7 Memory mechanisms for *wh*-dependency formation and their implications for islandhood

Matthew W. Wagers

1 Introduction

The goal of this chapter is to provide a functionally specified model of how long-distance dependencies are understood in real time. In particular its emphasis will be how long-distance dependencies interact with working memory (WM) and how that interaction might inform our understanding of the form or origin of grammatical constraints. By long-distance dependencies, I refer primarily to unbounded displacement dependencies, such as those exhibited in *wh*-questions, relative clauses, topicalizations, comparative clauses, clefts, etc. The grammar appears to place no limit on the hierarchical distance that such dependencies can span. The head of the dependency – the displaced constituent itself – can occur many clauses away from the tail of the dependency: a gap[1] or pronoun. The example in (1) illustrates this basic observation with relativization: the bold-face constituent, *cookies*, is only one clause distant from its gap in (1a). In (1b) and (1c), however, it is two and three clauses away, respectively.

(1) a. The **cookies** [$_S$ that Phil bakes ___] contain pistachios.
b. The **cookies** [$_S$ that Billy bothers Phil to bake ___**]** contain pistachios.
c. The **cookies** [$_S$ that everyone knows that Billy bothers Phil to bake ___] contain pistachios.

There has long been a mutual influence between the study of long-distance dependencies and the study of WM constraints in language processing (Miller and Chomsky 1963). Between its head and tail, a large number of processing events could occur which are essentially irrelevant to the long-distance dependency itself. At the syntactic level, these include events like establishing other thematic dependencies, interpreting modifier relations, checking case and agreement, or resolving anaphora. The processing system must thus have

[1] The use of the terms "gap" or "gap site" is here intended to be neutral between grammars which posit an unpronounced constituent (a trace or copy; Brody 1995, Chomsky 1995, Stabler 2000, Frank 2002) and those which combine the displaced constituent with its subcategorizing host in other ways (e.g., Steedman 2000, Bresnan 2001, Sag *et al.* 2003). In my view, psycholinguistic theories and data make no useful distinction between the alternatives at present (Phillips and Wagers 2007).

a means of retaining information about the head of the long-distance dependencies until its "tail" can be constructed, at which point it must be able to effectively recover the head. And it must do so without suspending the intervening processing events that span the dependency. These requirements describe exactly the sort of cognitive juggling act that working memory systems are proposed to accomplish (e.g., Baddeley 1986, Miyake and Shah 1999).

An important question to ask about how linguistic information interacts with WM is whether the grammar can generate structures that overwhelm WM capacity. For the moment, we will speak of capacity roughly as the amount of information that can be encoded and later recovered in a relatively loss-free fashion, as well as the extent of time for which that information can be maintained. Whether it is possible to give a general or useful characterization of WM capacity has driven considerable research in cognitive psychology for much of its modern history (Miller 1956, Cowan 2005). For language, we can imagine trivial examples that might overwhelm WM capacity by dint of their length. For example, I take it that no one can sensibly interpret a 256-clause sentence. The more interesting examples, however, feature syntactic or semantic complexities that seem to resist comprehension, even when the sentence is short. The prime example is the center self-embedded sentence:

(2) The cookies that the dog that Phil scolded tasted were burned.

Few speakers of English find (2) to be an acceptable sentence, even though it is straightforwardly generated by the language's phrase structure rules. However, if we remove just one layer of embedding, the sentence becomes unremarkably acceptable:

(3) The cookies that the dog tasted were burned.

The extreme unacceptability of center self-embeddings is believed to stem from how the application or recognition of grammatical rules is constrained by processing capacity, though there are a diversity of proposals for the exact nature and locus of such a constraint (Yngve 1961, Frazier and Fodor 1978, Stabler 1994, Lewis 1996, Gibson and Thomas 1999, Vasishth *et al.* 2010, among many others).

We can raise an analogous question about *wh*-movement dependencies. A minor modification of sentence (1b) transforms it from an acceptable sentence to an unacceptable one:

(4) The cookies that Billy bothers Phil after he bakes ____ contain pistachios.

The unacceptability of a sentence like (4) has been standardly attributed in generative grammar to the violation of an island constraint. In this case, the dependency spans the boundary of an adjunct clause and this violates a condition on extraction (Huang 1982a). However, by analogy to center self-embedded sentences, we can ask whether some island-violating sentences might be freely generable by the grammar, yet unable to be processed during comprehension because of a capacity constraint. A number of proposals have related the unacceptability of particular island condition sentences to a confluence of factors, among which WM capacity figures strongly (Givón 1979, Deane 1991, Kluender and Kutas 1993b, Kluender 2004, Hofmeister and Sag 2010, among others; see chapters 2, 3, 4, and 8 in this volume). In order to evaluate these theories and their competitors, it is important to have a precise notion of how WM capacity is related to language comprehension and how long-distance dependencies do or do not strain it.[2]

In the next section, I will present a theory of working memory which has been emerging as a consensus among many memory theorists in the past decade (Nairne 2002, McElree 2006, Jonides *et al.* 2008) as well as some psycholinguists (Vasishth and Lewis 2005). It has two interesting features: Firstly, it does not assume there is any strong mechanistic discontinuity between memory in the short term and memory in the long term. The means of retention and retrieval are largely the same, whether they take place on the timescale of 500 milliseconds or 15 days. As a consequence, this theory of WM largely eschews specialized buffers and storage subsystems that were central in other theories (Baddeley 1986). Secondly, the mechanisms for searching and retrieving information are optimized for the inherent features of stored representations and not the relations that hold between them (McElree 2000, Van Dyke 2007). This is, at first glance, problematic for language processing since constraints on grammatical dependencies are often characterized in both terms. For example, identifying the appropriate antecedent for a verb's reflexive argument is constrained both by independent properties of a potential antecedent ("feminine and singular and a noun phrase") and relational ones (e.g., "closest, c-commanding clause-mate"). In section 3, I will describe an empirically grounded model for processing *wh*-dependencies which is compatible with the unitary model of memory. Finally, in section 4, I will attempt to identify whether any interactions in the model could support an account of island constraints which reduces, in part, to working memory capacity or efficiency.

But let me first preview where the reader of this paper will end up. The question of how dependencies might strain the working memory system can be decomposed into two questions: what components of the dependency are required to be actively maintained, and what components must be recovered later via retrieval mechanisms. The argument developed in section 4, which addresses the problem of islandhood, takes the form of a "threat assessment":

[2] Before doing so, it is important to stress that WM capacity is only one piece of the puzzle in understanding why certain sentences are easy to process and why certain sentences are difficult. It is, however, a necessary piece of the puzzle.

how strong is the case that the WM system can be overwhelmed by dependency completion in the major island contexts? The threat, I believe, is too weak and too diffuse to heavily implicate a strain on the working memory system in the low acceptability of island-violating unbounded dependencies. I conclude that strong evidence is lacking that much information, if any, pertaining to the dependency is actively maintained across its span. Therefore, if islandhood depends on WM difficulty, the explanatory burden must rest with retrieval difficulty. There are several reasons to be skeptical, however, that the retrieval difficulty is severe enough to provide a useful explanation of islandhood. Some of these reasons are more conceptual in nature – such as analogy to the observation that retrieval pathways can be optimized when behavior is well practiced or routinized (Ericsson and Kintsch 1995). Other reasons are more directly empirical – such as the fact that direct manipulations of retrieval difficulty have produced measurable but quite small effects on long-distance dependency completion (Van Dyke and McElree 2006), or the fact that measures of WM efficiency show no correlation with the perceived severity of island violations (Sprouse *et al.* 2012).

Of course the validity of any threat assessment depends not only on the "known knowns" but also the "known unknowns" (and, more perniciously, the "unknown unknowns"; Rumsfeld 2002). For this reason I try to highlight throughout where my assessment could be misguided, and in particular, what kinds of data are missing that might improve our understanding or shift the balance of explanation.

2 Working memory

2.1 Introduction

Miller's famous 1956 paper discussed a variety of experiments testing the "span of immediate memory." In these experiments (Hayes 1952, Pollack 1953), adult participants were challenged to recall lists of various symbols: binary digits, decimal digits, letters, and words. Regardless of symbol type, there was a remarkably narrow range of variation in performance: recall was generally quite accurate until list length exceeded 5–9 items (i.e., the "magic number" 7 ± 2). This was striking to Miller because each of the symbol types conveyed different amounts of information: for example, a decimal digit conveys approximately 3.3 bits of information, while an English word conveys approximately 10 bits. Therefore he concluded that the capacity limitation on immediate memory was stated not in terms of information conveyed but in terms of a limited number of task-relevant encodings that could be successfully maintained and recalled in the short-term. Cowan (2005) has since argued that four is a more accurate estimate of the typical span of immediate memory for a variety of tasks.

The impact of Miller's paper on linguistic research was immediate. Perhaps most notably, it motivated Yngve's (1961) depth hypothesis, a proposal that the unacceptability of double center self-embedding stemmed from a restriction on the depth of the parser's stack.[3] But the question immediately arises whether span sizes translate directly into an architectural notion, like number of buffers or slots in a stack. Here it is useful to draw a distinction between two concepts: (1) "working memory," broadly, as the sum of the many component parts and processes that support the encoding, retention, and recollection of recently encountered information; (2) "working memory," narrowly, as the particular mechanism for maintaining a specific piece of information over the shortest term.[4] Inferring from a span number, be it 4 or 7, that there are a number of distinguished memory cells makes a commitment to the second concept. But it is logically possible that a relatively constant span derives from the first concept: that is, it derives from the interaction of many components and does not directly reflect the read-out of *n* working memory buffers, where *n* is the measure of span.

In recent years much evidence has accrued that supports the latter interpretation of span (see Conway *et al.* 2007). In most cognitive tasks, the amount of information that can be concurrently maintained and made directly available to ongoing processing is extremely limited (Broadbent 1958, McElree and Dosher 1989, Cowan 1995, 2005, McElree 2006, Jonides *et al*. 2008), smaller than the number of items that can be successfully recalled in a span task (McElree 2006). This small amount of information is accessible to ongoing cognitive processes with effectively very little delay. Information in this state is said to be in the *focus of attention*.[5] Information outside of this state must be restored to the focus of attention to be useful, a process referred to very generally as retrieval (Anderson and Neely 1996). Whether or not information needs to be retrieved is of considerable functional significance, since retrieved information takes more time to impact processing and retrieval is prone to error. There is consequently a functionally important interaction between the maintenance of the readily accessible focal representations and the retrieval of other representations to displace or transform the current contents of focal attention.

[3] The depth hypothesis is a grounded account of a grammatical constraint (see, e.g., Fodor 1978) since Yngve proposed to incorporate the depth-of-embedding limitation in the grammar as an adaptation to the memory limitation.

[4] This is sometimes called, more simply, short-term memory or STM. However, there seems to me to be enough confusion in psycholinguistics over the use of these terms that no terminological distinction will be introduced here.

[5] There is a strong analogy between a focal/non-focal split in information state and James's (1890) distinction between primary memory, evocatively called "the trailing edge of the conscious present," and secondary memory.

In the next three sections we will attempt to untangle this interaction by answering the following questions:

(5) **Q1** When must information be retrieved?
Q2 How is information retrieved?
Q3 What factors determine the success of retrieval?

2.2 *Question 1: When must information be retrieved?*

Q1 amounts to the narrow capacity question: How much information can be concurrently maintained in focal attention? The greater the capacity of focal attention, the less often its current contents will have to be shunted to make way for new information. Therefore the capacity of focal attention contributes to the expected frequency with which memory retrieval operations occur. There is broad agreement that focal capacity is restricted. Specific estimates vary depending on task and stimulus structure (Cowan 1995, 2001, Garavan 1998, McElree 2001, 2006, Oberauer 2002). McElree and colleagues have argued that only one task-relevant representation is typically maintained in focal attention (McElree and Dosher 1989, McElree 2006, Wagers and McElree 2011). Information not contained in that representation can only directly influence processing if it is retrieved to replace the focal representation. Their evidence comes from measuring the dynamics with which participants recognize or recall recently encountered data. This can best be illustrated with a concrete example from that research.

In a series of experiments, McElree (1996) asked participants to study five-word lists, presented word by word. After the final word, a visual mask was first displayed followed by a test probe. Depending on the trial, participants judged either whether the test probe was in the memory list, whether it rhymed with an item in the list, or whether it was a synonym of an item in the list. Responses were collected at a variety of lags after presentation of the test probe, so that the point at which information begins accumulating and the rate at which that accumulation occurs could be measured. This approach, known as a response-signal method, measures the speed–accuracy trade-off (SAT) participants make in giving a response. It is more revealing than simply collecting reaction times (RTs), because RTs are a unidimensional measure of processing efficiency which conflate the speed at which cognitive processes run with the accuracy criteria that participants (implicitly) set in completing tasks (Wickelgren 1976). What was consistent among all the trial types in McElree's experiment was that the most recently presented word not only achieved highest accuracy but it began its rise to accuracy the soonest. The remainder of the words in the list, regardless of their exact serial position, began their rise to asymptotic accuracy at the same later time, about 200 ms later. These data make the case that what matters for obtaining the speed advantage is that the test response depend on the last *task-relevant* representation to occupy focal attention.[6] For the rhyme test, focal attention must include information about the word's phonology. For the synonym test, it must include information about the word's semantic features. Note that these results do not necessarily imply that all conceivable information about a word was available – upon presentation of the mask, participants were cued about which judgment was required so they could have transformed the last representation accordingly.

Finally, several findings indicate that the speed advantage is not uniquely linked to an item's being in final position. If experimental procedures are used which encourage participants to rehearse items from particular list positions immediately prior to test, then the focal advantage accrues to those items (McElree 2006). Moreover, the structure of the word lists matters. If a word list can be parsed into multiple categories – for example, names for furniture and names for flowers – then the focal advantage accrues to the most recently encountered category, not the most recently encountered name (McElree 1998). Finally, evidence for focal attention comes from data other than SAT studies, including RT distributional analyses (Oberauer 2002, 2006) and fMRI studies of activation in hippocampus and inferior frontal gyrus (Öztekin *et al.* 2008; see Cabeza *et al.* 2003).

2.3 Question 2: How is information retrieved?

Q2 asks for the mechanism by which other previously constructed representations are restored to the focus of attention. The answer to **Q2** depends in part on the architecture of memory: that is, what lies "beyond" the focus of attention. McElree (2006) draws a distinction between bipartite and tripartite working memory architectures. In bipartite architectures, information can only occupy two states: *active*, in the focus of attention, or *passive*, in its durably encoded long-term memory (LTM) state. In tripartite architectures, an intermediate state exists for information not currently being processed which is nonetheless not simply in LTM. For some models, like Baddeley (1986), this intermediate state corresponds to a short-term memory store that is separate from long-term memory – like a buffer. Other models, like Cowan (1995, 2001) or Oberauer (2002), reject a separate short-term store but claim that a small amount of recently used information is so highly activated as to be functionally distinct from other LTM encodings.

The bipartite architecture makes a clear prediction about the mechanisms of retrieval. Because encodings outside of the focus of attention are in the same

[6] Additionally, the fact that study and test were interrupted by a visual mask allows us to dismiss any low-level perceptual account of the advantage.

state as LTM, it is predicted that the retrieval of recently encoded information will proceed in largely the same way as information encoded in LTM. Though it is not likely the case that there is a "single mechanism" of retrieving from LTM, the dominant mechanism appears to be associative, or content-addressable, retrieval (Murdock 1982, Gillund and Shiffrin 1984, Hintzman 1988, Hinton 1989, Clark and Gronlund 1996). Content-addressable retrieval refers to the use of the contents of memory encodings themselves in the access procedure. Inherent features of the desired encoding are used as probes to identify matching encodings. Thus content-addressable access may be contrasted with access to information that proceeds by iteratively inspecting storage locations, i.e. a search. Content-addressability is implementable in a variety of architectures and is characteristic of most contemporary memory models (see Clark and Gronlund 1996 for a review).[7] The key advantage of building content-addressability into a memory system is that when information of a certain type is desired, it is often not necessary for the system to consult or otherwise be influenced by irrelevant information. This contrasts with search procedures which require comparisons of the desired information with each memory record in the search set. Content-addressable retrievals yield retrieval times that are independent of the size of the search set. Searches, in contrast, yield retrieval times that are directly proportional to set size. The second advantage of content-addressability is thus its speed.

2.4 *Question 3: What factors determine the success of retrieval?*

The major determinant of retrieval success is the match between the information used at retrieval – the cues – and the desired encoding. If the combination of cues used at retrieval is sufficiently distinct, then retrieval success will be high. However, if they apply to many different encodings in memory, then the wrong encoding may be retrieved. This phenomenon is referred to as similarity-based interference (Anderson and Neely 1996). This is the trade-off for the fast access times associated with content-addressable memories: irrelevant encodings that are similar to the desired encodings can negatively impact processing. This problem is potentially very acute for linguistic representations. Because linguistic representations are recursive, compositional objects built out of a relatively small repertoire of atomic parts, they contain highly self-similar subparts.

[7] The address labels in a random-access memory (RAM) are themselves arbitrary, but direct, content-addressable access to specific encodings can nonetheless be achieved using hash-coding (Kohonen 1977). The crucial issue is thus not how the memory is physically implemented, but how many memory accesses are necessary to retrieve a desired encoding.

On the one hand, much available evidence suggests that content-addressability nonetheless prevails in memory retrieval for language processing. There is growing evidence of similarity-based interference in certain environments (Gordon *et al.* 2001, Van Dyke and Lewis 2003, Lewis and Vasishth 2005). As an example of both similarity-based interference and its relevance to language processing, consider one of the experimental sentences from Van Dyke and Lewis (2003), in (6). In this sentence, the subject of the embedded clause, *the student*, should be paired with the predicate, *was standing*.

(6) The secretary forgot that . . .
the student who thought that the exam was important was standing in the hallway.

