrases), and facilitation due to reactivation of the filler caused by the presence of a second gap in close proximity (parasitism).

Drawing from insights due to Engdahl (1983), Kluender (2004), and Hawkins (2004), I propose that although extraction from subjects is grammatical, it is typically low in acceptability because parses with subject-internal gaps are strongly preempted by language-specific parsing expectations. The ameliorative effect of prosody, specificity and pragmatics can be seen at work in other complex structures such as center-embedded relative clauses, and arises because such factors cue the parser to consider the (otherwise strongly preempted) subject-internal gap parse. Such parsing expectations are ultimately a consequence of independent cognitive and

pragmatic constraints. More specifically, the significant processing burden caused by complex subject phrases with sentence-medial gaps and Gricean economy conditions lead speakers to avoid subject-internal gaps under real-time sentence processing pressure. In turn, the extremely low frequency of subject-internal gaps results in a 'no subject-internal gaps' parsing heuristic expectation. Such an expectation effectively aids processing by pruning the search space of filler–gap propagation possibilities. Hence, any gaps are instead expected to be located in the subject's sister phrase, the VP. Parasitic effects obtain from the fact that only one expectation is violated, not two, and from the facilitatory effect caused by the existence of multiple gaps in close proximity that reactivate the same filler. In that case, the gapless-subject parse is subject to a weaker 'digging-in' effect, and it becomes easier for the parser to backtrack. Nominal and finite subjects consume more cognitive resources because they are referential (Kluender 1998), which explains why the only attested cases of non-parasitic extraction from subject phrases involve non-finite subjects and why the latter are the only kind of subject known to exhibit 'filled-gap' effects (Stowe 1986, Ellis 1991, Pickering et al. 1994, Phillips 2006).

REFERENCES

Abrúsan, Márta. 2007. *Contradiction and crammar: The case of weak islands.* Ph.D. dissertation, MIT.

Ackerman, Lauren, Masaya Yoshida & Janet Pierrehumbert. 2011. Prosodic structure of center-embedded sentences. Poster presented at the 24th Conference on Human Sentence Processing, Stanford University.

Amy, Gérard & Georges Noziet. 1978. Memory requirements and local ambiguities of parsing strategies. *Journal of Psycholinguistic Research* 20.3, 233–250.

Berwick, Robert & Amy S. Weinberg. 1984. *The grammatical basis of linguistic performance.* Cambridge, MA: MIT Press.

Bloom, Lois. 1990. Subjectless sentences in child language. *Linguistic Inquiry* 21.4, 491–504.

Boeckx, Cedric A. 2008. Islands. *Language and Linguistics Compass* 2, 151–167.

Borgonovo, Claudia & Ad Neeleman. 2000. Transparent adjuncts. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 45, 199–224.

Bouma, Gosse, Robert Malouf & Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language & Linguistic Theory* 19.1, 1–65.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina & R. Harald Baayen. 2007. Predicting the dative alternation. In Gerlof Boume, Irene Kraemer & Joost Zwarts (eds.), *Cognitive foundations of interpretation*, 69–94. Amsterdam: Royal Netherlands Academy of Science.

Bresnan, Joan & Marilyn Ford. 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 81.1, 186–213.

Brysbaert, Marc & Don C. Mitchell 1996. Modifier attachment in sentence parsing: Evidence from Dutch. *The Quartely Journal of Experimental Psychology* 49A.3, 664–695.

Cattell, Ray. 1976. Constraints on movement rules. *Language* 52, 18–59.

Chaves, Rui P. 2012. On the grammar of extraction and coordination. *Natural Language & Linguistic Theory* 30.2, 465–512.

Chaves, Rui P. & Jeruen E. Dery. In press. Which subject islands will the acceptability of improve with repeated exposure? *West Coast Conference on Formal Linguistics* (WCCFL) 31, Arizona State University.

Chen, Evan, Edward Gibson & Florian Wolf. 2005. Online syntactic storage costs in sentence comprehension. *Journal of Memory and Language* 52, 144–169.

Chomsky, Noam. 1962. The logical basis of linguistic theory. In Horace Lunt (ed.), *The Ninth International Congress of Linguistics*, 914–978. The Hague: Mouton, 1964. [Reprinted as 'Current issues in linguistic theory' in *The structure of language: Readings in the philosophy of language*, edited by Jerry A. Fodor & Jerrold J. Katz, New York: Prentice-Hall, 1964; expanded in Janua Linguarum, Series Minor 38, 7–119, The Hague: Mouton, 1964.]

Chomsky, Noam. 1973. Conditions on transformations. In Stephen R. Anderson & Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, 232–286. New York: Holt, Reinhart & Winston.

Chomsky, Noam. 1977. *Essays on form and interpretation*. New York: North-Holland.

Chomsky, Noam. 1982. *Concepts and consequences of the theory of government and binding*. Cambridge, MA: MIT Press.

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

Chomsky, Noam. 2008. On phases. In Robert Freidin, David Michaels, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–165, Cambridge, MA: MIT Press.

Chung, Sandra. 1991. Sentential subjects and proper government in Chamorro. In Carol Georgopoulos & Roberta Ishihara (eds.), *Interdisciplinary approaches to language: Essays in honor of S.-Y. Kuroda*, 75–99. Dordrecht: Kluwer.

Chung, Sandra. 1994. *Wh*-agreement and 'referentiality' in Chamorro. *Linguistic Inquiry* 25, 1–44.

Chung, Sandra & James McCloskey. 1983. On the interpretation of certain island facts in GPSG. *Linguistic Inquiry* 14, 703–714.

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

Clark, Herbert H. & Thomas Wasow. 1998. Repeating words in spontaneous speech. *Cognitive Psychology* 37, 201–242.

Clausen, David R. 2010. Processing factors influencing acceptability in extractions from complex subjects. Ms., Stanford University.

Clausen, David R. 2011. Informativity and acceptability of complex subject islands. Poster presented at the 24th Annual CUNY Sentence Processing Conference, Stanford University.

Crawford, Jean. 2011a. Syntactic satiation of subject islands. Poster presented at the 24th Annual CUNY Sentence Processing Conference, Stanford University.

Crawford, Jean. 2011b. Using syntactic satiation effects to investigate subject islands. In Jaehoon Choi, E. Alan Hogue, Jeffrey Punske, Deniz Tat, Jessamyn Schertz & Alex Trueman (eds.), *West Coast Conference on Formal Linguistics* (WCCFL) 29, 38–45. Somerville, MA: Cascadilla Proceedings Project.

Cuetos, Fernando & Don C. Mitchell. 1988. Cross-linguistic differences in parsing: Restrictions on the use of the late closure strategy in Spanish. *Cognition* 30, 73–105.

Culicover, Peter W. 1999. *Syntactic nuts: Hard cases in syntax* (Foundations of Syntax 1). Oxford: Oxford University Press.

Culicover, Peter W. & Paul M. Postal (eds.). 2001. *Parasitic gaps*. Cambridge, MA: MIT Press.

Deane, Paul D. 1991. Limits to attention: A cognitive theory of island phenomena. *Cognitive Linguistics* 2.1, 1–63.

Deane, Paul D. 1992. *Grammar in mind and brain*. Berlin: Mouton de Gruyter.

Eady, Stephen J. & Janet Dean Fodor. 1981. Is center embedding a source of processing difficulty? Presented at the Linguistics Society of America Annual Meeting, New York.

Ehrlich, Karen, Eva M. Fernández, Janet Dean Fodor, Eric Stenshoel & Mihai Vinereanu. 1999. Low attachment of relative clauses new data from Swedish, Norwegian and Romanian. Poster presented at the 12th Conference on Human Sentence Processing, New York.

Elliot, Malcom, James Kirby, Osamu Sawada, Eleni Staraki & Suwon Yoon (eds.). 2007. *Chicago Linguistic Society* (CLS) 43.1. Chicago, IL: Chicago Linguistic Society.

Ellis, Rod. 1991. Grammaticality judgments and second language acquisition. *Studies in Second Language Acquisition* 13.2, 161–186.