However, the presence of a full lexical subject (*the exam*) in the intervening relative clause can impact the dependency formation process, rendering it slower or less accurate. This finding has been replicated in a number of contexts (Van Dyke 2007, Wagers 2008), with the major determinant of retrieval success appearing to be whether or not a linearly intervening, grammatically inappropriate [Spec,TP] position is occupied. Consistent with this evidence, Lewis and Vasishth (2005) and Lewis *et al.* (2006), on the basis of their ACT-R model of sentence processing, have argued that similarity-based retrieval interference is a significant determinant of comprehension success. In ACT-R, or in any model that incorporates content-addressability, the fact that the encodings of the grammatically appropriate subject in (6) and the grammatically inaccessible subject overlap in some of their features is what renders them liable to retrieval interference. Many instances of grammatically inaccurate performance seem to yield nicely to a retrieval-based account, such as patterns of case and agreement attraction (Wagers *et al.* 2009).

On the other hand, there is a large body of evidence indicating that much of sentence processing is grammatically accurate, and interference-robust (Phillips *et al.* 2011). An important research question is how these two sets of empirical observations might be reconciled. It may be that different linguistic phenomena are processed by distinct memory mechanisms. According to this view, fast, interference-prone memory operations could characterize some kinds of dependencies in language; while slower, search-style operations characterize others. A second (non-exclusive) possibility is that the systems which encode linguistic representations and manage the cues used at retrieval have become effectively adapted to the nature of linguistic representations, such that optimally diagnostic cue sets are used in retrieval. Ericsson and Kintsch (1995), in their theory of long-term working memory, have made essentially this proposal to account for expert performance in well-practiced, narrow domains – for example, chess. A skilled player can accurately recall complex (legal) chess positions from her recent games, but not because she has an exceptional memory. Instead, it is argued she has a well-developed skill at encoding specific episodes of chess in a way that will make them selectively targetable with an effective and frugal retrieval structure, despite the fact that there are many abstract similarities from chess game to chess game. Ericsson and Kintsch write, “the acquired memory skill involves the development of encodings for which the subject can provide controlled access to significant aspects of the encoding context and thus indirectly to the desired information in a manner consistent with the encoding-specificity principle (Tulving, 1983)” (1995: 216). It seems likely the same skills are developed in language comprehension. Though the component memory processes are themselves fast, limited, and error-prone, comprehenders may strategically encode each new dependency to guard against similarity-based interference. In section 4, I will illustrate one possible strategy for doing so with long-distance dependencies.

In this section, I have sketched an account of the memory architecture which forms one component of the language-processing system. To summarize: this architecture allows for minimal concurrent activation of cognitive representations; it is therefore frequently cycling representations between the active and passive states, guided by content-addressable retrieval. By virtue of its content-addressability, cognitive operations are liable to fail when similar representations compete at retrieval. In the next section I shall turn to what the parsing of *wh*-dependencies looks like in such an architecture.

3 The time course of long-distance dependency formation

3.1 The basic generalization: island-sensitive, active dependency formation

The last thirty years of psycholinguistic research have set several important empirical boundaries on any theory of *wh*-processing. The first observation is that *wh*-processing is a highly incremental, predictive process: a property I’ll refer to as *active dependency formation*. An important cue to a long-distance dependency’s interpretation is the absence of a particular constituent, i.e., the gap. The parser must “detect” the gap – that is, identify its subcategorizing syntactic head – in order to correctly thematically integrate the filler. Jackendoff and Culicover (1971), in discussing patterns of acceptability in dative questions, proposed that gaps were only hypothesized as a last resort: only when an obligatory constituent was absent would the gap be “detected.” The Augmented Transition Network model of Wanner and Maratsos (1978) embodied this strategy by only analyzing and attaching displaced NPs if it could not otherwise recognize a legal VP constituent. Fodor (1978), however, argued that the last resort strategy was too strong. For example, it predicted comprehension difficulty that does not exist with extraction from optional transitives. The sentence in (7) illustrates this observation with the verb *read*: there is no intuitive difficulty in recognizing the filler-gap dependency, yet "read again to her son" is a legal VP even without a gap.

(7) Which book did the tired mother have to read again to her son?

Fodor proposed that gaps should be able to be postulated in advance of direct evidence for their location. In a case like (7), Fodor argues, readers preferentially posit a direct object gap because of the likelihood that *read* will take a direct object.

Direct experimental evidence has broadly supported the idea that the parser completes long-distance dependencies without waiting for unambiguous evidence of the gap position: that is, it is an active dependency completion parser. Phillips and Wagers (2007) present a review of the evidence that supports this conclusion, evidence which spans a diverse array of experimental methodologies as well as one of the broadest cross-linguistic samples I know of in psycholinguistics.[8] Here I will mention only a few experiments. One of the earliest demonstrations comes from the *filled-gap effect* (Crain and Fodor 1985, Stowe 1986). Stowe (1986) compared self-paced reading times for sentences containing a displaced *wh*-phrase, like (8a), with matched sentences without a *wh*-dependency, like (8b):

(8) a. My brother wanted to know **who** Ruth will bring *us* home to ____ at Christmas.
b. My brother wanted to know if Ruth will bring *us* home to Mom at Christmas.

The direct object NP, *us*, led to longer-reading times in (8a) compared to the same NP in (8b). Stowe construed this contrast as a surprise effect, which derived from an initial direct object interpretation of the filler. Using the same logic, Lee (2004) has shown that prior to its direct object interpretation, the filler is interpreted as originating in subject position. Electrophysiological studies have provided convergent evidence. For example, Garnsey *et al.* (1989) varied the plausibility of filler–verb combinations, as in (9a–b), and observed immediate detection of the semantic anomaly on the verb which hosts the gap, as indexed by an N400 evoked response.

[8] CROSS-LINGUISTICALLY: Dutch (Frazier 1987a, Frazier and Flores d'Arcais 1989, Kaan 1997), German (Schlesewsky *et al.* 2000), Hungarian (Radó 1999), Italian (de Vincenzi 1991), Japanese (Aoshima *et al.* 2004), Russian (Sekerina 2003).

CROSS-METHODOLOGICALLY: Electrophysiology using EEG (Garnsey *et al.* 1989, Kaan *et al.* 2000, Phillips *et al.* 2005) and MEG (Lau *et al.* 2006), the "stops making sense" task (Tanenhaus *et al.* 1985, Boland *et al.* 1995), eye-tracking (Traxler and Pickering 1996), cross-modal lexical priming (Nicol and Swinney 1989, Nicol *et al.* 1994), anticipatory eye movements (Sussman and Sedivy 2003).

(9) a. The businessman knew which customer the secretary called ___ at home.
b. The businessman knew which article the secretary called ___ at home.

Traxler and Pickering (1996) showed that reading times measured in eye-tracking experiments increase under a similar manipulation.

There are a number of possible (non-exclusive) motivations for active dependency formation. As Fodor (1978) argued, a non-active parser – one which is cued by failure to detect an obligatory constituent – will often unknowingly make mistakes. An active parser might often make mistakes – which is the very premise of the filled-gap effect – but it will nonetheless obtain a clear error signal. And when it does not make mistakes, the active parser will be able to establish some crucial aspects about the sentence's interpretation sooner rather than later. Wagers and Phillips (2009) argued that an important and distinct motivation for active dependency formation derives from a pressure to satisfy open grammatical licensing requirements as soon as possible, a view which can be identified with principle-based parsing (Pritchett 1992, Weinberg 1992). Using a plausibility manipulation, they contrasted the processing of across-the-board raising extractions from conjoined VPs (as in 10a–b) with potential parasitic gap hosts (as in 11a–b). While extraction is essentially obligatory from both conjuncts of conjoined VPs, it is optional in the case of parasitic gaps. Wagers and Phillips found evidence for active dependency completion for the second gap in the obligatory extractions (10), but not for the potential, optional extractions (11).

(10) *Coordinated VP, Plausible*
a. The **wines** which the gourmets were energetically discussing __ or slowly *sipping* __ during the banquet were rare imports from Italy.

Coordinated VP, Implausible
b. The **cheeses** which the gourmets were energetically discussing __ or slowly *sipping* __ during the banquet were rare imports from Italy.

(11) *Potential parasitic gap, Plausible*
a. The **wines** which the gourmets were energetically discussing __ before slowly *sipping* the samples during the banquet were rare imports . . .

Potential parasitic gap, Implausible
b. The **cheeses** which the gourmets were energetically discussing __ before slowly *sipping* some wine during the banquet were rare imports . . .

The parser thus appears to be sensitive to what the grammar requires of well-formed long-distance dependencies in a way that affects whether or not the dependency is completed actively.

This finding leads us to a broader generalization, which is the second – and for this chapter, more interesting – property of long-distance dependencies completion: *island sensitivity*. Though the parser completes dependencies actively, it does not do so at all costs. The across-the-board extraction facts explored by Wagers and Phillips (2009) strengthens a much larger body of observations that island domains are respected in online processing. Across many studies the parser overwhelmingly does not posit gaps inside island domains, evidence that Phillips (2006) reviews. Traxler and Pickering (1996), an eye-tracking study, provides an illustrative example of this fact for relative clause islands (see also Phillips, this volume, chapter 4). Firstly they consider a plausibility contrast in a simple relativization:

(12) We like **the book / the city** that the author *wrote* unceasingly and with great dedication about ___ while waiting for a contract.

In the case of either filler, the ultimately correct analysis is unexceptionable: books and cities are both plausible things to write about. In the initial analysis, however, the active parser interprets the filler as the direct object of *wrote*. While it is plausible to write books, it seems impossible to "write cities." Consistent with similar experiments, like Garnsey *et al.* (1989), Traxler and Pickering observe a slowdown at the verb *wrote* for the implausible filler compared to the plausible one. In a second contrast, they subordinated *wrote* inside a relative clause, as in (13), thus making any potential gap hosted by *wrote* inaccessible to the filler in the matrix clause.

(13) We like **the book / the city** that the author [$_{\text{RC}}$ who *wrote* unceasingly and with great dedication] saw ___ while waiting for a contract.

In this case, no effect of filler plausibility is observed at the lure verb *wrote*. The authors conclude that comprehenders never entertain the island-violating dependency as an analysis.

One straightforward interpretation of the island sensitivity findings is that the grammatical restriction against long-distance dependencies terminating inside certain domains causes the parser to refrain from positing a dependency inside of them. However, a potential worry can be raised about such findings: because island sensitivity is reflected by the absence of any effect (with the exception of Wagers and Phillips 2009), it may be that the parser does not specifically refrain from positing dependencies, but that it is simply unable to do so. In other words, demonstration of island sensitivity in long-distance dependencies completion experiments is also potentially evidence for the proposal that islands are real-time epiphenomena (Deane 1991, Pritchett 1991b, Kluender and Kutas 1993b). Phillips (2006) attempted to respond to this proposal for subject islands (Pritchett 1991, Kluender 2004). Extraction from subjects normally appears to be ill-formed (14).

(14) *What did the attempt to repair ___ ultimately damage the car?

However, as with VP adjuncts, parasitic gaps are legal in certain subjects, particularly those whose head takes an infinitival complement, as in (15).

(15) What did the attempt to repair $\_\_\__{PG}$ ultimately damage ___?

Phillips (2006) shows that comprehenders posit gaps inside just those subject phrases which can support a parasitic gap. Thus, he argues, it cannot be the inability of comprehenders to create a dependency in a subject environment that accounts for the ill-formedness of sentences like (14).[9]

3.2 *Integration with the working memory architecture*

With these two basic facts about long-distance dependency completion in hand – dependency completion is active but it is island-sensitive – we can attempt to integrate the theory of memory explored in section 2 with the time-course facts. There are two essential questions to answer:

(16) *Maintenance*
Is any information actively maintained in memory while a long-distance dependency remains unresolved?

(17) *Retrieval success*
What information is used to cue the retrieval of the filler, once the gap is postulated?

The maintenance question (16) has often suggested itself as a source of difficulty in comprehension (Gibson 1998, 2000; Fiebach *et al*. 2002; Kluender and Kutas 1993b). To use a phrase sometimes encountered in discussions of this phenomenon, the trouble with long-distance dependencies may be that they require a filler to be "carried" forward in time. The retrieval success question (17) has figured somewhat less prominently in discussion of long-distance dependencies comprehension but it is closely related to the view that fillers have to be reactivated. In some recent research, Hofmeister and Sag (2010) have fingered retrieval success as a major determinant of complexity in long-distance dependency formation.

I will first attempt to answer the maintenance question. While the answer here is unfortunately probably the least clear, the data suggest that, overall, only very little is actively maintained of the filler's contents. If this is the correct conclusion, it is convergent with the focus-of-attention limitation on concurrent

[9] Phillips (2006) and Wagers and Phillips (2009) present an interesting contrast: comprehenders are willing to undertake a subject parasitic gap analysis, but not an adjunct parasitic gap analysis. Wagers and Phillips (2009) argue that the relative linear ordering of the parasitic and licensing gap may explain this difference: comprehenders are willing to undertake the parasitic gap analysis while they are still searching for the licensing gap, but they will not actively undertake a parasitic gap analysis if the licensing gap is already resolved. That claim remains to be directly tested.

maintenance introduced in section 2.2. It also suggests that the hypothetical cost of carrying filler information forward in time is itself not a major contributor to complexity to be reckoned with in prospective theories of islandhood.

3.2.1 Maintenance Wanner and Maratsos (1978) was an early proponent of the idea that there is a storage cost for the filler in an incomplete long-distance dependency. In their ATN ("augmented transition networks") model, the cost stemmed from the fact that the filler was kept in a distinguished memory register, called the HOLD cell,[10] until it was analyzed at its gap location. The Active Filler Strategy, developed in various forms in Frazier (1987a), Frazier and Flores d'Arcais (1989), and Frazier and Clifton (1989), can be seen as inspired by the spirit of the HOLD cell hypothesis, though it is not explicitly committed to maintenance. The Active Filler Strategy states that the parser should prefer attaching gaps to full NPs while a long-distance dependency is being resolved. However, there is a more interesting dynamic underpinning to the Active Filler Strategy than merely enforcing a preference: if the filler could somehow effectively outcompete the bottom-up input for attachment at potential argument sites, then the observed "preference" would follow as a consequence of this competition. It would, of course, be necessary to have a mechanism to promote this competition, to covertly enter the filler into the parser's workspace. In Frazier's terms, the filler had to be kept non-"inert."

Consistent with a maintenance hypothesis, long-distance dependencies do robustly exact a cost on linguistic performance, as measured by almost any method (e.g., Wanner and Maratsos 1978, King and Just 1991, Sprouse *et al.* 2012). However, Gibson and Grodner (2005), among others, have tied this cost to the re-integration of the filler at the gap site, and not the openness of the dependency per se. Of course it is possible that the reading time evidence supporting that conclusion may simply not be sensitive to memory load in the right way. Indeed, researchers working with electrophysiological methods have occasionally reached a seemingly opposite conclusion. EEGs obtained while participants read object-extracted filler-gap dependencies or long-distance questions show a sustained anterior negativity (SAN) while the dependency remains open (King and Kutas 1995, Fiebach *et al.* 2002, Phillips *et al.* 2005). The SAN has previously been implicated in explicit memory load tasks (Ruchkin *et al.* 1990) and thus its presence in open filler-gap dependencies has been interpreted as a

[10] More specifically, the proposed cost stemmed from the fact that the filler phrase was not stored as an analyzed constituent, but only as a string-to-be-analyzed. The assumption that the filler was unanalyzed was a consequence, in part, of the architecture of augmented transition networks; but linking the cost to its string encoding seemed to fit nicely with an early finding in psycholinguistics that imposing a syntactic analysis significantly improved the effective span of immediate free recall (see Fodor *et al.* 1974).

rather direct reflection of the memory load consumed by actively maintaining the filler.

This interpretation is qualified by two points. Firstly, the SAN does not reflect a cumulative effect that accrues or is renewed at each word, but instead it derives mainly from the first words of the dependency (King and Kutas 1995, Phillips *et al.* 2005). More importantly, the SAN studies do not speak to what the actual contents of the filler are when the dependency is unsatisfied. In discussing the SAN, Fiebach *et al.* (2002) are careful to point out that the electrophysiological effect itself does not discriminate between alternative accounts of what is being maintained. It could be a full semantic or syntactic representation of the filler, or perhaps a subset of the features of those representations. Alternatively, it may not contain the filler's content at all but instead register the existence of an (unsatisfied) prediction for a syntactic environment that allows completion of the dependency, as in Dependency Locality Theory (Gibson 2000).

More direct evidence about the maintained contents of the fillers come from cross-modal lexical activation studies and probe recognition tasks (Bever and McElree 1988, Nicol and Swinney 1989, Nicol *et al.* 1994). In cross-modal lexical decision, auditory word-by-word sentence comprehension is interrupted unpredictably by an on-screen lexical decision task. For example, a test sentence containing a long-distance dependency might be given as follows:

(18) The doctor that the visitor from the prestigious University met in the lobby seemed to be in a hurry.

The lexical decision task would then probe with either a semantic associate of the filler *doctor* – for example, "nurse" – or a suitable control word – for example, "mechanic." By arranging to interrupt sentence comprehension at enough points between the filler and the gap (over the course of the experiment), it is possible to derive a temporal profile of the filler's activation in memory. In all such studies, there is evidence that semantic features of the filler are not maintained and must be reactivated later in a state suitable for integration. In particular the filler's semantic associates are primed immediately following the introduction of the filler in the sentence (Nicol and Swinney 1989, Nicol *et al.* 1994). However, the priming effect declines sharply during the rest of the sentence and only rises again at the gap site. McElree (2001) documented a similar pattern using a probe recognition task.[11] Thus the reactivation studies

[11] In this task, participants had to report whether the interrupting word was a synonym of an adjective contained in the filler phrase. For example, if the filler were "the brave puppy," a related probe word might be "courageous." This design feature avoids a contention raised by McKoon and Ratcliff (1994) about earlier studies. There they argued that the RT facilitation taken to be evidence for reactivation could reflect the fact that interruption points near the gap were also good candidate integration sites for the lexical decision target. Targets happened to be nouns and the intervening interruption points were less good fits for introducing a noun than

suggest that the filler is not concurrently maintained in a rich enough form to prime its lexical associates or otherwise facilitate processing of related words.

Finally, Wagers and Phillips (2009) reported that the slowdown effect which is normally robustly observed at the verb when the filler is implausible can be seriously attenuated if the distance between filler and verb is increased only slightly. When a relative-clause long-distance dependency was serially lengthened by attaching a five-word prepositional phrase to the intervening subject, the slowdown effect no longer surfaced on the verb and instead moved to a post-gap position. Wagers and Phillips (2009) proposed that, for longer long-distance dependencies, the anomaly detection effect could be greatly delayed because semantic integration might no longer take place immediately at the verb, but only after the semantic details of the filler could be adequately retrieved.[12] Wagers and Phillips (2012) replicated their original finding for the plausibility effect but reported that at dependency lengths comparable or longer, the filled-gap effect was preserved. The modified filled-gap paradigm they employed contrasted DP and PP extraction (as in Lee 2004). The fact that this index of active dependency completion survived multiple dependency lengths suggests that, unlike fine semantic details, at least coarse-grained syntactic category information about the filler may be maintained (consistent also with Gibson 2000).

3.2.2 Retrieval If the filler's contents are not maintained in a privileged state, they must be retrieved once a suitable gap host is identified in the input or constructed based on expectations of upcoming input. There are several possibilities for how the correct filler constituent is then retrieved. Lexical information, such as the syntactic or semantic restrictions the verb places on its arguments, may be one source of retrieval cues for the filler. Van Dyke and McElree (2006) provided some direct evidence that the filler is retrieved at the site of integration based on verb properties, and that this retrieval is subject

the direct object position. According to this interpretation of the data, the cross-modal lexical decision task does not reflect activation state. The McElree (2001) study partially blunts this criticism by probing with a word which could not syntactically or semantically be integrated at the point of interruption.

12 In light of the generalizations developed in section 3.1, this finding might seem at first surprising. It is useful to note, however, that past studies of filler-gap dependency construction have been heavily skewed to very short distances. In a sample of twenty-one influential experiments performed between 1986 and 2004, Wagers (2008) found there were on average only 2.9 linear interveners, or 1.8 constituent interveners, between filler and gap. These experiments spanned different paradigms: 10 self-paced reading studies, 5 eye-tracking studies, 3 sensicality monitoring studies, 3 cross-modal priming studies: Tanenhaus *et al.* (1985); Stowe (1986); Swinney *et al*. (1988); Frazier and Clifton (1989); Pickering *et al.* (1994); Boland *et al*. (1995); Traxler and Pickering (1996); Clahsen and Featherston (1999); McElree (2000); Sussmann and Sedivy (2003); Aoshima *et al.* (2004); Conklin *et al.* (2004); Lee (2004).

to interference. They examined the processing of clefts under a memory load manipulation. In half the experimental conditions, participants were presented with a list of three nouns at the start of the trial, which would have to be recalled after the sentence comprehension task (Load conditions). For example:

(19) table – sink – truck

Participants then read sentences like the following:

(20) It was **the boat** that the guy who lived by the sea *sailed / fixed* ___ in two sunny days.

Two possible critical verb types could occur in the sentence. In half the conditions, exemplified by "sailed," the critical verb was not a good fit for the memory load nouns and thus generated low interference: it is not plausible to sail a table, sink, or truck. In the other half of the conditions, the critical verb was a good fit and thus generated high interference: tables, sinks, and trucks are, like boats, fixable things. In the critical verb region they reported an interaction of interference and load conditions. Reading times between the two verb types were identical when there was no memory list; however, under Load conditions, high-interference verbs were read more slowly. Thus the goodness of the match between the verb and the filler as its object determined the ease of comprehension in this experiment.

The plausibility effect studies, like Garnsey *et al*. (1989) and Traxler and Pickering (1996), provide an interesting constraint on the retrieval process. Recall that in those studies, an anomalous verb–filler combination was detected very early at the verb. This suggests that even if a filler is not a good object for the verb, it may nonetheless be retrieved. However, previous studies have not always carefully distinguished between purely selectional restrictions and violations of real-world expectations, so some caution is required. For example, in Traxler and Pickering (1996), the verb–filler combination "shoot the garage" is certainly unusual, but it does not obviously violate a selectional restriction. Boland *et al*. (1995) found that active dependency formation was not pursued for anomalous verb–object combinations if the verb was a verb like "remind" – which combines with an animate object and a clause controlled by that object – and the filler was an inanimate noun like "movie." Pickering and Traxler (2001) obtained convergent evidence for this finding and argued that certain simple selectional features, particularly animacy, could be used to filter dependency completion. One mechanism for achieving this filtering is to block retrieval of the filler if it is known not to match the gap host's selectional restrictions or if an initial retrieval based on a positive selectional requirement fails.

Subcategorization information may also be well suited as a cue. For example, intransitive verbs like "talk" or "arrive" do not combine with a direct object, so one can ask what kind of gap positions are actively projected in VPs headed by intransitive verbs.[13] The subcategorization frame for such verbs, containing no second argument, may prevent – or perhaps only make less effective – the retrieval of the filler. The evidence is currently somewhat mixed. If we pay attention to two quite recent studies, we find Staub (2007a) on the one hand, finding no evidence for active dependency formation at intransitive verbs, and Omaki *et al.* (2011) on the other, finding positive evidence for active dependency formation. The empirical landscape is somewhat clearer with regards to relative subcategorization frequency. Both Pickering and Traxler (2003) and Frazier and Clifton (1989) find that, when DP and PP subcategorization frames compete, the filler is always initially analyzed as being extracted from DP position, regardless of preference (but cf. Stowe *et al*. 1991, and discussion in Fodor 1978).

Finally, although the gap host can provide a rich set of cues with which to retrieve the filler, we also know that it cannot be the only source of information. Verb-final languages like Japanese show evidence of active long-distance dependency formation well in advance of the verb (Aoshima *et al*. 2004, Nakano *et al*. 2002). If Omaki *et al.* (2011) are correct that the parser sometimes constructs gaps hosted by intransitive verbs, then even English may be a language in which long-distance dependency formation occurs without verb information. Nothing should seem outlandish in that proposal, as long as it is allowed that the construction of a particular syntactic phrase is not strictly dependent on its head having been pronounced. Interestingly Nakano *et al*. (2002) provide evidence from cross-modal lexical priming for pre-verbal activation in the argument field of a gapped Japanese VP. Moreover, they find evidence that reactivation is strongest among individuals who have high WM scores. These two pieces of evidence suggest that pre-verbal dependency formation may actually involve retrieval of the displaced argument, and not merely formation of a content-free syntactic representation. If there is retrieval from outside of the focus of attention, then the cues must derive from parsing rules and expectations and not from information contained in a specific lexeme.

[13] The necessary distinction may need to be finer than simply whether the verb takes a second argument. Unergative verbs, like run, which only take an external argument, may differ from unaccusative verbs, like arrive, whose single argument may originate within the VP. Likewise, the possibility of diathesis phenomena such as causative/inchoative and unergative/transitive alternations could affect the retrieval; in the psycholinguistic literature these alternations have usually been addressed only indirectly, under the cover term of a particular verb's "optional intransitivity."

4 Implications for island constraints

4.1 Recap

If one attempts to put together the theoretical and empirical pieces from sections 2 and 3, the following sort of picture emerges. When a displaced phrase is encountered during language comprehension, the parser actively seeks to integrate it with its gap host. However, other kinds of processing events must occur in the meanwhile. Because of the strong narrow-capacity limitation on working memory, all or almost all of the displaced phrase's encoding must be displaced from the focus of attention. There is thus relatively little maintenance possible during the processing of a filler-gap dependency and the evidence suggests that there is very little maintenance required. This does not mean that no filler information is preserved in advance of retrieval. The proposal, by Pickering and Traxler (2001), that some kinds of selectional requirements set early filters on dependency formation suggests that it could be useful to preserve a bare-bones feature set for the filler. Furthermore, the evidence from Wagers and Phillips (2012) that the filled-gap effect survives long dependency lengths suggests at least coarse syntactic category information may be preserved. Stronger evidence is needed to substantiate these claims. The kind of focus-of-attention studies described in section 2 would provide the most convincing evidence.

The parser cues retrieval at some point in advance of the gap site. Although it is typically taken to be the verb that cues retrieval in a verb-medial language like English, the evidence is scant that no dependency formation occurs in advance of the verb. The cue set or retrieval structure used during this process can come not only from (1) a syntactic head, like the verb, but also (2) the internal rules of the parser, (3) any information maintained in the focus of attention, and potentially (4) non-linguistic information from the general context. The best case scenario is one in which the cue set compiled is sufficiently specific to activate one encoding. In general the parser can improve the likelihood of successful retrieval by ensuring that the information from multiple (uncorrelated) sources contributes to the retrieval set. As mentioned in section 2, experts in particular domains, like chess or music, develop large and accurate memories in their domain of expertise by learning how to adaptively assemble retrieval structures. It seems reasonable, though as yet undemonstrated, to expect that language comprehenders do the same.

In Figure 7.1, I lay out one strategy comprehenders could use to assemble their retrieval structures in a way that would strengthen the likelihood of a successful grammatical outcome. The basic idea is that comprehenders should maintain a very sparse feature set from the original filler: indeed, it

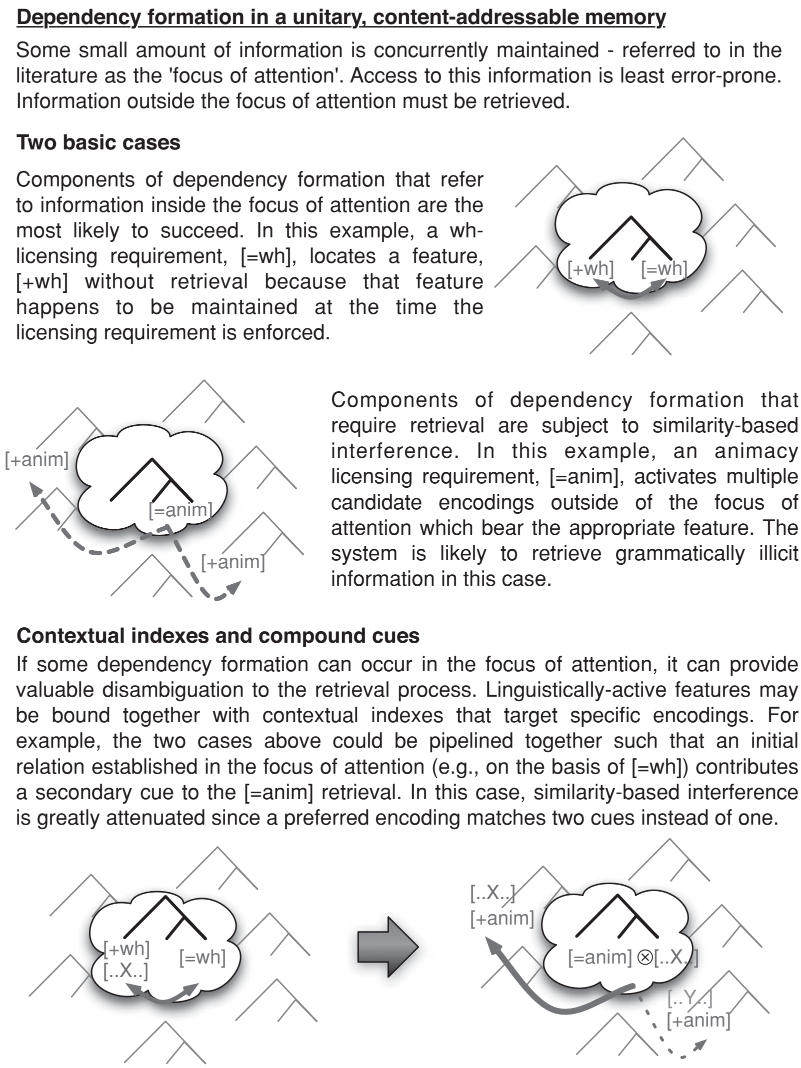

Figure 7.1 Dependency formation in a unitary, content-addressable memory

is less important to retain grammatical features, and more important to retain features that are specific to the event of encoding the filler, something like a randomly generated key or a hash (see Polyn *et al.* 2009 for a formalization of this idea, and Tulving and Thomson 1973 for theoretical grounding).

If comprehenders combine the maintained features with the retrieval cues generated at the retrieval site, then retrieval interference would be significantly dampened.[14]

4.2 *Implications for islandhood*

Thus far I have been mostly concerned with clarifying the memory processes involved in constructing a filler-gap dependency. Now it is time to return to the original question: can any properties of the memory system contribute to an explanation of island effects? In other words, is the inability of filler-gap dependencies to cross certain boundaries attributable to an inability to construct the dependency in real time because of memory failures? I emphasize that I am restricting the question of interest to the contribution of the working memory system to parsing failure. Under a multi-componential account, like Kluender and Kutas (1993b) or Kluender (2004), many difficulties might combine together and those difficulties may relate to other properties of the system.

I think that the totality of the evidence suggests that any failures must lie in the retrieval process – that is, any failures must stem from the parser's inability to *retrieve* filler information at some point during dependency completion. At present the evidence for maintenance in normal, uneventful long-distance dependency completion is too weak to attribute variation in acceptability among structures to that cause. Pernicious effects of retrieval interference, on the other hand, and the importance of well-chosen retrieval structures, are abundantly clear throughout both the memory literature and the psycholinguistics literature.

Retrieval interference is thus the strongest candidate to underlie any hypothetical memory failures in long-distance dependency completion. However, at the same time, it seems highly unlikely that retrieval interference is ever severe enough in sentence comprehension to guarantee actual retrieval failures. There are at least four reasons to believe this is the case. The first is simply a theoretical observation, and the next three are empirical. The theoretical

[14] As a reviewer points out, it would be important to establish whether sparsification of the original filler representation would solve the problem set by the limited maintenance capacity of the system. This is clearly an empirical question. One can imagine, following Kluender and Kutas (1993b), that encountering another A-bar operator, as in *wh*-islands or relative clause islands, leads to displacement of even the sparse representation of the filler. If that were true, then it would be important that retrieval still be able to occur when this information is unavailable, and that it can generally succeed. As I discuss below, retrieval interference effects themselves vary in severity, so it is plausible that retrieval outcomes would be successful without the "robustness" provided by carrying forward some episodic information. Note that this is likely true since the mere existence of an island domain in the midst of an unbounded dependency that does not terminate in that domain does not seem to affect the resolution of that dependency. See Wagers (2008), Experiments 9–10.

argument concerns expert performance and practiced skills, the cases Ericsson and Kintsch (1995) were worried about, and the cases for which robustness to retrieval interference is well documented. It would seem odd that sentence comprehension is not well practiced enough to allow comprehenders to develop retrieval structures to overcome any interference which island configurations might generate (for example, the strategy I outline in Figure 7.1). To see that it is likely not the case that comprehenders cannot in principle wield such retrieval structures, we consider the phenomenon of resumption.

Many languages allow resumptive pronouns in the place of gaps. Interestingly, the use of resumptive pronouns often allows the grammar to form an unbounded long-distance dependency that would otherwise violate an island constraint (McCloskey 2002). It has sometimes been speculated that resumptive pronouns might aid in the retrieval of filler material. However, recent evidence indicates that resumption does not obviously improve comprehension performance on (non-island) long-distance dependences (Alexopoulou and Keller 2007). But quite apart from the (scant) empirical work on resumptives in real-time processing, it is actually unclear why resumptives, compared to gaps, should facilitate dependency formation. Across languages it has been observed that resumptive pronouns come from the same pronoun series as anaphoric pronouns (McCloskey 2002). This poses an ambiguity resolution problem: the comprehender must decide whether a pronoun within the scope of an A-bar operator should function as the tail of the A-bar dependency, or whether it should be related to another syntactic or discourse element. Thus resumptives require an ambiguity resolution process just as in gap finding, during which the comprehender must decide whether the absence of a constituent is due to displacement or to argument optionality/lexical ambiguity. Resumptive pronouns do bear *phi*-features, which could conceivably aid retrieval. But using *phi*-features in retrieval is helpful to the extent they are distinctive. If they were not distinctive, and multiple constituents in the sentence bore them, then including the resumptive's features as a retrieval cue could harm comprehension. But I emphasize it remains an open empirical question how resumptives participate in dependency formation. This is particularly true since almost all the research to date on resumption has been in English, for which resumptives clearly have a marginal grammatical status (McCloskey 2006, Heestand *et al.* 2011). However, if it does turn out to be the case that resumption does not enhance the retrievability of the filler, then the existence of island-crossing dependencies with resumptive pronouns strongly suggests that it is possible to retrieve filler material inside an island when it is a grammatical option.[15]

[15] Some very recent evidence by Yoshida *et al.* (in press b) has shown that other unbounded dependencies are readily constructed inside islands, as long as they are not A-bar dependencies. In particular, they show that the resolution of cataphoric dependencies (in which a pronoun

The second reason to doubt that retrieval interference leads to significant retrieval failures comes from the studies that directly document its existence. Consider, in particular, Van Dyke and McElree (2006), an experiment which explicitly introduces memory load items. In that experiment, comprehension accuracy declined only 4 percent between interfering and non-interfering conditions. Recall accuracy – how well items on the load list were recalled – only declined 2 percent, a non-significant effect. Interestingly, in experiments documenting interference effects for subject–verb attachment, comprehension decrements due to interference were comparable in magnitude (Van Dyke and Lewis 2003, Van Dyke 2007, Wagers 2008). In those experiments, the interference manipulation was to vary the number of subject phrases embedded inside the matrix subject and the measure was performance on the matrix verb. Yet despite the existence of comparable interference effects on subject–verb attachment, I am not aware of any proposed grammatical constraints on the relative complexity of pre-verbal subjects in English that are comparable in their severity to island constraints. It is true that other kinds of dependency formation show much larger interference effects than the studies above: in particular, higher rates of parsing or interpretation error can be found in establishing subject–verb agreement (Wagers *et al.* 2009) or resolving reflexive anaphora (eventually) (Sturt 2003). However, I take the Van Dyke and McElree (2006) study, with its relatively low rates of interference, as the most relevant to the question at hand since it deals specifically with a *wh*-dependency.