Engdahl, Elisabet. 1983. Parasitic gaps. *Linguistics and Philosophy* 6, 3–34.

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

Erteschik-Shir, Nomi. 1981. More on extractability from quasi-NPs. *Linguistic Inquiry* 12, 665–670.

Erteschik-Shir, Nomi. 2006. What's what? In Fanselow et al. (eds.), 317–335.

Erteschik-Shir, Nomi. 2007. *Information structure: The syntax–discourse interface*. Oxford: Oxford University Press.

Erteschik-Shir, Nomi & Shalom Lappin. 1979. Dominance and the functional explanation of island phenomena. *Theoretical Linguistics* 6, 41–86.

Fanselow, Gisbert, Caroline Féry, Ralf Vogel & Matthias Schlesewsky (eds.). 2006. *Gradience in grammar*. Oxford: Oxford University Press.

Fernández, Ángel Jiménez. 2009. On the composite nature of subject islands: A phase-based approach. *SKY Journal of Linguistics* 22, 91–138.

Ferreira, Fernanda. 1991. Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language* 30.2, 210–233.

Ferreira, Fernanda & John M. Henderson. 1991. Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language* 31, 725–745.

Ferreira, Fernanda & John M. Henderson. 1993. Basic reading processes during syntactic analysis and reanalysis. *Reading and Language Processing:* Special Issue of *Canadian Journal of Psychology* 47, 247–275.

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.

Fodor, Janet Dean. 1992. Islands, learnability and the lexicon. In Goodluck & Rochemont (eds.), 109–180.

Fodor, Janet Dean. 2002a. Prosodic disambiguation in silent reading. In Mako Hirotani (ed.), *Northeast Linguistic Society* (NELS) 32, 113–132, Amherst, MA: GLSA.

Fodor, Janet Dean. 2002b. Psycholinguistics cannot escape prosody. In Bernard Bel & Isabelle Marlien (eds.), *Speech Prosody 2002*, Université de Provence, Aix-en-Provence, 83–88.

Francom, Jerid. 2009. *Experimental syntax: Exploring the effect of repeated exposure to anomalous syntactic structure – evidence from rating and reading tasks*. Ph.D. thesis, University of Arizona.

Frank, Robert E. 1992. *Syntactic locality and Tree Adjoining Grammar: Grammatical, acquisition and processing perspectives*. Ph.D. dissertation, University of Pennsylvania.

Frazier, Lyn. 1987. Syntactic processing: Evidence from Dutch. *Natural Language & Linguistic Theory* 5, 519–559.

Frazier, Lyn & Carles Clifton Jr. 1996. *Construal*. Cambridge, MA: MIT Press.

Garnsey, Susan M. 1985. *Function words and content words: Reaction time and evoked potential measures of word recognition* (Cognitive Science Technical Report URCS-29). Rochester, NY: University of Rochester.

Garnsey, Susan M., Michael K. Tanenhaus & Robert M. Chapman. 1989. Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research* 18, 51–60.

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68, 1–76.

Gibson, Edward. 2000. The Dependency Locality Theory: A distance-based theory of linguistic complexity. In Alec Marantz, Yasushi Miyashita & Wayne O'Neil (eds.), *Image, language, brain*, 95–126. Cambridge, MA: MIT Press.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning and use of English interrogative constructions*. Stanford, CA: CSLI Publications.

Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.

Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Goodluck, Helen & Michael Rochemont (eds.). 1992. *Island constraints: Theory, acquisition and processing*. Dordrecht: Kluwer.

Grosu, Alexander. 1981. *Approaches to island phenomena*. Amsterdam: North-Holland.

Hawkins, John A. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press.

Hawkins, John A. 1999. Processing complexity and filler–gap dependencies across grammars. *Language* 75.2, 244–249.

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

323

Hegarty, Michael. 1990. On adjunct extraction from complements. In Lisa Cheng & Hamida Demirdash (eds.), *Papers on* wh-*movement* (MIT Working Papers in Linguistics 13), 101–124. Cambridge, MA: MIT.