The third reason to doubt that retrieval interference underlies islandhood comes from recent studies seeking to correlate measures of WM efficiency/capacity with acceptability ratings (Tokimoto 2009, Sprouse *et al.* 2012). Sprouse *et al.* (2012) attempted to correlate scores from two kinds of WM-sensitive test with the acceptability decrements island-violating sentences earned in ratings tasks: performance on an *n*-back task, as well as performance on a word-span task (immediate free recall). The results were unequivocal: there was no useful covariation between memory scores and the severity of island violations. Of course it is always possible that the wrong index of WM was selected – as a host are available – and other candidate WM indices should be tested. Yet it is important to note that the simple recall task Sprouse and colleagues employed shares considerable variance (40%) with more complex span tasks, like operation span (Kane *et al*. 2004). Therefore it seems less likely that merely choosing another WM index will uncover significant covariation. Pursuing more direct measures of WM operation (e.g., the techniques in section 2) is preferable in any case.

is introduced prior to any accessible discourse referent) proceeds actively into relative clause islands. There are clearly differences in the representation of these two dependency types, but the functional demands on the comprehender are similar.

4.3 Closing

In this chapter I have attempted to interpret existing psycholinguistic theory and data in a way that clarifies the time course of long-distance dependency formation and aligns it with the component processes of the working memory system. I hasten to add that our empirical database is rich in many aspects, but in others it still requires us to make a few educated guesses. Where we are data-rich concerns word-by-word time-course questions: what kinds of dependencies are formed, and at what landmark in the sentence. Such information has been derived primarily from carefully designed reading studies (with both behavioral and electrophysiological measures). Where we remain theory- and data-poor is in charting the contents of information maintained not only from past events but also information about expected future events. Such information requires relatively resource-intensive probe recognition experiments, implemented as cross-modal lexical priming with multiple interruption or in speed–accuracy trade-off studies. However, I suspect that coming to grips with the contents of the relevant internal representations will require not only that we do more of the relevant kinds of experiments; we must also push more strongly on the important distinction between the underlying events of comprehension, i.e., the instantiation and manipulation of internal linguistic representations, and the more indirect events, i.e., the incoming words, which supply the system with evidence for those representations. The active formation of long-distance dependencies reminds us that the two need not align in time.

References

Abrusán, M. (2008a). A semantic analysis of negative islands with manner questions. In A. Grønn (ed.), *Proceedings of Sinn und Bedeutung* 12: 1–16.

Abrusán, M. (2008b). Islands of contradiction: Presuppositional and negative islands. In *Proceedings of the 18th Semantics and Linguistics Theory Conference* (SALT 18), pp. 1–18.

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

Abrusán, M. (2011b). *Wh*-islands in degree questions: A semantic approach. *Semantics and Pragmatics* 4: 1–44.

Adams, M. (1985). Government of empty subjects in factive clausal complements. *Linguistic Inquiry* 16: 305–313.

Aldwayan, S., Fiorentino, R., and Gabriele, A. (2010). Evidence of syntactic constraints in the processing of *wh*-movement: A study of Najdi Arabic learners of English. In B. Van Patten and J. Jegerski (eds.), *Research in Second Language Processing and Parsing*, pp. 65–86. Amsterdam: John Benjamins.

Alexopoulou, T., and Keller, F. (2007). Locality, cyclicity and resumption: At the interface between the grammar and the human sentence processor. *Language* 83: 110–160.

Allen, S. E. M. (1999). Learning about Argument Realization in Inuktitut and English: Graduate development in the use of non-ellipsed forms. Unpublished MS, Max Planck Institute for Psycholinguistics.

Allwood, J. (1982). The complex NP constraint in Swedish. In E. Engdahl and E. Ejerhed (eds.), *Readings on Unbounded Dependencies in Scandinavian Languages*, pp. 15–32. Stockholm: Almqvist & Wiksell.

Ambridge, B., and Goldberg, A. (2008). The island status of clausal complements: Evidence in favor of an information structure explanation. *Cognitive Linguistics* 19: 357–389.

Ambridge, B., Pine, J. M., Rowland, C. F., and Young, C. R. (2008). The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgements of argument-structure overgeneralization errors. *Cognition* 106: 87–129.

Anagnostopoulou, E. (1994). *Clitic Dependencies in Modern Greek*. Ph.D. thesis, University of Salzburg.

Anderson, M. C., and Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork and R. A. Bjork (eds.), *Handbook of Perception and Memory*. Vol. 10: *Memory*, pp. 237–313. San Diego, CA: Academic Press.

Andersson, L-G. (1982). What is Swedish an exception to? In E. Engdahl and E. Ejerhed (eds.), *Readings on Unbounded Dependencies in Scandinavian Languages*, pp. 33–46. Stockholm: Almqvist & Wiksell.

Aoun, J., and Choueiri, L. (1996). Resumption and last resort. Unpublished MS, University of Southern California.

Aoun, J., Choueri, L., and Hornstein, N. (2001). Resumption, movement, and derivational Economy. *Linguistic Inquiry* 32: 371–403.

Aoshima, S., Phillips, C., and Weinberg, A. S. (2004). Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language* 51: 23–54.

Aoshima, S., Yoshida, M., and Phillips, C. (2009). Incremental processing of coreference and binding in Japanese. *Syntax* 12: 93–134.

Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics* 24: 65–87.

Ariel, M. (1990). *Accessing Noun Phrase Antecedents*. London: Routledge.

Ariel, M. (1999). Cognitive universals and linguistic conventions: the case of resumptive pronouns. *Studies in Language* 23 (2): 217–269.

Arnold, J. (1998). *Marking Salience: The Similarity of Topic and Focus*. University of North Carolina.

Arnold, J., Wasow, T., Asudeh, A., and Alrenga, P. (2004). Avoiding attachment ambiguities: The role of constituent ordering. *Journal of Memory and Language* 51: 55–70.

Arnold, J., Wasow, T., Losongco, T., and Ginstrom, R. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language* 76: 28–55.

Arnon, I., Estigarribia, B., Hofmeister, P., Jaeger T. F., Pettibone, J., Sag, I. A., and Snider, N. (2005). Long-distance dependencies without island constraints. Paper presented to the HOWL 3: Hopkins Workshop on Language, Johns Hopkins University.

Arnon, I., Snider, N., Hofmeister, P., Jaeger, T. F., and Sag, I. A. (2006). Crosslinguistic variation in a processing account: The case of multiple wh-questions. *Proceedings of 32nd Annual Meeting of the Berkeley Linguistics Society*.

Asher, N., and Wada, H. (1988). A computational account of syntactic, semantic, and discourse principles of anaphora resolution. *Journal of Semantics* 6: 309–344.

Aslin, R., Saffran, J., and Newport, E. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science* 9(4): 321–324.

Asudeh, A. (2012). *The Logic of Pronominal Resumption*. Oxford University Press.

Aylett, M., and Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence and duration in spontaneous speech. *Language and Speech* 47: 31–56.

Aylett, M., and Turk, A. (2006). Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. *Journal of the American Statistical Association* 119: 3048–3058.

Baayen, R. H. (2007). *Analyzing Linguistic Data: A practical introduction to statistics*. Cambridge University Press.

Babby, L. (1980). *Existential Sentences and Negation in Russian*. Ann Arbor, MI: Karoma Publishers.

Babko-Malaya, O. (2003). Perfectivity and prefixation in Russian. *Journal of Slavic Linguistics* 11: 5–36.

Babyonyshev, M., Ganger, J., Pesetsky, D., and Wexler, K. (2001). The maturation of grammatical principles: Evidence from Russian unaccusatives. *Linguistic Inquiry* 32: 1–43.

Baddeley, A. D. (1986). *Working Memory*. Oxford: Clarendon Press.

Badecker, W., and Straub, K. (2002). The processing role of structural constraints on the interpretation of pronouns and anaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28: 748–769.

Bader, M., and Häussler, J. (2010). Toward a model of grammaticality judgments. *Journal of Linguistics* 46: 273–330.

Bailyn, J. (1995). Underlying phrase structure and short verb movement in Russian. *Journal of Slavic Linguistics* 3: 13–58.

Bailyn, J. (2001). On scrambling: A reply to Bošković and Takahashi. *Linguistic Inquiry* 32: 635–658.

Bailyn, J. (2003). Does Russian scrambling exist? In S. Karimi (ed.), *Word Order and Scrambling*, pp. 156–176. Oxford: Blackwell.

Bailyn, J. (2004). Generalized Inversion. *Natural Language and Linguistic Theory* 22: 1–49.

Baker, M. C. (1988). *Incorporation: A theory of grammatical function changing*. The University of Chicago Press.

Baltin, M. (1982). A landing site theory of movement rules. *Linguistic Inquiry* 13: 1–38.

Bard, E. G., Robertson, D., and Sorace, A. (1996). Magnitude estimation of linguistic acceptability. *Language* 72: 32–68.

Basilico, D. (2003). The topic of small clauses. *Linguistic Inquiry* 34: 1–35.

Beck, S. 2006. Intervention effects follow from focus interpretation. *Natural Language Semantics* 14: 1–56.

Bennett, R. (2008). English resumptive pronouns and the highest-subject restriction: A corpus study. Paper presented at the Trilateral (TREND) Linguistics Weekend, University of California Santa Cruz.

Berwick, R. C., and Weinberg, A. S. (1984). *The Grammatical Basis of Linguistic Performance*. Cambridge, MA: The MIT Press.

Bever, T. G. and McElree, B. (1988). Empty categories access their antecedent during comprehension. *Linguistic Inquiry* 19: 35–43.

Bever, T., and Sanz, M. (1997). Empty categories access their antecedents during comprehension: Unaccusatives in Spanish. *Linguistic Inquiry* 28: 69–91.

Bhatt, R. (1999). *Covert Modality in Non-finite Contexts.* PhD thesis, University of Pennsylvania.

Bianchi, V. (2000). The raising analysis of relative clauses: A reply to Borsley. *Linguistic Inquiry* 31: 123–140.

Biber, D., Johansson, S., Leech, G., Conrad, S., and Finegan, E. (1999). *The Longman Grammar of Spoken and Written English*.

Bley-Vroman, R. W., Felix, S. W., and Ioup, G. L. (1988). The accessibility of universal grammar in adult language learning. *Second Language Research* 4(1): 1–32.

Bock, K. and Miller, C. A. (1991). Broken agreement. *Cognitive Psychology* 23: 45–93.

Boeckx, C. (2003). *Islands and Chains: Resumption and stranding*. Amsterdam: John Benjamins.

Boeckx, C. (2008a). Islands. *Language and Linguistics Compass* 2: 151–167.

Boeckx, C. (2008b). *Bare syntax*. Oxford University Press.

Boland, J., Tanenhaus, M., Garnsey, S., and Carlson, G. N. (1995). Verb argument structure in parsing and interpretation, Evidence from wh-questions. *Journal of Memory and Language* 34: 774–806.

Bolinger, D. (1963). Length, vowel, juncture. *Linguistics* 1: 5–29.

Borovikoff, N. (2001). *Unaccusativity and Movement in Russian: Integrating formal syntax and discourse functions*. PhD Dissertation, Indiana University.

Bourdages, J. (1992). Parsing complex NPs in French. In H. Goodluck and M. Rochemont (eds.), *Island Constraints: Theory, acquisition, and processing*, pp. 61–87. Dordrecht: Kluwer.

Brandi, L., and Cordin, P. (1989). Two Italian dialects and the null subject parameter. In O. Jaeggli and K. Safir (eds.), *The Null Subject Parameter*, pp. 111–142. Dordrecht: Kluwer.

Braze, F. (2002). *Grammaticality, Acceptability, and Sentence Processing: A psycholinguistic study*. PhD Dissertation, University of Connecticut.

Brennan, S., and Clark, H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory and Cognition* 22: 1482–1493.

Bresnan, J. (2001). *Lexical-Functional Syntax*. Oxford: Blackwell.

Bresnan, J., and Moshi, L. (1990). Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21: 147–185.

Broadbent, D. E. (1958). *Perception and communication*. New York: Oxford University Press.

Brody, M. (1995). *Lexico-Logical Form*: A radically minimalist theory. Cambridge, MA: The MIT Press.

Brown, J. D. (2008). Effect size and eta squared, *JALT Testing and Evaluation Newsletter* 12: 38–43.

Brown, R. (1973). *A First Language: The early stages*. Cambridge, MA: Harvard University Press.

Bybee, J. (2007). Review of constructions at work: The nature of generalizations in language. *Journal of Child Language* 34: 692–697.

Cabeza, R., Dolcos, R., Prince, S. E., Rice, H. J., and Nyberg, L. (2003). Attention-Related Activity During Episodic Memory Retrieval: A Cross-Function fMRI Study. *Neuropsychologia* 41: 390–399.

Cann, R., Kaplan, T., and Kempson, R. (2004). Data at the grammar–pragmatics interface: The case of resumptive pronouns. *Lingua* 115: 1551–1577.

Caplan, D., and Waters, G. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences* 22: 77–94

Cardinaletti, A, and Guasti, M. T., eds. (1995). *Syntax and Semantics 28: Small clauses*. San Diego, CA: Academic Press.

Carlson, K. (2001). The effects of parallelism and prosody in the processing of gapping structures. *Language and Speech* 44: 1–26.

Carlson, K. (2002). Parallelism and prosody in the processing of ellipsis sentences [Mar]. *Dissertation Abstracts International, A: The Humanities and Social Sciences* 62: 3026-A.

Carpenter, P. A., and Just, M. A. (1975). Sentence comprehension: A psycholinguistic model of verification. *Psychological Review* 82: 45–76.

Carpenter, P. A., Just, M. A., Keller, A. K., Eddy, W. F., and Thulborn, K. R. (1999). Time course of fMRI activation in language and spatial networks during sentence comprehension. *NeuroImage* 10: 216–224.

Cattell, R. (1976). Constraints on Movement Rules. *Language* 52: 18–50.

Cattell, R. (1978). On the source of interrogative adverbs. *Language* 54: 61–77.

Chafe, W. L. (1976). Givenness, contrastiveness, subject, topic, and point of view, in C. Li (ed.), *Subject and Topic*, pp. 27–55. New York: Academic Press.

Chafe, W. L. (1987). Cognitive constraints on infomation flow. In R. Tomlin (ed.), *Coherence and Grounding in Discourse*, pp. 21–51. Amsterdam: Benjamins.

Chafe, W. (1994). *Discourse, Consciousness and Time: The flow and displacement of conscious experience in speaking and writing*. University of Chicago Press.

Chambers, S. M., and Forster, K. I. (1975). Evidence for lexical access in a simultaneous matching task. *Memory and Cognition* 3: 549–559.

Chao, W., and Sells, P. (1983). On the interpretation of resumptive pronouns. In *Proceedings of the 13th Annual Meeting of the North East Linguistic Society*, pp. 47–61.

Chen, E., Gibson, E., and Wolf, F. (2005). Online syntactic storage costs in sentence comprehension. *Journal of Memory and Language* 52: 144–169.

Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics* 6: 339–405.

Chomsky, N. (1962). The logical basis of linguistic theory. In M. Halle (ed.), *Preprints of the Papers for the Ninth International Congress of Linguists, August 27–31, 1962*, pp. 509–574.

Chomsky, N. (1964). Current issues in linguistic theory. In J. A. Fodor and J. J. Katz (eds.), *The Structure of Language: Readings in the Philosophy of Language*, pp. 50–118. Englewood Cliffs, NJ: Prentice Hall.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press.

Chomsky, N. (1972). Some empirical issues in the theory of Transformational Grammar. In Paul Stanley Peters (ed.), *Goals of Linguistic Theory*, pp. 63–130. Englewood Cliffs, NJ: Prentice Hall.

Chomsky, N. (1973). Conditions on transformations. In S. Anderson and P. Kiparsky (eds.), *A Festschrift for Morris Halle*, pp. 232–286. New York: Holt, Rinehart, & Winston.

Chomsky, N. (1977). On Wh-Movement. In P. Culicover, T. Wasow, and A. Akmajian (eds.), *Formal Syntax*, pp. 71–132. New York: Academic Press.

Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. (1986). *Barriers*. Cambridge, MA: The MIT Press.

Chomsky, N. (1993). A Minimalist Program for linguistic theory. In Kenneth Hale and Samuel Jay Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, pp. 1–52. Cambridge, MA: MIT Press. (Also appears as chapter 3 of Chomsky 1995.)

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. (2000). Minimalist inquiries: The framework. In Roger Martin, David Michaels, and Juan Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, pp. 89–157. Cambridge, MA: MIT Press.

Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A life in language*, pp. 1–52. Cambridge, MA: MIT Press.

Chomsky, N. (2008). On phases. In R. Freidin, C. P. Otero, and M.-L. Zubizarreta (eds.), *Foundational Issues in Linguistics*, pp. 133–166. Cambridge, MA: MIT Press.

Chomsky, N., and Lasnik, H. (1977). Filters and control. *Linguistic Inquiry* 8(3): 425–504.

Chomsky, N., and Miller, G. (1963). Introduction to the formal analysis of natural languages. In R. Luce, R. Bush, and E. Galanter (eds.), *Handbook of Mathematical Psychology*, pp. 269–321. New York: Wiley.

Chung, S. (1982). Unbounded dependencies and Chamorro grammar. *Linguistic Inquiry* 13: 39–77.

Chung, S. (1994). Wh-agreement and referentiality. *Linguistic Inquiry* 25: 1–44.

Chung, S. (2005). Sluicing and the lexicon: The point of no return. In Rebecca T. Cover and Yuni Kim (eds.), *Proceedings of the 31st Annual Meeting of the Berkeley Linguistics Society*, pp. 73–91. Berkeley, CA: Berkeley Linguistics Society.

Chung, S., Ladusaw, W. A., and McCloskey, J. (1995). Sluicing and Logical Form. *Natural Language Semantics* 3: 239–282.

Chung, S., Ladusaw, W. A., and McCloskey, J. (2010). Sluicing: Between structure and inference. Unpublished MS, University of California Santa Cruz.

Chung, S. and McCloskey, J. (1983). On the interpretation of certain island facts in GPSG. *Linguistic Inquiry* 14: 704–713.

Cinque, G. (1990). *Types of A'-dependencies*. Cambridge, MA: MIT Press.

Cinque, G. (1995). The pseudo-relative and ACC-ing constructions after verbs of perception. In G. Cinque (ed.), *Italian Syntax and Universal Grammar*, pp. 244–275. Cambridge University Press.

Cinque, G. (1999). *Adverbs and Functional Heads: A cross-linguistic perspective*. Oxford University Press.

Cinque, G. (2010). On a selective "violation" of the complex NP constraint. In C. J. W. Zwart and de M. vries (eds.), *Structure Preserved: Studies in syntax for Jan Koster*, pp. 81–89. Amsterdam: John Benjamins.

Clahsen, H., and Featherston, S. (1999). Antecedent priming at trace positions, evidence from German scrambling. *Journal of Psycholinguistic Research* 28: 415–437.

Clark, H. H. (1976). *Semantics and Comprehension*. The Hague: Mouton.

Clark, S., and Gronlund, S. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin and Review* 3(1): 37–60.

Clemens, L. E., Morgan, A., Polinsky, M., and Xiang, M. (2012). Listening to resumptives: An auditory experiment. Poster presented at the 25th Annual CUNY Conference on Human Sentence Processing, New York, March 2012.

Clements, G. N. (1984). Binding domains in Kikuyu. *Studies in the Linguistic Sciences* 14: 37–56.

Clifton, C., Fanselow, G., and Frazier, L. (2006). Amnestying superiority violations: Processing multiple questions. *Linguistic Inquiry* 37(1): 51–68.

Clifton, C. E., Jr. and Frazier, L. (1989). Comprehending sentences with long-distance dependencies. In M. K. Tanenhaus and G. N. Carlson (eds.), *Linguistic Structure in Language Processing*, pp. 273–317. Dordrecht: Kluwer.

Cole, P., Harbert, W., Sridhar, S., Hashimoto, S., Nelson, C., and Smietana, D. (1977). Noun Phrase accessibility and Island Constraints. In P. Cole and J. M. Sadock (eds.), Syntax and Semantics Volume 8: Grammatical Relations, pp. 27–46. New York: Academic Press.

Cole, P., and Hermon, G. (1994). Is there LF *wh*-movement? *Linguistic Inquiry* 25: 239–262.

Comorovski, I. (1989a). *Discourse and the Syntax of Multiple Constituent Questions*. PhD thesis, Cornell University.

Comorovski, I. (1989b). Discourse-linking and the wh-island constraint. In J. Carter and R.-M. Déchaine (eds.), *Proceedings of the 19th Annual Meeting of the North East Linguistic Society*, pp. 78–96.

Comrie, B. (1987). *Language Universals and Linguistic Typology*. University of Chicago Press.

Conklin, K., Koenig, J. P., and Mauner, G. (2004). The role of specificity in the lexical encoding of participants. *Brain and Language* 90: 221–30.

Conway, A. R. A., Jarrold, C., Kane, M. J., Miyake, A., and Towse, J. (2007). *Variation in Working Memory*. Oxford University Press.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambridge, D. Z., Wilhelm, O., and Engle, R. W. (2005). Working memory span tasks: A methodological review and users guide. *Psychonomic Bulletin and Review* 12: 769–786.

Cooke, A., Zurif, E. B., DeVita, C., Alsop, D., Koenig, P., and Detre, J. (2002). Neural basis for sentence comprehension: Grammatical and short-term memory components. *Human Brain Mapping* 15(2): 80–94.

Cornish, E. R., and Wason, P. C. (1970). The recall of affirmative and negative sentences in an incidental learning task. *Quarterly Journal of Experimental Psychology* 22: 109–114.

Coulson, S., King, J. W., and Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes* 13: 21–58.

Cowan, N. (1995). *Attention and Memory: An integrated framework*. New York: Oxford University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences* 24: 87–114.

Cowan, N. (2005). *Working Memory Capacity*. Hove, East Sussex: Psychology Press.

Cowart, W. (1997). *Experimental Syntax: Applying objective methods to sentence judgments*. Thousand Oaks, CA: Sage Publications.

Crain, S., and Fodor, J. D. (1985). How can grammars help parsers? In D. Dowty, L. Karttunen, and A. Zwicky (eds.), *Natural Language Parsing: Psycholinguistic, computational, and theoretical approaches*, pp. 94–128. Cambridge University Press.

Crain, S., and Fodor, J. D. (1987). Sentence matching and overgeneration. *Cognition* 26: 123–169.

Crain, S., and Thornton, R. (1998). *Investigations in Universal Grammar*. Cambridge, MA: MIT Press.

Crawford, J. (2011). Using syntactic satiation to investigate subject islands. In *Proceedings of the 29th West Coast Conference on Formal Linguistics (WCCFL 29)*, pp. 38–45. Somerville, MA: Cascadilla Press.

Creswell, C. (2002). Resumptive pronouns, wh-island violations, and sentence production. In *Proceedings of the Sixth International Workshop on Tree-Adjoining Grammar and Related Frameworks (TAG+6)*, pp. 101–109. Venice: Università di Venezia.

Crocker, M., and Brants, T. (2000). Wide-coverage probabilistic sentence processing. *Journal of Psycholinguistic Research* 29: 647–669.

Croft, W. (2000). *Explaining Language Change: An evolutionary approach*. Harlow, Essex: Longman.

Croft, W. (2001). *Radical Construction Grammar*. Oxford University Press.

Culicover, P. (2001). Parasitic gaps: A history. In P. Culicover and P. Postal (eds.), *Parasitic Gaps*, pp. 3–68. Cambridge, MA: The MIT Press.

Culicover, P. W., and Jackendoff, R. (1997). Semantic subordination despite syntactic coordination. *Linguistic Inquiry* 28: 195–217.

Culicover, P., and Postal, P., eds. (2001). *Parasitic Gaps*. Cambridge, MA: MIT Press.

Dabrowska, E. (2004). *Language, Mind and Brain*. Washington, DC: Georgetown University Press.

Dabrowska, E., Rowland, C., and Theakston, A. (2009). The acquisition of questions with long-distance dependencies. *Cognitive Linguistics* 20: 571–97.

Dalrymple, M., Shieber, S. M., and Pereira, F. C. N. (1991). Ellipsis and higher-order unification. *Linguistics and Philosophy* 14: 399–452.

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior* 19: 450–466.

Deane, P. (1991). Limits to attention: A cognitive theory of island phenomena. *Cognitive Linguistics* 2: 1–63.

Demirdache, H. (1991). *Resumptive Chains in Restricted Relatives, Appositives and Dislocation Structures*. PhD Dissertation, MIT.

den Besten, H. (1985). The ergative hypothesis and free word order in Dutch and German. In Jindrich Toman (ed.), *Studies in German Grammar*, pp. 23–64. Dordrecht: Foris.

den Dikken, M. (2003). *The Structure of the Noun Phrase in Rotuman*. Munich: LINCOM Europa.

den Dikken, M., Bernstein, J., Tortora, C., and Zanuttini, R. (2007). Data and grammar: Means and individuals. *Theoretical Linguistics* 33(3): 335–352.

den Dikken, M., and Naess, A. (1993). Case dependencies: The case of predicate inversion. *The Linguistic Review* 10: 303–336.

de Villiers, J. (1995). Questioning minds and answering machines. In D. MacLaughlin and S. McEwan (eds.), *Proceedings of the 19th Boston University Conference on Language Development*, pp. 20–36. Somerville, MA: Cascadilla Press.

de Vincenzi, M. (1991). *Syntactic Parsing Strategies in Italian*. Dordrecht: Kluwer Academic.

Dickey, M. W. (1996). Constraints on the sentence processor and the distribution of resumptive pronouns. In M. W. Dickey and S. Tunstall (eds.), *Linguistics in the Laboratory*, pp. 157–192. Amherst, MA: University of Massachusetts.

Dickey, M. W., and Bunger, A. (2010). Comprehension of elided structure: Evidence from sluicing. *Language and Cognitive Process* 26: 63–78.

Diesing, M. (1992). *Indefinites*. Cambridge, MA: MIT Press.

Diessel, H. (2001). The ordering distribution of main and adverbial clauses: A typological study. *Language* 77: 345–365.

Diessel, H. (2009). The emergence of relative clauses in early child language. Unpublished MS, University of Jena.

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

Dresher, E. (1999). Charting the learning path: Cues to parameter setting. *Linguistic Inquiry* 30: 27–67.

Dryer, M. S. (1986). Primary Objects, Secondary Objects and Antidative. *Language* 62: 808–845.

Eberhard, K. M., Cutting, J. C., and Bock, K. (2005). Making sense of syntax: Number agreement in sentence production. *Psychological Review* 112: 531–559.

Eide, K. and Åfarli, T. (1999). The syntactic disguises of the predication operator. *The Linguistic Review* 52: 155–181.

Ellefson, M. R., and Christiansen, M. H. (2000). Subjacency constraints without Universal Grammar: Evidence from artificial language learning and connectionist modeling. In L. R. Gleitman and A. K. Joshi (eds.), *The Proceedings of the 22nd Annual Conference of the Cognitive Science Society*, pp. 645–50. Mahwah, NJ: Lawrence Erlbaum.

Ellis, N. C. (2002). Frequency Effects in language processing: A review with implications for theories of implicit and explicit. *Studies in Second Language Acquisition* 24: 143–188.

Engdahl, E. (1980). Wh-constructions in Swedish and the relevance of subjacency. In J. T. Jensen (ed.), *Cahiers Linguistiques d'Ottawa: Proceedings of the Tenth Meeting of the North East Linguistic Society*, pp. 89–108. Ottawa, ONT: University of Ottawa Department of Linguistics.

Engdahl, E. (1982). Restrictions on unbounded dependencies in Swedish. In E. Engdahl and E. Ejerhed (eds.), *Readings on Unbounded Dependencies in Scandinavian Languages*, pp. 151–174. Stockholm: Almqvist & Wiksell.

Engdahl, E. (1983). Parasitic gaps. *Linguistic Inquiry* 6: 5–34.

Engdahl, E. (1986). *Constituent Questions: The syntax and semantics of questions with special reference to Swedish*. Dordrecht: Reidel.

Engdahl, E. (1997). Relative clause extractions in context. *Working Papers in Scandinavian Syntax* 60: 51–79.

Engelhardt, P., Bailey, K., and Ferreira, F. (2006). Do speakers and listeners observe the Gricean Maxim of Quantity? *Journal of Memory and Language* 54: 554–573.

Ericsson, K. A., and Kintsch, W. (1995). Long-term working memory. Psychological Review 102: 211–245.

Eriksen, C. W., and Schultz, D. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception and Psychophysics* 25: 249–263.

Erteschik-Shir, N. (1973). *On the Nature of Island Constraints*. PhD dissertation, MIT.

Erteschik-Shir, N. (1977). *On the Nature of Island Constraints*. Bloomington: Indiana University Linguistics Club.

Erteschik-Shir, N. (1979). Discourse constraints on dative movement. In S. Laberge and G. Sankoff (eds.), *Syntax and Semantics*, pp. 441–467. New York: Academic Press.

Erteschik-Shir, N. (1992). Resumptive pronouns in islands. In H. Goodluck and M. Rochemont (eds.), *Island Constraints: Theory, acquisition and processing*, Volume 15 of *Studies in Theoretical Psycholinguistics*, pp. 89–108. Dordrecht: Kluwer Academic Publishers.

Erteschik-Shir, N. (2007). *Information Structure: The syntax–discourse interface*. Oxford University Press.

Evans, N., and Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32: 429–492.

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

Featherston, S. (2008). Thermometer judgments as linguistic evidence. In M. Claudia and A. Rothe (eds.), *Was ist linguistische Evidenz?*, pp. 69–89. Aachen: Shaker Verlag.

Feldman, N., Griffiths, T., and Morgan, J. (2009). The influence of categories on perception: Explaining the perceptual magnet effect as optimal statistical inference. *Psychological Review* 116: 752–782.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology* 47: 164–203.

Ferreira, F., and Swets, B. (2005). The production and comprehension of resumptive pronouns in relative clause "island" contexts. In Anne Cutler (ed.), *Twenty-first Century Psycholinguistics: Four cornerstones*, pp. 263–278. Mahwah, NJ: Lawrence Erlbaum Associates.

Ferreira, V. S. (2008). Ambiguity, accessibility, and a division of labor for communicative success. *Psychology of Learning and Motivation* 49: 209–246.

Ferreira, V. S., and Dell, G. S. (2000). The effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology* 40: 296–340.

Fiebach, C. J., Schlesewsky, M., and Friederici, A. D. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German *wh*-questions. *Journal of Memory and Language* 47: 250–272.

Fiengo, R., and Higginbotham, J. (1981). Opacity in NP. *Linguistic Analysis* 7: 395–421.

Fillmore, C. J. (1965). *Indirect Object Constructions in English and the Ordering of Transformations*. The Hague: Mouton.

Fodor, J. A., Bever, T. G., and Garrett, M. F. (1974). *The Psychology of Language: An introduction to psycholinguistics and generative grammar*. New York: McGraw-Hill.

Fodor, J. D. (1978). Parsing strategies and constraints on transformations. *Linguistic Inquiry* 9: 427–473.

Fodor, J. D. (1983). Phrase structure parsing and the island constraints. *Linguistics and Philosophy* 6: 163–223.

Fodor, J. D. (1998a). Unambiguous triggers. *Linguistic Inquiry* 29: 1–36.

Fodor, J. D. (1998b). Parsing to learn. *Journal of Psycholinguistic Research* 27(3): 339–374.

Fodor, J. D. (2009). Syntax acquisition: An evaluation measure after All? In M. Piatelli Palmarini, J. Uriagereka, and P. Salaburu (eds.), *Of Minds and Language: The Basque country encounter with Noam Chomsky*, pp. 256–277. New York: Oxford University Press.

Fodor, J. D., and Inoue, A. (1994). The diagnosis and cure of garden paths. *Journal of Psycholinguistic Research* 23(5): 407–434.

Fodor, J. D., and Inoue, A. (1998). Attach anyway. In J. D. Fodor and F. Ferreira (eds.), *Reanalysis in Sentence Processing*, pp. 101–141. Dordrecht: Kluwer Academic Publishers.

Foraker, S., Regier, T., Khetarpal, A., Perfors, A., and Tenenbaum, J. (2009). Indirect evidence and the poverty of the stimulus: The case of anaphoric *one. Cognitive Science* 33: 287–300.

Ford, M. (1983). A method of obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior* 22: 203–218.

Forster, K. I. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper and E. Walker (eds.), *Sentence Processing: Psycholinguistic essays presented to Merrill Garrett*, pp. 27–75. Hillsdale, N.J.: Erlbaum.

Forster, K. I., and Stevenson, B. J. (1987). Sentence matching and well-formedness. *Cognition* 26: 171–186.

Fox, D., and Hackl, M. (2006). The universal density of measurement. *Linguistics and Philosophy* 29: 537–586.

Fox, D., and Nissenbaum, J. (1999). Extraposition and scope: a case for overt QR. In S. Bird, A. Carnie, J. D. Haugen, and P. Norquest (eds.), *Proceedings of the 18th West Coast Conference on Formal Linguistics*, pp. 25–36. Sommerville, MA: Cascadilla Press.

Francis, H. S., and Michaelis, L. A. (2007). Lexical subjects and the conflation strategy. In N. Hedberg and R. Zacharski (eds.), *Topics in the Grammar–Pragmatics Interface: Papers in Honor of Jeanette K. Gundel*, pp. 19–48. Amsterdam: Benjamins.

Francom, J. (2009). *Experimental Syntax: Exploring the effect of repeated exposure to anomalous syntactic structure – evidence from rating and reading tasks*. PhD dissertation, University of Arizona.

Frank, M. C., Goodman, S., and Tenenbaum, J. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science* 20(5): 578–585.

Frank, R. (2002). *Phrase Structure Composition and Syntactic Dependencies*. Cambridge, MA: The MIT Press.

Frazier, L. (1985). Syntactic complexity. In D. Dowty, L. Karttunen, and A. Zwicky (eds.), *Natural Language Processing: Psychological, computational, and theoretical perspectives*, pp. 129–189. Cambridge University Press.

Frazier, L. (1987a). Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory* 5: 519–560.

Frazier, L. (1987b). Sentence processing: A tutorial review. In M. Coltheart (ed.), *Attention and Performance XII: The psychology of reading*, pp. 559–586. London: Lawrence Erlbaum Associates.

Frazier, L., and Clifton, C. (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes* 4: 93–126.

Frazier, L., and Clifton, C. Jr. (1998). Comprehension of sluiced sentences. *Language and Cognitive Process* 13: 499–520.

Frazier, L., and Clifton, C. Jr. (2000). Parsing coordinates and ellipsis: Copy alpha. *Syntax* 4: 1–22.

Frazier, L., and Clifton, C. Jr. (2001). The syntax–discourse divide: Processing ellipsis. *Syntax* 8: 121–174.

Frazier, L., and Clifton, C. Jr. (2002). Processing 'd-linked' phrases. *Journal of Psycholinguistic Research* 31: 633–660.

Frazier, L., and Flores d'Arcais, G. (1989). Filler-driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language* 28: 331–344.

Frazier, L., and Fodor, J. D. (1978). The sausage machine. A new two-stage parsing model. *Cognition* 6(4): 291–325.

Frazier, L., Lori, T., Roeper, T., Clifton, C. Jr., and Ehrlich, K. (1984). Parallel structure: A source of facilitation in sentence comprehension. *Memory and Cognition* 12: 421–430.

Frazier, L., Munn, A., and Clifton, C. Jr. (2000). Processing Coordinate Structures. *Journal of Psycholinguistic Research* 29: 343–370.

Freedman, S., and Forster, K. I. (1985). The psychological status of overgenerated sentences. *Cognition* 19: 101–131.