Hemforth, Barbara, Lars Konieczny, Christoph Scheepers & Gerhard Strube. 1985. Syntactic ambiguity resolution in German. In Dieter Hillert (ed.), *Syntax and semantics*, vol. 31, 293–312. San Diego, CA: Academic Press.

Hiramatsu, Kazuko. 1999. Subject and adjunct island asymmetries: Evidence from syntactic satiation. In Sonya Bird, Andrew Carnie, Jason D. Haugen & Peter Norquest (eds.), *West Coast Conference on Formal Linguistics* (WCCFL) 18, 183–192. Sommerville, MA: Cascadilla Press.

Hofmeister, Philip. 2007. Retrievability and gradience in filler–gap dependencies. In Elliott et al. (eds.), 109–123.

Hofmeister, Philip & Ivan A. Sag. 2010. Cognitive constraints and island effects. *Language* 86.2, 366–415.

Hofmeister, Philip, Laura Staum Casasanto & Ivan A. Sag. 2012a. How do individual cognitive differences relate to acceptability judgments? A reply to Sprouse, Wagers, and Phillips. *Language* 88.2, 390–400.

Hofmeister, Philip, Laura Staum Casasanto & Ivan A. Sag. 2012b. Misapplying working-memory tests: A reductio ad absurdum. *Language* 88.2, 408–409.

Hofmeister, Philip, Laura Staum Casasanto & Ivan A. Sag. In press. Islands in the grammar? Standards of evidence. In Jon Sprouse & Norbert Hornstein (eds.), *Experimental syntax and the islands debate*. Cambridge: Cambridge University Press.

Hornstein, Norbert, Howard Lasnik & Juan Uriagereka. 2006. The dynamics of islands: Speculations on the locality of movement. *Linguistic Analysis* 33.1–2, 149–175.

Horvath, Julia. 1992. The anti c-command and case-compatibility in the licensing of parasitic chains. *The Linguistic Review* 9, 183–218.

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

Huddleston, Rodney, Geoffrey K. Pullum & Peter Peterson. 2002. Relative clause constructions and unbounded dependencies. In Rodney Huddleston, Geoffrey K. Pullum et al., *The Cambridge Grammar of the English Language*, 1031–1096. Cambridge: Cambridge University Press.

Hukari, Thomas & Robert D. Levine. 1991. On the disunity of unbounded dependency constructions. *Natural Language & Linguistic Theory* 9, 97–144.

Jackendoff, Ray S. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

Jurafsky, Daniel. 1996. A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science* 2.20, 137–194.

Jurka, Johannes, Chizuru Nakao & Akira Omaki. 2011. It's not the end of the CED as we know it: Revisiting German and Japanese subject islands. In Mary Byram Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer & Barbara Tomaszewicz (eds.), *West Coast Conference on Formal Linguistics* (WCCFL) 28, 124–132. Somerville, MA: Cascadilla Proceedings Project.

Kamide, Yuki & Don C. Mitchell 1997. Relative clause attachment: Non-determinism in Japanese parsing. *Journal of Psycholinguistic Research* 26, 247–254.

Kayne, Richard S. 1981. ECP extensions. *Linguistic Inquiry* 12, 93–133.

Kayne, Richard S. 1983. Connectedness. *Linguistic Inquiry* 14, 223–249.

Kehler, Andrew. 2002. *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI Publications.

Kemper, Susan. 1986. Imitation of complex syntactic constructions by elderly adults. *Applied Psycholinguistics* 7, 277–288.

Kemper, Susan. 1987. Life-span changes in syntactic complexity. *Journal of Gerontology* 42.3, 323–328.

Kimball, John. 1975. Predictive analysis and over-the-top parsing. In John Kimball (ed.), *Syntax and semantics*, vol. 4, 155–179. New York: Academic Press.

Kitagawa, Yoshihisa & Janet Dean Fodor. 2006. Prosodic influence on syntactic judgements. In Fanselow et al. (eds.), 336–354.

Kluender, Robert. 1992. Deriving island constraints from principles of predication. In Goodluck & Rochemont (eds.), 223–258.