Friedmann, N., Taranto, G., Shapiro, L., and Swinney, D. (2008). The leaf fell (the leaf): The online processing of unaccusatives. *Linguistic Inquiry* 39: 355–377.

Garavan, H. (1998). Serial attention within working memory. *Memory and Cognition* 26: 263–276.

Garnsey, S. M., Tanenhaus, M. K., and Chapman, R. M. (1989). Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research* 18: 51–60.

Gazdar, G. (1981). Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12: 155–184.

Gazdar, G., Klein, E., Pullum, G. K., and Sag, I. A. (1985). *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.

Georgopoulos, C. (1985). Variables in Palauan syntax. *Natural Language & Linguistic Theory* 3: 59–94.

Georgopoulos, C. (1991). *Syntactic Variables: Resumptive pronouns and binding in Palauan*. Dordrecht: Kluwer.

Gerken, L. (2006). Decision, decisions: Infant language learning when multiple generalizations are possible. *Cognition* 98: B67–B74.

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

Gibson, E. (2000). *The dependency locality theory, a distance-based theory of linguistic complexity*. In Y. Miyashita, A. Marantaz, and W. O'Neil (eds.), *Image, Language, Brain*, pp. 95–126. Cambridge, MA: MIT Press.

Gibson, E., and Fedorenko, E. (2013). The need for quantitative methods in syntax and semantics research. *Language and Cognitive Processes* 28(1–2): 88–124.

Gibson, E., and Grodner, D. J. (2005). Consequences of the serial nature of linguistic input for sentential complexity. *Cognitive Science* 29: 261–291.

Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E., and Hickok, G. (1996). Recency preference in the human sentence processing mechanism. *Cognition* 59: 2359.

Gibson, E., and Thomas, J. (1999). Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes* 14: 225–248.

Gibson, E., and Wexler, K. (1994). Triggers. *Linguistic Inquiry* 25: 355–407.

Gieselman, S., Kluender, R., and Caponigro, I. (2011). Pragmatic processing factors in negative island contexts. In *Proceedings of the 22nd Annual Meeting of the Western Conference on Linguistics*, pp. 65–76.

Gieselman, S., Kluender, R., and Caponigro, I. (in press). Isolating processing factors in negative island contexts. *Proceedings of the 41st Annual Meeting of the North East Linguistic Society*.

Gilligan, G. (1987). *A cross-linguistic approach to the pro–drop parameter*. Ph.D. Dissertation, University of Southern California.

Gillund, G., and Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review* 91(1): 1–67.

Givón, T. (1979). *On Understanding Grammar*. New York: Academic Press.

Givón, T. (1984). *Syntax: A functional-typological introduction*. Amsterdam: John Benjamins.

Glenberg, A. M., Robertson, D. A., Jansen, J. L., and Johnson-Glenberg, M. C. (1999). Not propositions. *Journal of Cognitive Systems Research* 1: 19–33.

Glushan, Z. (2006). *Japanese style scrambling in Russian: Myth and reality*. Master's thesis, University of Tromsoe.

Goldberg, A. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford University Press.

Goldsmith, J. (1985). A principled exception to the Coordinate Structure Constraint. In *Papers from the Twenty-first Annual Regional Meeting of the Chicago Linguistic Society*, pp. 133–143. Chicago Linguistic Society.

Goldwater, S., Griffiths, T., and Johnson, M. (2009). A Bayesian framework for word segmentation: Exploring the effects of context. *Cognition* 112(1): 21–54.

Gomez, R. (2002). Variability and detection of invariant structure. *Psychological Science* 13: 431–436.

Goodall, G. (2001). The EPP in Spanish. In W. Davies and S. Dubinsky (eds.), *Objects and Other Subjects: Grammatical functions, functional categories, and configurationality*, pp. 193–223. Dordrecht: Kluwer.

Goodall, G. (2002). On preverbal subjects in Spanish. In T. Satterfield, C. Tortora, and D. Cresti (eds.), *Current Issues in Romance Languages: Selected papers from the 29th Linguistic Symposium on Romance languages*, pp. 95–109. Amsterdam: John Benjamins.

Goodall, G. (2004). On the syntax and processing of wh-questions in Spanish. In B. Schmeiser, V. Chand, A. Kelleher, and A. Rodriguez (eds.), *Proceedings of the 23rd West Coast Conference on Formal Linguistics*, pp. 101–114. Somerville, MA: Cascadilla Press.

Goodall, G. (2011). Syntactic satiation and the inversion effect in English and Spanish *wh*-questions. *Syntax* 14: 29–47.

Goodluck, H., and Tavakolian, S. L. (1982). Competence and processing in children's grammar of relative clauses. *Cognition* 11: 1–27.

Gordon, P. and Hendrick, R. (1998). The representation and processing of coreference in discourse. *Cognitive Science* 22: 389–424.

Gordon, P. C., Hendrick, R., and Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27: 1411–1423.

Gouvea, A., Phillips, C., Kazanina, N., and Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes* 25: 149–188.

Graf Estes, K., Evans, J., Alibali, M., and Saffran, J. (2007). Can infants map meaning to newly segmented words? *Psychological Science* 18(3): 254–260.

Grewendorf, G. (1989). *Ergativity in German*. Dordrecht: Foris.
Gribanova, V. (2010). *Composition and locality: The morphosyntax and phonology of the Russian verb complex*. PhD dissertation, University of California Santa Cruz.
Grice, P. (1969). Utterer's meaning and intentions. *Philosophical Review* 78: 147–177.
Grice, P. (1975). Logic and conversation. In P. Cole and J. Morgan (eds.), *Syntax and Semantics 3: Speech Acts*, pp. 41–58. New York: Academic Press.
Grice, P. (1978). Further notes on logic and conversation. In P. Cole (ed.), *Syntax and Semantics 9: Pragmatics*, pp. 113–128. New York: Academic Press.
Griffiths, T., and Tenenbaum, J. (2005). Structure and strength in causal induction. *Cognitive Psychology* 51: 334–384.
Grimshaw, J. (1986). Subjacency and the S/S′ Parameter. *Linguistic Inquiry* 17: 364–369.
Grosz, B. J., Joshi, A. K., and Weinstein, S. (1983). Providing a unified account of definite noun phrases in discourse. In *Proceedings, 21st Annual Meeting of the Association of Computational Linguistics*, pp. 44–50.
Guasti, M. T. (2002). *Language Acquisition: The growth of grammar*. Cambridge, MA: MIT Press.
Gundel, J. (1974). *The role of topic and comment in linguistic theory*. PhD dissertation, University of Texas at Austin.
Gundel, J. K., Hedberg, N., and Zacharski R. (1993). Cognitive status and the form of referring expressions in discourse. *Language* 69: 274–307
Haegeman, L. (1992). Negation in west Flemish and the Neg criterion. In *Proceedings of the 22nd Annual Meeting of the North East Linguistic Society*, pp. 195–211.
Haegeman, L. (1995). *The Syntax of Negation*. Cambridge University Press.
Hagstrom, P. (1998). *Decomposing Questions*. Cambridge, MA: MIT dissertation.
Hahne, A., and Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience* 11: 194–205.
Haider, H. (1983). Connectedness effects in German. *Groninger Arbeiten zur Germanistischen Linguistik* 25: 83–117.
Haider, H. (1993). *Deutsche Syntax-generativ*. Tübingen: Narr.
Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics*, pp. 159–166.
Halliday, A. K. (1967). Notes on transitivity and theme in English: Part II. *Journal of Linguistics* 3: 199–244.
Han, C., and Kim, J. (2004). "Double relative clauses" in Korean? *Linguistic Inquiry* 35: 315–337.
Han, C., Kirby, S., Elouazizi, N., Göorgülü, N., Hedberg, N., Dyakonova, M., Galeano, C., Hinnell, J., Jeffrey, M., and Kim, K. (2012). Subject-object asymmetry in English resumption. Paper presented at the 86th Annual Meeting of the Linguistic Society of America, Portland, OR, January 2012.
Hardt, D. (1993). *Verb Phrase ellipsis: Form meaning and processing*. PhD Dissertation, University of Pennsylvania, Philadelphia.
Hardt, D. (1999). Dynamic interpretation of Verb Phrase ellipsis. *Linguistics and Philosophy* 22: 187–221.

Harizanov, B. (2011). The acceptability of true resumption: An acceptability rating study in Bulgarian. Unpublished MS, University of California Santa Cruz.

Hart, B., and Risley, T. R. (1995). *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore, MD: P. H. Brookes.

Harves, S. (2002). *Unaccusative syntax in Russian*. PhD dissertation, Princeton University.

Harves, S. (2003). Getting impersonal: Case, agreement, and distributive *po*-phrases in Russian. In W. Browne *et al.* (eds.), *Formal Approaches to Slavic Linguistics, Amherst Meeting*, pp. 235–254. Ann Arbor: Michigan Slavic Publications.

Hasegawa, N. (1981). *A lexical interpretive theory with emphasis on the role of subject*. PhD dissertation, University of Washington.

Hawkins, J. A. (1999). Processing complexity and filler-gap dependencies across grammars. *Language* 75: 244–285.

Hawkins, J. (2004). *Efficiency and Complexity of Grammars*. Oxford University Press.

Hayes, J. R. M. (1952). Memory span for several vocabularies as a function of vocabulary size. In *Quarterly Progress Report*, pp. 5–9. Cambridge, MA: Acoustics Laboratory, MIT.

Heestand, D. (2010). *Conditions on the cross-linguistic acceptability of resumptive pronouns*. Senior thesis, Harvard University.

Heestand, D., Xiang, M., and Polinsky, M. (2011). Resumption still does not rescue islands. *Linguistic Inquiry* 42: 138–152.

Hendrick, R. (2005). Resumptive and bound variable pronouns in Tongan. In J. Heinz and D. Ntheliteos (eds.), *Proceedings of AFLA-XII*, pp. 103–115. Los Angeles: UCLA Dept. of Linguistics.

Hinton, G. E. (1989). Connectionist learning procedure. *Artificial Intelligence* 40: 185–234.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review* 95: 528–551.

Hiramatsu, K. (1999). Subject and adjunct island asymmetries: Evidence from syntactic satiation. In S. Bird, A. Carnie, J. D. Haugen, and P. Norquest (eds.), *Proceedings of the West Coast Conference on Formal Linguistics* 18, pp. 183–192. Somerville, MA: Cascadilla Press.

Hiramatsu, K. (2000). *Accessing linguistic competence: Evidence from children's and adults' acceptability judgments*. PhD dissertation, University of Connecticut.

Hofmeister, P. (2007b). *Representational complexity and memory retrieval in language comprehension*. Stanford, CA: Stanford University dissertation.

Hofmeister, P. (2011). Representational complexity and memory retrieval in language comprehension. *Language and Cognitive Processes* 26 (3): 376–405.

Hofmeister, P., Jaeger, T. F., Arnon, I., Sag, I. A., and Snider, N. (2013). The source ambiguity problem: Distinguishing the effects of grammar and processing on acceptability judgments. *Language and Cognitive Processes* 28(1–2): 48–87.

Hofmeister, P., Jaeger, T. F., Sag, I. A., Arnon, I., and Snider, N. (2007). Locality and accessibility in wh-questions. In S. Featherston and W. Sternefeld (eds.), *Roots: Linguistics in Search of its Evidential Base*, pp. 185–206. Berlin: Mouton de Gruyter.

Hofmeister, P., and Norcliffe, E. (2013). Does resumption facilitate sentence comprehension. In P. Hofmeister and E. Norcliffe (eds.), *The Core and the Periphery: Data-driven Perspectives on Syntax Inspired by Ivan A. Sag*, pp. 225–246. Stanford, CA: CSLI Publications.

Hofmeister, P., and Sag, I. A. (2010). Cognitive constraints on syntactic islands. *Language* 86: 366–415.

Hoshi, K. (2004). Parameterization of the external D-system in relativization. *Language, Culture, and Communication* 33: 1–50.

Hsu, C.-C. N. (2006). *Issues in head-final relative clauses in Chinese: Derivation, processing, and acquisition*. PhD dissertation, University of Delaware.

Huang, C-T. J. (1982a). *Logical relations in Chinese and the theory of grammar*. Cambridge, MA: MIT dissertation.

Huang, C.-T. J. (1982b). Move wh in a language without wh-movement. *The Linguistic Review* 1: 369–416.

Huang, C.-T. J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15: 531–574.

Huang, Y. (2000). *Anaphora: A cross-linguistic approach*. Oxford University Press.

Hunter, T. (2010). *The syntax and semantics of movement and adjunction in natural language*. Doctoral dissertation, University of Maryland.

Inoue, K. (1976). *Henkei-bunpo to nihongo*. Tokyo: Taishukan.

Ishizuka, T. (2009). CNPC violations and possessor raising in Japanese. Unpublished MS, UCLA.

Jackendoff, R. S., and Culicover, P. (1971). A reconsideration of dative movements. *Foundations of Language* 7(3): 397–412.

Jaeger, F. (2006). *Redundancy and syntactic reduction in spontaneous speech*. PhD dissertation, Stanford University.

Jaeger, T. F., and Snider, N. (2008). Implicit learning and syntactic persistence: Surprisal and cumulativity. In *Proceedings of the 30th Annual Meeting of the Cognitive Science Society*, pp. 1061–1066.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., and Perrig, W. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences* 105: 6829–6833.

James, D. (1972). Some aspects of the syntax and semantics of interjections. Paper presented to the 8th Regional Meeting of the Chicago Linguistic Society.

James, W. (1890). *The Principles of Psychology*, 2 vols. New York: Holt.

Jha, A. P., and McCarthy, G. (2000). The influence of memory load upon delay-interval activity in a working-memory task: An event-related functional MRI study. *Journal of Cognitive Neuroscience* 12(2): 90–105.

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., and Moore, K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology* 59(1): 193–224.

Junghanns, U., and Zybatow, G. (1997). Syntax and information structure of Russian clauses. In W. Browne *et al*. (eds.), *Formal Approaches to Slavic Linguistics* 4: *Cornell Meeting*, pp. 289–319. Ann Arbor: Michigan Slavic Publications.

Jurka, J. (2010). *The importance of being a complement: CED effects revisited*. PhD dissertation, University of Maryland.

Just, M., and Carpenter, P. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review* 87: 123–154.

Just, M., and Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review* 99: 122–149.

Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., and Thulborn, K. R. (1996). Brain activation modulated by sentence comprehension. *Science* 274: 114–116.

Just, M. A., Carpenter, P. A., and Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology* 111: 228–238.

Kaan, E. (1997). *Processing subject–object ambiguities in Dutch*. Doctoral dissertation, University of Groningen.

Kaan, E., Harris, A., Gibson, E., and Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes* 15: 159–201.

Kallestinova, E. (2007). *Aspects of Word Order in Russian*. PhD Dissertation, University of Iowa.

Kane, M. J., Conway, A. R. A., Miura, T. K., and Colflesh, G. J. H. (2007). Working memory, attention control, and the n-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33: 615–622.

Kane, M. J., and Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual differences perspective. *Psychonomic Bulletin and Review* 9: 637–671.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., and Engle, R. W. (2004). The generality of working memory capacity, a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General* 133: 189–217.

Kang, Y.-S. (1986). *Korean syntax and Universal Grammar*. PhD dissertation, Harvard University.

Kanno, K., and Nakamura, M. (2001). Processing of relative clauses by Japanese native speakers and L2 learners. *Journal of Association for Japanese Language Education* 4: 134–148.

Karttunen, L. (1976). Discourse referents. In J. D. McCawley (ed.), *Syntax and Semantics* 7: *Notes From the Linguistic Underground*, pp. 48–75. New York: Academic Press.

Karttunen, L. (1977). Syntax and semantics of questions. *Linguistics and Philosophy* 1: 3–44.

Kaschak, M. P., and Glenberg, A. M. (2004). This construction needs learned. *Journal of Experimental Psychology: General* 133: 450–467.

Kaup, B. (2006). What psycholinguistic negation research tells us about the nature of the working memory representations utilized in language comprehension. In H. Pishwa (ed.), *Language and Memory: Aspects of knowledge representation*, pp. 313–355. Berlin: Mouton.

Kayne, R. (1983). Connectedness. *Linguistic Inquiry* 14: 223–249.

Kayne, R. (1993). Toward a modular theory of auxiliary selection. *Studia Linguistica* 47: 3–31.

Kayne, R. S. and Pollock, J.-Y. (1978). Stylistic Inversion, Successive Cyclicity, and Move NP in French. *Linguistic Inquiry* 9(4): 595–621.

Kazanina, N., Lau, E. F., Lieberman, M., Yoshida, M., and Phillips, C. (2007). The effect of syntactic constraints on the processing of backwards anaphora. *Journal of Memory and Language* 56: 384–409.

Kazanina, N., and Phillips, C. (2010). Differential effects of constraints in the processing of Russian cataphora. *Quarterly Journal of Experimental Psychology* 63: 371–400.

Keenan, E. L., and Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1): 63–99.

Kehler, A. (2002). *Coherence, Reference, and the Theory of Grammar*. Stanford: CSLI Publications.

Keller, F. (2000). *Gradience in grammar: Experimental and computational aspects of degress of grammaticality*. Doctoral Dissertation, University of Edinburgh.

Keller, F., Gunasekharan, S., Mayo, N., and Corley, M. (2009). Timing accuracy of web experiments: A case study using the WebExp software package. *Behavior Research Methods* 41 (1): 1–12.

Kempen, G., and Harbusch, K. (2004). The relationship between grammaticality ratings and corpus frequencies: A case study into word order variability in the midfield of German clauses. In S. Kepser and M. Reis (eds.), *Linguistic Evidence: Empirical, theoretical, and computational Perspectives*, pp. 329–349. Berlin: Mouton De Gruyter.

Kenstowicz, M. (1989). The null subject parameter in modern Arabic dialects. In O. Jaeggli and K. Safir (eds.), *The Null Subject Parameter*, pp. 263–275. Dordrecht: Kluwer.

King, T. H. (1995). *Configuring Topic and Focus in Russian*. Stanford: CSLI.

King, J., and Just, M. A. (1991). Individual differences in syntactic processing, the role of working memory. *Journal of Memory and Language* 30: 580–560.

King, J., and Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience* 7: 376–395.

Kiparsky, P., and Kiparsky, C. (1971). *Fact*. In M. Bierwisch and K. Heidolph (eds.), *Progress in Linguistics*, pp. 143–173. The Hague: Mouton.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology* 55: 352–358.

Kiss, K. É. (1987). *Configurationality in Hungarian*. Dordrecht: Reidel.

Kiss, K. É (1995). *Introduction*. In K. É Kiss (ed.), *Discourse Configurational Languages*, pp. 3–35. Oxford University Press.

Kitagawa, Y., Tamaoka, K., and Tomioka, S. (2013). Prosodic matters in intervention effects in Japanese: An experimental study. *Lingua* 124: 41–63.

Kjellmer, G. (1975). Are relative infinitives modal? *Studia Neophilologica* 47: 323–332.

Klein, D. and Manning, C. (2002). A generative constituent-context model for improved grammar induction. In *Proceedings of the 40th Annual Meeting for the Association for Computational Linguistics*, pp. 128–135. Stroudsburg, PA: Association for Computational Linguistics.

Klima, E. (1964). Negation in English. In Fodor and Katz (eds.), *The Structure of language*, pp. 246–323. Englewood Cliffs, NJ: Prentice Hall.

Kluender, R. (1991). *Cognitive constraints on variables in syntax*. Doctoral dissertation, University of California, San Diego, La Jolla.

Kluender, R. (1992). Deriving island constraints from principles of predication. In H. Goodluck and M. Rochemont (eds.), *Island Constraints: Theory, acquisition and processing*, Volume 15 of *Studies in Theoretical Psycholinguistics*, pp. 223–258. Dordrecht: Kluwer Academic Publishers.

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

Kluender, R. (2004). Are Subject islands subject to a processing account? In V. Chand, A. Kelleher, A. Rodriguez, and B. Schmeiser (eds.), *Proceedings of the West Coast Conference on Formal Linguistics 23*, pp. 475–499. Somerville, MA: Cascadilla Press.

Kluender, R., and Kutas, M. (1993a). Bridging the gap – evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience* 5: 196–214.

Kluender, R., and Kutas, M. (1993b). Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8: 573–633.

Knoeferle, P., and Crocker, M. W. (2009). Constituent order and semantic parallelism in online comprehension: Eye-tracking evidence from German. *Quarterly Journal of Experimental Psychology* 62: 2338 – 2371.

Kohonen, T. (1977). *Associative Memory: A system theoretical approach*. New York: Springer-Verlag.

Koopman, H. (1984). *The Syntax of Verbs: From verb movement rules in the Kru languages to Universal Grammar*. Dordrecht: Foris.

Koopman, H., and Sportiche, D. (1986). A note on long extraction in Vata and the ECP. *Natural Language and Linguistic Theory* 4: 357–374.

Kothari, A. (2008). Frequency-based expectations and context influence bridge quality. Unpublished MS, Stanford University.

Kravtchenko, E., Polinsky, M., and Xiang, M. (2009). Are all subject islands created equal? Poster presented at CUNY 2009.

Kroch, A. (1981). On the role of resumptive pronouns in amnestying island constraint violations. In R. A. Hendrick, C. S. Masek, and M. F. Miller (eds.), *Papers from the 17th Regional Meeting of the Chicago Linguistic Society*, pp. 125–135. Chicago Linguistic Society, University of Chicago.

Kroch, A. (1989). Amount quantification, referentiality, and long *wh*-movement. Unpublished MS, University of Pennsylvania.

Kuno, S. (1972). Functional sentence perspective: A case study from Japanese and English. *Linguistic Inquiry* 3: 269–320.

Kuno, S. (1973). *The Structure of the Japanese Language*. Cambridge, MA: MIT Press.

Kuno, S. (1976). Subject, theme, and speaker's empathy: A reexamination of relativization phenomena. In C. Li (ed.), *Subject and Topic*, pp. 417–444. New York: Academic Press.

Kuno, S. (1987). *Functional Syntax: Anaphora, discourse and empathy*. Chicago: University of Chicago Press.

Kuno, S., and Takami, K. (1997). Remarks on negative islands. *Linguistic Inquiry* 28: 553–576.

Kuroda, S.-Y. (1988). Whether we agree or not: A comparative syntax of English and Japanese. In W. Poser (ed.), *Papers from the Second International Workshop on Japanese Syntax*, pp. 103–143. Stanford, CA: CSLI.

Kush, D. (2011). On the (in-)escapability of islands in Scandinavian. Unpublished MS, University of Maryland.

Kwon, N., Lee, Y., Gordon, P. C., Kluender, R., and Polinsky, M. (2010). Cognitive and linguistic determinants of the subject–object asymmetry: An eye-tracking study of pre-nominal relative clauses in Korean. *Language* 86: 546–582.

Lakoff, G. (1986). Frame semantic control of the Coordinate Structure Constraint. In A. Farley, P. Farley, and K.-E. McCullough (eds.), *Papers from the Parasession on Pragmatics and Grammatical Theory at the 22nd Regional Meeting of the Chicago Linguistics Society*, pp. 152–167. Chicago Linguistic Society, University of Chicago.

Lambrecht, K. (1994). *Information Structure and Sentence Form*. Cambridge University Press.

Lane, L. W., and Ferreira, V. (2008). Speaker-external versus speaker-internal forces on utterance form: Do cognitive demands override threats to referential success? *Journal of Experimental Psychology: Learning, Memory & Cognition* 34: 1466–1481.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar*, Volume 1. Stanford, CA: Stanford University Press.

Langendoen, D. T., Kalish-Landon, N., and Dore, J. (1973). Dative questions: a study in the relation of acceptability to grammaticality of an English sentence type. *Cognition* 2: 451–78.

Larson, R. K. (1988). On the double object construction. *Linguistic Inquiry* 19: 335–391.

Lasnik, H. (1992). Case and expletives. *Linguistic Inquiry* 23: 381–405.

Lasnik, H. (1999). On the locality of movement. In M. Darnell, E. Moravcsik, F. Newmeyer, M. Noonan, and K. Wheatley (eds.), *Functionalism and Formalism in linguistics*, Volume 1: *General Papers*, pp. 33–54. Amsterdam: John Benjamins.

Lasnik, H. (2001). When can you save a structure by destroying It? In M. Kim and U. Strauss (eds.), *Proceedings of the 31st Annual Meeting of the North East Linguistic Society*, pp. 301–320.

Lasnik, H. (2005). Review of the *Syntax of Silence* by Jason Merchant. *Language* 81: 259–265.

Lasnik, H., and Park, M.-K. (2003). The Epp and the Subject Condition under Sluicing. *Linguistic Inquiry* 34: 649–660.

Lasnik, H., and Saito, M. (1984). On the nature of proper government. *Linguistic Inquiry* 15: 235–289.

Lasnik, H., and Saito, M. (1992). *Move Alpha: Conditions on its application and output*. Cambridge, MA: MIT Press.

Lau, E. F., Yeung, H. H., Hashimoto, R., Braun, A., and Phillips, C. (2006). Time-course and localization of syntactic and semantic anomaly responses in sentence processing, a within-subjects fMRI/MEG design. Poster presented at the 36th Society for Neuroscience Annual Meeting, Atlanta, GA.

Leddon, E. M., and Lidz, J. L. (2006). Reconstruction effects in child language. In D. Bamman, T. Magnitkaia, and C. Zaller (eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development*, pp. 328–339. Somerville, MA: Cascadilla Press.

Lee, M.-W. (2004). Another look at the role of empty categories in sentence processing (and grammar). *Journal of Psycholinguistic Research* 33: 51–73.

Lehmann, C. (1987). *Der Relativsatz: Typologie seiner Strukturen, Theorie seiner Funktionen, Kompendium seiner Grammatik*. Tübingen: Narr.

Levin, B., and Rappaport Hovav, M. (1995). *Unaccusativity: At the Syntax–Lexical Semantics Interface*. Cambridge, MA: MIT Press.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition* 106: 1126–1177.

Levy, R. (2011a). Integrating surprisal and uncertain-input models in online sentence comprehension: formal techniques and empirical results. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pp. 1055–1065.

Levy, R. (2011b). Resumption is not more common where it is least acceptable. Unpublished MS, University of California San Diego.

Lewis, R. L. (1996). Interference in short-term memory: The magical number two (or three) in sentence processing. *Journal of Psycholinguistic Research* 25(1): 93–115.

Lewis, R. L. (2000). Specifying architectures for language processing: Process, control, and memory in parsing and interpretation. In M. W. Crocker, M. J. Pickering, and C. Clifton (eds.), *Architectures and Mechanisms for Language Processing*, pp. 56–89. Cambridge University Press.

Lewis, R. L., and Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29: 375–419.

Lewis, R. L., Vasishth, S., and Van Dyke, J. A. (2006). Computational principles of working memory in sentence processing. *Trends in Cognitive Sciences* 10: 447–454.

Li, Y.-H. A. (2002). Word order, Structure and Relativization. In Sze-Wing Tang and C.-S. Luther Liu (eds.), *On the Formal Way to Chinese Languages*, pp. 45–73. Stanford: CSLI Publications.

Lidz, J., and Williams, A. (2009). Constructions on holiday: Commentary on Goldberg (2006). *Cognitive Linguistics* 20: 177–189.

Lightfoot, D. (1991). *How to Set Parameters: Arguments from language change*. Cambridge, MA: MIT Press.

Lightfoot, D. (2010). Language acquisition and language change. *Wiley Interdisciplinary Reviews: Cognitive Science* 1: 677–684.

Linkins, J. (2011). What do journalists think about the debt crisis? They think that Americans got what they deserved. *Huffington Post*, August 1, 2011. [www.huffingtonpost.com/2011/08/01/journalists-debt-ceiling-crisis_n_915531.html

Lohndal, T. (2009). Comp-t effects: Variation in the position and features of C. *Studia Linguistica* 63: 204–232.

Lohndal, T. (2011). Freezing effects and objects. *Journal of Linguistics* 47: 163–199.

Lüdtke, J., Friedrich, C. K., De Filippis, M., and Kaup, B. (2008). Event-related potential correlates of negation in a sentence-picture verification paradigm. *Journal of Cognitive Neuroscience* 20: 1355–1370.

Lüdtke, J. and Kaup, B. (2006). Context effects when reading negative and affirmative sentences, in R. Sun (ed.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, pp. 1735–1740. Mahwah, NJ: Lawrence Erlbaum.

Luka, B. J., and Barsalou, L. W. (2005). Structural facilitation: Mere exposure effects for grammatical acceptability as evidence for syntactic priming in comprehension. *Journal of Memory and Language* 52: 436–459.

Lutz, U. (2001). *Studien zu Extraktion und Projektion im Deutschen*. Doctoral dissertation, Universität Tübingen.

MacDonald, M. C., Pearlmutter, N. J., and Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review* 101: 676–703.

Maclay, H., and Sleator, M. D. (1960). Responses to language: Judgements of grammaticalness. *Journal of American Linguistics* 26: 275–282.

MacWhinney, B. (1977). Starting points. *Language* 53: 152–68.

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mahajan, A. (1990). *The A/A-bar distinction and movement theory*. PhD dissertation, MIT.

Malhotra, S. (2009). On *wh*-quantifier interactions. Unpublished MS, University of Maryland.

Maling, J. (1978). An asymmetry with respect to wh-islands. *Linguistic Inquiry* 9(1): 75–89.

Maling, J., and Zaenen, A. (1982). A phrase structure account of Scandinavian extraction phenomena. In P. Jacobson and G. K. Pullum (eds.), *The Nature of Syntactic Representation*, pp. 229–282. Dordrecht: Reidel.

Manning, C., and Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. Cambridge, MA: MIT Press.

Manzini, M. R. (1992). *Locality: A theory and some of its empirical consequences*. Cambridge, MA: MIT Press.

Marcus, M., Santorini, B., Marcinkiewicz, M., and Taylor, A. (1999). *Treebank-3*. Philadelphia: Linguistic Data Consortium.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. Cambridge, MA: MIT Press.

Martin, A., and McElree, B. (2008). A content-addressable pointer mechanism underlies comprehension of verb-phrase ellipsis. *Journal of Memory and Language* 58: 879–906.

Maxwell, S. E., and Delaney, H. D. (2003). *Designing Experiments and Analyzing Data: A model comparison perspective*. Mahwah, NJ: Lawrence Erlbaum Associates.

May, R. (1985). *Logical Form: Its structure and derivation*. Cambridge, MA: MIT Press.

Maye, J., Werker, J. F., and Gerken, L.-A. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82: B101–B111.

McCawley, J. D. (1981). The syntax and semantics of English relative clauses. *Lingua* 53: 99–149.

McCloskey, J. (1990). Resumptive pronouns, A′-binding and levels of representation in Irish. In R. Hendrick (ed.), *Syntax and Semantics* 23: *Syntax of the Modern Celtic languages* pp. 199–248. New York: Academic Press.

McCloskey, J. (2000). Quantifier float and *wh*-movement in an Irish English. *Linguistic Inquiry* 31: 57–84.

McCloskey, J. (2001). The morphosyntax of WH-extraction in Irish. *Journal of Linguistics* 37: 67–100.

McCloskey, J. (2002). Resumption, successive cyclicity, and the locality of operations. In S. D. Epstein and T. D. Seely (eds.), *Derivation and Explanation in the Minimalist Program*, pp. 184–226. Oxford: Blackwell.

McCloskey, J. (2006). Resumption. In Martin Everaert and Henk van Riemsdijk (eds.), *The Blackwell Companion to Syntax*, pp. 94–117. Oxford: Blackwell Publishing.

McDaniel, D., and Cowart, W. (1999). Experimental evidence for a minimalist account of English resumptive pronouns. *Cognition* 70: B15–B24.

McElree, B. (1996). Accessing short-term memory with semantic and phonological information, a time-course analysis. *Memory & Cognition* 24(2): 173–187.

McElree, B. (1998). Attended and nonattended states in working memory, accessing categorized structures. *Journal of Memory & Language* 38: 225–252.

McElree, B. (2000). Sentence comprehension is mediated by content-addressable memory. *Journal of Psycholinguistic Research* 29(2): 111–123.

McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory & Cognition* 27: 817–835.

McElree, B. (2006). Accessing Recent Events. *The Psychology of Learning and Motivation* 46(6): 155–200.

McElree, B., and Dosher, B. A. (1989). Serial position and set size in short-term memory: Time course of recognition. *Journal of Experimental Psychology: General* 118: 346–373.

McElree, B., Foraker, S., and Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language* 48: 67–91.

McElree, B., and Griffth, T. (1995). Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21: 134–157.

McElree, B., and Griffith, T. (1998). Structural and lexical constraints on filling gaps during sentence processing: A time-course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24: 432–460.

McKinnon, R., and Osterhout, L. (1996). Event-related potentials and sentence processing: Evidence for the status of constraints on movement phenomena. *Language and Cognitive Processes* 11: 495–523.

McKoon, G., and Ratcliff, R. (1994). Sentential context and on-line lexical decision tasks. *Journal of Experimental Psychology: Language, Memory and Cognition* 20: 1239–1243.

McMahon, L. E. (1963). *Grammatical analysis as part of understanding a sentence*. PhD thesis, Harvard University.

Mehler, J. (1963). Some effects of grammatical transformations on the recall of English sentences. *Journal of Verbal Learning and Verbal Behavior* 2(4): 346–351.

Merchant, J. (2001). *The Syntax of Silence*. Oxford University Press.

Merchant, J. (2009a). Phrasal and clausal comparatives in Greek and the abstractness of syntax. *Journal of Greek linguistics* 9: 134–164.

Merchant, J. (2009b). Islands and Lf-Movement in Greek Sluicing. *Journal of Greek Linguistics* 1: 39–62.

Michel, D. (2011). Individual differences inform the syntax-processing island debate. Talk at *Islands in Contemporary Linguistic Theory*. Vitoria-Gasteiz, University of the Basque Country.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63: 81–97.

Miller, G. A., and Chomsky, N. (1963). Finitary models of language users. In R. Luce, R. Bush, and E. Galanter (eds.), *Handbook of Mathematical Psychology*, Volume 2. New York: Wiley and Sons.

Milsark, G. L. (1976). *Existential sentences in English*. Doctoral dissertation, MIT.

Mintz, T. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition* 90: 91–117.

Mintz, T. (2006). Finding the verbs: Distributional cues to categories available to young learners. In K. Hirsh-Pasek and R.M. Golinkoff (eds.), *Action Meets Word: How children learn verbs*, pp. 31–63. New York: Oxford University Press.

Miyake, A., and Shah, P. (1999). *Models of Working Memory: Mechanisms of active maintenance and executive control*. New York: Cambridge University Press.

Miyamoto, E. T., and Nakamura, M. (2003). Subject/object asymmetries in the processing of relative clauses in Japanese. In G. Garding and M. Tsujimura (eds.), *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, pp. 342–355. Somerville, MA: Cascadilla Press.

Miyamoto, E. T., and Takahashi, S. (2004). Filler–gap dependencies in the processing of scrambling in Japanese. *Language and Linguistics* 5: 153–166.

Montalbetti, M. (1984). *After binding*. PhD dissertation, MIT.

Morgan, J. L. (1975). Some interactions of syntax and pragmatics. In P. Gole and J. L. Morgan (eds.), *Syntax and Semantics: Speech Acts*, Volume 3, pp. 289–303. New York: Academic Press.

Müller, G. (2010). On deriving CED effects from the PIC. *Linguistic Inquiry* 41(1): 35–82.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review* 89: 609–626.

Musolino, J., and Lidz, J. (2006). Why children aren't universally successful with quantification. *Linguistics* 44: 817–852.

Myers, J. (2006). An experiment in minimalist experimental syntax. Unpublished MS, National Chung Cheng University.

Nairne, J. (2002). The myth of the encoding-retrieval match. *Memory* 10(5–6): 389–95.

Nakano, Y., Felser, C., and Clahsen, H. (2002). Antecedent priming at trace positions in Japanese long-distance scrambling. *Journal of Psycholinguistic Research* 31: 531–571.