Kluender, Robert. 1998. On the distinction between strong islands and weak islands: A processing perspective. In Peter W. Culicover & Louise McNally (eds.), *The limits of syntax* (Syntax and Semantics 29), 241–279, New York: Academic Press.

Kluender, Robert. 2004. Are subject islands subject to a processing account? In Vineeta Chand, Ann Kelleher, Angelo J. Rodríguez & Benjamin Schmeiser (eds.), *West Coast Conference on Formal Linguistics* (WCCFL) 23, 101–125, Somerville, MA: Cascadilla Press.

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

Kravtchenko, Ekaterina, Maria Polinsky & Ming Xiang. 2009. Are all subject islands created equal? Poster at the 22nd CUNY Conference on Human Sentence Processing, Davis, CA.

Kroch, Anthony. 1989. Amount quantification, referentiality, and long *wh*-movement. Ms., Department of Linguistics, University of Pennsylvania.

Kuno, Susumu. 1973a. Constraints on internal clauses and sentential subjects. *Linguistic Inquiry* 4, 363–386.

Kuno, Susumu. 1973b. *The structure of the Japanese language*. Cambridge, MA: MIT Press.

Kuno, Susumu. 1987. *Functional syntax – anaphora, discourse and empathy*. Chicago, IL & London: The University of Chicago Press.

Kuno, Susumu & Ken-Ichi Takami. 1993. *Grammar and discourse principles: Functional syntax and GB theory*. Chicago, IL & London: The University of Chicago Press.

Kutas, Marta, Mireille Besson & Cyma Van Petten. 1988. Event-related potential asymmetries during the reading of sentences. *Electroencephalography and Neurophysiology* 69, 218–233.

Kynette, Donna & Susan Kemper. 1986. Aging and the loss of grammatical forms: A cross-sectional study of language performance. *Language and Communication* 6, 65–72.

Lascarides, Alex & Anne Copestake. 1999. Default unification in constraint-based frameworks. *Computational Linguistics* 25.1, 55–105.

Lasnik, Howard & Myung-Kwan Park. 2003. The EPP and the subject condition under sluicing. *Linguistic Inquiry* 34, 649–660.

Lasnik, Howard & Mamoru Saito. 1992. *Move α: Conditions on its application and output*. Cambridge, MA: MIT Press.

Levine, Robert D. 2001. The extraction riddle: Just what are we missing? *Journal of Linguistics* 37, 145–174.

Levine, Robert D. & Thomas E. Hukari. 2006. *The unity of unbounded dependency constructions*. Stanford, CA: CSLI Publications.

Levine, Robert D., Thomas E. Hukari & Michael Calcagno. 2001. Parasitic gaps in English: Some overlooked cases and their theoretical implications. In Culicover & Postal (eds.), 181–222.

Levine, Robert D. & Ivan A. Sag. 2003. Some empirical issues in the grammar of extraction. In Stefan Müller (ed.), *HPSG-2003 Conference*, Michigan State University, East Lansing, 236–256. Stanford, CA: CSLI Publications.

Morgan, Jerry L. 1973. Sentence fragments and the notion of sentence. In Braj B. Kachru, Robert B. Lees, Yakov Malkiel, Angelina Pietrangeli & Sol Sapotra (eds.), *Issues in linguistics: Papers in honor of Henry and Renee Kahane*, 719–751. Urbana, IL: University of Illinois Press.

Müller, Gereon. 2010. On deriving CED effects from the PIC. *Linguistic Inquiry* 41.1, 35–82.

Norcliffe, Elisabeth. 2009. *Head-marking in usage and grammar: A study of variation and change in Yucatec Maya*. Ph.D. dissertation, Stanford University.

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

Pesetsky, David. 2000. *Phrasal movement and its kin*. Cambridge, MA: MIT Press.

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

Pickering, Martin J., Stephen Barton & Richard Shillcock. 1994. Unbounded dependencies, island constraints and processing complexity. In Charles Clifton Jr., Lyn Frazier & Keith Rayner (eds.), *Perspectives on sentence processing*, 199–224. London: Lawrence Erlbaum.

Pollard, Carl & Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Chicago, IL: The University of Chicago Press & Stanford, CA: CSLI Publications.

Postal, Paul M. 1974. *On raising*. Cambridge, MA: MIT Press.

Postal, Paul M. 1993. Parasitic gaps and the across-the-board phenomenon. *Linguistic Inquiry* 24, 735–754.

Postal, Paul M. 1994. Parasitic and pseudoparasitic gaps. *Linguistic Inquiry* 25, 63–117.

325

Postal, Paul M. 1998. *Three investigations of extraction*. Cambridge, MA: MIT Press.

Pustejovsky, James. 1995. *The generative lexicon*. Cambridge, MA: MIT Press.

Quinn, Deirdre, Hala Abdelghany & Janet Dean Fodor. 2000. More evidence of implicit prosody in silent reading: French, English and Arabic relative clauses. Poster presented at the 13th Conference on Human Sentence Processing, La Jolla, CA.

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

Rogers, James & Geoffrey K. Pullum. 2011. Aural pattern recognition experiments and the subregular hierarchy. *Journal of Logic, Language and Information* 20, 329–342.

Ross, John R. 1967. *Constraints on variables in syntax*. Ph.D. dissertation, MIT. [Published 1986 as *Infinite Syntax!*, Norwood, NJ: Ablex.]

Sabel, Joachim. 2002. A minimalist analysis of syntactic islands. *The Linguistic Review* 19, 271–315.

Sag, Ivan A. 1992. Taking performance seriously. In Carlos Martin-Vide (ed.), *VII Congreso de Languajes Naturales y Languajes Formales*, Barcelona, 61–74.

Sag, Ivan A. 2010a. English filler–gap constructions. *Language* 86.3, 486–545.

Sag, Ivan A. 2010b. Sign-based Construction Grammar: An informal synopsis. Ms., Stanford University. Available at http://lingo.stanford.edu/sag/papers/theo-syno.pdf (retrieved 9 October 2012). [To appear in Hans Boas & Ivan A. Sag (eds.), *Sign-based Construction Grammar*. Stanford, CA: CSLI Publications.]

Sag, Ivan A., Philip Hofmeister & Neal Snider. 2007. Processing complexity in subjacency violations: The complex noun phrase constraint. In Elliot et al. (eds.), 215–229.

Santorini, Beatrice. 2007. (Un?)expected movement. Ms., University of Pennsylvania. http://www.ling.upenn.edu/~beatrice/examples/movement.html, retrieved 2 January 2012.

Sauerland, Uli & Paul D. Elbourne. 2002. Total reconstruction, PF movement, and derivational order. *Linguistic Inquiry* 33.2, 283–319.

Sells, Peter. 1984. *Syntax and semantics of resumptive pronouns*. Ph.D. dissertation, University of Massachusetts, Amherst.

Sheehan, Michelle. 2009. The resuscitation of CED? In Seda Kan, Claire Moore-Cantwell & Robert Staubs (eds.), *North East Linguistic Society* (NELS) 40. Amherst, MA: GLSA.

Shimojo, Mitsuaki. 2002. Functional theories of island phenomena: The case of Japanese. *Studies in Language* 26, 67–123.

Snyder, William. 2000. An experimental investigation of syntactic satiation effects. *Linguistic Inquiry* 31, 575–582.

Sprouse, Jon. 2009. Revisiting satiation: Evidence for an equalization response strategy. *Linguistic Inquiry* 40.2, 329–341.

Sprouse, Jon, Matt Wagers & Colin Phillips. 2012a. A test of the relation between working memory capacity and syntactic island effects. *Language* 88.1, 82–123.

Sprouse, Jon, Matt Wagers & Colin Phillips. 2012b. Working-memory capacity and island effects: A reminder of the issues and the facts. *Language* 88.2, 401–407.

Steedman, Mark. 1996. *Surface structure and interpretation*. Cambridge, MA: MIT Press.

Stepanov, Arthur. 2007. The end of CED? Minimalism and extraction domains. *Syntax* 10.1, 80–126.

Stowe, Lauriea. 1986. Parsing WH-constructions: Evidence for on-line gap location. *Language and Cognitive Processes* 1, 227–245.

Szabolcsi, Anna. 2006. Strong vs. weak islands. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*, vol. 4, 479–531. Oxford: Blackwell.

Szabolcsi, Anna & Marcel den Dikken. 1999. Islands. *GLOT International* 4/6, 3–8.

Tabor, Whitney & Sean Hutchins. 2004. Evidence for self-organized sentence processing: Digging-in effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30.2, 431–450.

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

Taraldsen, Knut Tarald. 1980. The theoretical interpretation of a class of marked extractions. In Adriana Belleti, Luciana Brandi & Luigi Rizzi (eds.), *The Third GLOW Conference*, 475–516. Pisa: Annali della Scuola Normale Superiore.

Taylor, Heather Lee. 2007. Movement from IF-clause adjuncts. In Anastasia Conroy, Chunyuan Jing, Chizuru Nakao & Eri Takahashi (eds.), *University of Maryland Working Papers in Linguistics* (UMWPiL) 15, 192–206. College Park, MD: Department of Linguistics, University of Maryland.

Traxler, Matthew J. & Martin J. Pickering. 1996. Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language* 35, 454–475.

Truswell, Robert. 2011. *Events, phrases and questions*. Oxford: Oxford University Press.

Tsiamtsiouris, Jim & Helen Smith Cairns. 2009. Effects of syntactic complexity and sentence-structure priming on speech initiation time in adults who stutter. *Journal of Speech, Language, and Hearing Research* 52.6, 1623–1639.

Uriagereka, Juan. 1999. Multiple spell-out. In Samuel D. Epstein & Norbert Hornstein (eds.), *Working Minimalism*, 251–282. Cambridge, MA: MIT Press.

Van Petten, Cyma & Marta Kutas. 1991. Influences of semantic and syntactic context on open and closed class words. *Memory and Cognition* 19, 95–112.

Van Valin, Robert D. Jr. 1986. Pragmatics, island phenomena, and linguistic competence. In Anne M. Farley, Peter T. Farley & Karl-Erik McCullough (eds.), *Chicago Linguistic Society (CLS) 22.2: The Parasession on Pragmatics and Grammatical Theory*, 223–233. Chicago, IL: Chicago Linguistic Society.

Van Valin, Robert D. Jr. 1995. Toward a functionalist account of so-called extraction constraints. In Betty Devriendt, Louis Goossens & Johan van der Auwera (eds.), *Complex structures: A functionalist perspective*, 26–60. Berlin: Mouton de Gruyter.

Vasishth, Shravan & Richard L. Lewis. 2006. Argument-head distance and processing complexity: Explaining both locality and antilocality effects. *Language* 82.4, 767–794.

Wanner, Eric & Michael Maratsos. 1978. An ATN approach to comprehension. In Morris Hale, Joan Bresnan & George A. Miller (eds.), *Linguistic theory and psychological reality*, 119–161. Cambridge, MA: MIT Press.

Wasow, Thomas. 1997. Remarks on grammatical weight. *Language Variation and Change* 9, 81–105.

Wasow, Thomas. 2002. *Postverbal behavior*. Stanford, CA: CSLI Publications.

Wasow, Thomas. 2007. Gradient data and gradient grammars. In Elliot et al. (eds.), 255–271.

Wexler, Kenneth & Peter W. Culicover. 1980. *Formal principles of language acquisition*. Cambridge, MA: MIT Press.

Yngve, Victor. 1960. A model and an hypothesis for language structure. *The American Philosophical Society* 104, 444–466.

Zagar, Daniel, Joel Pynte & Sylvie Rativeau. 1997. Evidence for early closure attachment on first pass reading times in French. *The Quarterly Journal of Experimental Psychology* 50A.2, 421–438.

Zahn, Daniela & Christoph Scheepers. 2011. Task effects in resolving RC attachment ambiguities. Poster presented at the 24th Conference on Human Sentence Processing, Stanford University.

Author's address: Linguistics Department, University at Buffalo,
609 Baldy Hall, Buffalo, NY 14260, USA
rchaves@buffalo.edu