Neville, H. J., Nicol, J., Barss, A., Forster, K. I., and Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience* 3: 151–165.

Newmeyer, F. (2004). Against a parameter-setting approach to language variation. *Linguistic Variation Yearbook* 4: 181–234.

Nicol, J. L., Fodor, J. D., and Swinney, D. (1994). Using cross-modal lexical decision tasks to investigate sentence processing. *Journal of Experimental Psychology: Learning, Memory and Cognition* 20: 1229–38.

Nicol, J., and Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research* 18: 5–19.

Nieuwland, M., and Kuperberg, G. (2008). When the truth is not too hard to handle: An event-related potential study on the pragmatics of negation. *Psychological Science* 19: 1213–1218.

Nishigauchi, T. (1990). *Quantification in the Theory of Grammar*. Dordrecht: Kluwer.

Niyogi, P., and Berwick, R. (1996). A language learning model for finite parameter spaces. *Cognition* 61: 161–193.

Nordgaard, T. (1985). *Word order, binding, and the empty category principle*. Cand. Philol. thesis, University of Trondheim.

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory and Cognition* 28: 411–421.

Oberauer, K. (2006). Is the focus of attention in working memory expanded through practice? *Journal of Experimental Psychology: Learning, Memory and Cognition* 32: 197–214.

O'Bryan, E. (2003). *Event structure in language comprehension*. PhD dissertation, University of Arizona.

Oehrle, R. (1976). *The grammatical status of the English dative alternation*. PhD dissertation, MIT.

Omaki, A. (2010). *Commitment and flexibility in the developing parser*. PhD dissertation, University of Maryland.

Omaki, A., Davidson White, I., Goro, T., Lidz, J., and Phillips, C. (2012). No fear of commitment: Children's incremental interpretation in English and Japanese *wh*-questions. Unpublished MS.

Omaki, A., Lau, E., Davidson White, I., and Phillips, C. (2013). Hyper-active gap filling, pre-verbal object gap creation in English filler-gap dependency processing. Unpublished MS.

Omaki, A., and Nakao, C. (2010). Does English resumption really help to repair island violations? *Snippets* 21: 11–12.

Omaki, A., and Schulz, B. (2011). Filler-gap dependencies and island constraints in second language sentence processing. *Studies in Second Language Acquisition* 33: 563–588.

Öztekin, I., McElree, B., Staresina, B. P., and Davachi, L. (2008). Working memory retrieval: Contributions of left prefrontal cortex, left posterior parietal cortex and hippocampus. *Journal of Experimental Psychology: General* 138: 64–87.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin* 116: 220–244.

Pearl, L., Goldwater, S., and Steyvers, M. (2011). Online learning mechanisms for Bayesian models of word segmentation. *Research on Language and Computation*, special issue on computational models of language acquisition.

Pearl, L., and Lidz, J. (2009). When domain-general learning fails and when it succeeds: Identifying the contribution of domain-specificity. *Language Learning and Development* 5(4): 235–265.

Pearl, L., and Mis, B. (2011). How far can indirect evidence take us? Anaphoric One revisited. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, pp. 879–884. Boston, MA: Cognitive Science Society.

Pearl, L., and Mis, B. (2013). What indirect evidence can tell us about universal grammar: Anaphoric one revisited. *Unpublished MS*, University of California, Irvine.

Pearl, L., and Sprouse, J. (2013). Syntactic islands and learning biases: Combining experimental syntax and computational modeling to investigate the language acquisition problem. *Language Acquisition* 20: 23–68.

Pelucchi, B., Hay, J., and Saffran, J. (2009a). Statistical learning in natural language by 8-month-old infants. *Child Development* 80(3): 674–685.

Pelucchi, B., Hay, J., and Saffran, J. (2009b). Learning in reverse: Eight-month-old infants track backward transitional probabilities. *Cognition* 113(2): 244–247.

Perfors, A., Tenenbaum, J., and Regier, T. (2011). The learnability of abstract syntactic principles. *Cognition* 118: 306–338.

Perlmutter, D. (1971). *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rinehart, and Winston.

Pesetsky, D. (1982). *Paths and categories*. PhD dissertation, MIT.

Pesetsky, D. (1987). Wh-in-situ: movement and unselective binding. In E. Reuland and A. ter Meulen (eds.), *The Representation of (In)definiteness*, pp. 98–129. Cambridge, MA: MIT Press.

Pesetsky, D. (2000). *Phrasal Movement and its Kin*. Cambridge, MA: MIT Press.

Pesetsky, D., and Torrego, E. (2001). T-to-C movement: Causes and consequences. In M. Kenstowicz (ed.), *Ken Hale: A life in language*, pp. 355–426. Cambridge, MA: MIT Press.

Phillips, C. (1996). *Order and structure*. PhD dissertation, Massachusetts Institute of Technology.

Phillips, C. (2006). The real-time status of island phenomena. *Language* 82: 795–823.

Phillips, C. (2013). Some arguments and non-arguments for reductionist accounts of syntactic phenomena. *Language and Cognitive Processes* 28: 156–187.

Phillips, C., Kazanina, N., and Abada, S. (2005). ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research* 22: 407–428.

Phillips, C., and Lewis, S. (2013). Derivational order in syntax: Evidence and architectural consequences. *Studies in Linguistics* 6: 11–47.

Phillips, C., and Wagers, M. (2007). Relating structure and time in linguistics and psycholinguistics. In G. Gaskell (ed.), *Oxford Handbook of Psycholinguistics*, pp. 739–756. Oxford University Press.

Phillips, C., Wagers, M. W., and Lau, E. F. (2011). Grammatical illusions and selective fallibility in real-time language comprehension. In J. Runner (ed.), *Syntax and Semantics 37: Experiments at the Interfaces*, pp. 153–186. Bingley, UK: Emerald.

Pickering, M. J., Barton, S., and Shillcock, R. (1994). Unbounded dependencies, island constraints, and processing complexity. In C. Clifton, Jr., L. Frazier, and K. Rayner (eds.), *Perspectives on Sentence Processing*, pp. 199–224. London: Erlbaum.

Pickering, M. J., and Traxler, M. J. (2001). Strategies for processing unbounded dependencies, lexical information and verb-argument assignment. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27: 1401–10.

Pickering, M. J., and Traxler, M. J. (2003). Evidence against the use of subcategorisation frequency in the processing of unbounded dependencies. *Language and Cognitive Processes* 18: 469–503.

Pinheiro, J. C., and Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.

Platzack, C. (1999). Satsfläta med Relativsats. In I. Haskå and C. Sandqvist (eds.), *Alla tiders språk: En vänskrift till Gertrud Pettersson*, pp. 189–199. Lund: Institution för nordiska språk, Lunds universitet.

Polinsky, M. (1998). A non-syntactic account of some asymmetries in the double object construction. In J.-P. Koenig (ed.), *Conceptual Structure and Language: Bridging the gap*, pp. 403–422. Stanford: CSLI.

Polinsky, M., and Potsdam, E. (in press). Diagnosing covert A-movement. In L. Cheng and N. Corver (eds.), *Diagnostics in Syntax*.

Polinsky, M., Gomez-Gallo, C., Graff, P., and Kravtchenko, E. (2012). Subject preference and ergativity. *Lingua* 122(3): 267–277.

Pollack, I. (1953). The assimilation of sequentially encoded information. *American Journal of Psychology* 66: 421–435.

Pollard, C., and Sag, I. A. (1994). *Head-driven Phrase Structure Grammar*. University of Chicago Press.

Polyn, S. M., Norman, K. A., and Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review* 116 (1): 129–156.

Postal, P. M. (1994). Contrasting extraction types. *Journal of Linguistics* 30: 159–186.

Postal, P. (1998). *Three Investigations of Extraction*. Cambridge, MA: MIT Press.

Potsdam, E. and Polinsky, M. (2011). Against covert A-movement in Russian unaccusatives. *Linguistic Inquiry* 42: 345–355.

Poulsen, M. (2008). Acceptability and processing of long-distance dependencies in Danish. *Nordic Journal of Linguistics* 31: 73–107.

Prince, A., and Smolensky, P. (1993). *Optimality Theory: Constraint interaction in generative grammar*. Technical Report #2, Center for Cognitive Science, Rutgers University.

Prince, E. F. (1981). Towards a taxonomy of given-new information. In P. Cole (ed.), *Radical Pragmatics*, pp. 223–255. New York: Academic Press.

Prince, E. (1990). Syntax and discourse: A look at resumptive pronouns. In *Berkeley Linguistic Society: Proceedings of the 16th Annual Meeting*, pp. 482–497.

Prince, E. F. (1992). The ZPG letter: Subjects, definiteness, and information status. In S. A. Thompson and W. C. Mann (eds.), *Discourse Description: Diverse analyses of a fund raising text*, pp. 295–325. Amsterdam/Philadelphia: John Benjamins.

Prince, E. (1998). On the limits of syntax, with reference to Left-Dislocation and Topicalization. In P. Culicover and L. McNally (eds.), *Syntax and Semantics*, *29: The Limits of Syntax*, pp. 281–302. New York: Academic Press.

Pritchett, B. L. (1991a). Subjacency in a principle-based parser. In R. C. Berwick, S. P. Abney, and C. Tenny (eds.), *Principle-based Parsing: Computation and psycholinguistics*, pp. 301–345. Dordrecht: Kluwer.

Pritchett, B. L. (1991b). Head position and parsing ambiguity. *Journal of Psycholinguistic Research* 20(3): 251–270.

Pritchett, B. L. (1992). *Grammatical Competence and Parsing Performance*. The University of Chicago Press.

Pullum, G. K., and Scholz, B. (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review* 19: 9–50.

Radó, J. (1999). Some effects of discourse salience on gap-filling. Poster presented at the 12th Annual CUNY Conference on Human Sentence Processing.

Rafel, J. (2000). From complementizer to preposition. *Probus* 12: 67–92.

Raposo, E., and Uriagereka, J. (1990). Long-distance Case assignment. *Linguistic Inquiry* 21: 505–538.

Reali, F., and Christiansen, M. (2005). Uncovering the statistical richness of the stimulus: Structure dependence and indirect statistical evidence. *Cognitive Science* 29: 1007–1028.

Redington, M., Chater, N., and Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science* 22: 425–469.

Regier, T., and Gahl, S. (2004). Learning the unlearnable: The role of missing evidence. *Cognition* 93: 147–155.

Reinhart, T. (1997). Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20: 335–397.

Richards, N. (2001). *Movement in Language: Interactions and architectures*. Oxford University Press.

Rizzi, L. (1978). Violations of the wh-island condition in Italian and the subjacency condition. *Montreal Working Papers in Linguistics*.

Rizzi, L., ed. (1982a). *Issues in Italian Syntax*. Dordrecht: Foris.

Rizzi, L. (1982b). Violations of the wh-island constraint and the subjacency condition. In L. Rizzi (ed.), *Issues in Italian Syntax*, pp. 49–76. Dordrecht: Foris.

Rizzi, L. (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.

Rizzi, L. (1996). Residual verb second and the wh-criterion. In A. Belletti and L. Rizzi (eds.), *Parameters and Functional Heads: Essays in comparative syntax*, pp. 63–90. Oxford University Press.

Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar: Handbook of generative syntax*, pp. 281–337. Dordrecht: Kluwer.

Rizzi, L. (2003). Relativized minimality effects. In M. Baltin and C. Collins (eds.), *Handbook of Contemporary Syntactic Theory*, pp. 89–110. Oxford: Blackwell.

Rizzi, L. (2004). Locality and left periphery. In A. Belletti (ed.), *Structures and Beyond: The cartography of syntactic structures*, Volume 3, pp. 223–251. Oxford University Press.

Rizzi, L. (2006). On the form of chains: Criterial positions and ECP effects. In L. L.-S. Cheng and N. Corver (eds.), *Wh-movement: Moving on*, pp. 97–133. Cambridge, MA: MIT Press.

Rizzi, L. (2007). On some properties of criterial freezing. In V. Moscati (ed.), *CISCL Working Papers on Language and Cognition, 1: StiL Studies in Linguistics*, pp. 145–158. University of Siena.

Rizzi, L., and Shlonsky, U. (2007). Strategies of subject extraction. In H.-M. Gärtner and U. Sauerland (eds.), *Interfaces + Recursion = Language? Chomsky's minimalism and the view from syntax-semantics*, pp. 115–160. Berlin: Mouton de Gruyter.

Roberts, I., and Holmberg, A. (2010). Introduction: Parameters in minimalist theory. In T. Biberauer, A. Holmberg, I. Roberts, and M. Sheehan (eds.), *Parametric Variation: Null subjects in minimalist theory*, pp. 1–58. Cambridge University Press.

Roberts, R., and Gibson, E. (2002). Individual differences in sentence memory. *Journal of Psycholinguistic Research* 31: 573–598.

Rohde, D. (2003). Linger: a flexible platform for language processing experiments, version 2.94. http://tedlab.mit.edu/~dr/Linger/

Rooryck, J. (1992). Negative and factive islands revisited. *Journal of Linguistics* 28: 343–373.

Romero, M. (1998). *Focus and reconstruction effects in wh-phrases*. PhD Dissertation, University of Massachusetts at Amherst.

Romero, M. (2003). Correlate Restriction and Definiteness Effect in Ellipsis. In K. Schwabe and S. Winkler (eds.), *The Interfaces: Deriving and interpreting omitted structures*, pp. 263–300. Amsterdam: Benjamins.

Rosengren, I. (2002). EPP: A syntactic device in the service of semantics. *Studia Linguistica* 56(2): 145–190.

Ross, B. (2010). *The Psychology of Learning and Motivation: Advances in research and theory*. San Diego, CA: Academic Press.

Ross, J. R. (1967). *Constraints on Variables in Syntax*. Cambridge, MA: MIT dissertation.

Ross, J. R. (1969). Guess Who? In R. I. Binnick, A. Davison, G. M. Green, and J. L. Morgan (eds.), *Papers from the 5th Regional Meeting of the Chicago Linguistic Society*, pp. 252–286. Chicago Linguistic Society.

Ross, J. R. (1984). Inner islands, in *Proceedings of the 10th Annual Meeting of the Berkeley Linguistics Society*, pp. 258–265. University of California. Berkeley.

Ross, J. R. (1987). Islands and syntactic prototypes. In B. Need, E. Schiller, and A. Bosch (eds.), *Papers from the 23rd Annual Regional Meeting of the Chicago Linguistic Society: The General Session* I, pp. 309–320.

Ruchkin, D. S., Johnson, R., Canoune, H., and Ritter, W. (1990). Short-term memory storage and retention: An event-related brain potential study. *Electroencephalography and Clinical Neurophysiology* 76: 419–439.

Rudin, C. (1988). On multiple questions and multiple *wh-fronting. Natural Language and Linguistic Theory* 6: 445–501.

Rullmann, H. (1995). *Maximality in the semantics of WH-constructions*. PhD thesis, University of Massachusetts, Amherst.

Rumsfeld, D. H. (Presenter). (2002, February 12). *Department of Defense News Briefing [TV Transcript]. Washington, D.C., Department of Defense*. Retrieved from www.defense.gov/transcripts/transcript.aspx?transcriptid=2636.

Runner, J. T., Sussman, R., and Tanenhaus, M. (2003). Assignment of reference to reflexives and pronouns in picture noun phrases: evidence from eye movements. *Cognition* 81: 1–13.

Saah, K., and Goodluck, H. (1995). Island effects in parsing and grammar: Evidence from Akan. *Linguistic Review* 12: 381–409.

Saffran, J., Aslin, R., and Newport, E. (1996). Statistical learning by 8-month-old infants. *Science* 274: 1926–1928.

Saffran, J. R., Hauser, M., Seibel, R. L., Kapfhamer, J., Tsao, F., and Cushman, F. (2008). Grammatical pattern learning by infants and cotton-top tamarin monkeys. *Cognition* 107: 479–500.

Sag, I. A. (2010). English filler-gap constructions. *Language* 86: 486–545.

Sag, I. A., Hofmeister, P., and Snider, N. (2007). Processing complexity in Subjacency violations: The complex noun phrase constraint. In M. Elliott, J. Kirby, O. Sawada, E. Staraki, and S. Yoon (eds.), *Proceedings of the 43rd Annual Meeting of the Chicago Linguistic Society*, pp. 215–229. University of Chicago.

Sag, I. A., Wasow, T., and Bender, E. M. (2003). *Syntactic theory: A formal introduction*. Number 152 in CSLI Lecture Notes. CSLI Publications.

Sagae, K., Davis, E., Lavie, A., MacWhinney, B., and Wintner, S. (2010). Morphosyntactic annotation of CHILDES transcript. *Journal of Child Language* 37(3): 705–729.

Saito, M. (1985). *Some asymmetries in Japanese and their theoretical implications*. PhD dissertation, MIT.

Sakai, H. (1994). Complex NP constraint and case conversions in Japanese. In M. Nakamura (ed.), *Current Topics in English and Japanese*, pp. 179–203. Tokyo: Hituzi Syobo.

Sakas, W. G., and Fodor, J. D. (2001). The structural triggers learner. In S. Bertolo (ed.), *Language Acquisition and Learnability*, pp. 172–233. Cambridge University Press.

Salzmann, M. (2006). *Resumptive Prolepsis: A study in indirect A.′-dependencies*. Utrecht: LOT.

Sanford, A. J., and Sturt, P. (2002). Depth of processing in language comprehension: Not noticing the evidence. *Trends in Cognitive Science* 6(9): 382–386.

Schlesewsky, M., Fanselow, G., Kliegl, R., and Krems. J. (2000). The subject preference in the processing of locally ambiguous WH-questions in German. In B. Hemforth and L. Konieczny (eds.), *German Sentence Processing*, pp. 65–93. Dordrecht: Kluwer.

Schneider, D., and Phillips, C. (2001). Grammatical search and reanalysis. *Journal of Memory and Language* 45: 308–336.

Schulte im Walde, S. (1998). *Automatic semantic classification of verbs according to their alternation behavior*. Diplomarbeit, University of Stuttgart.

Schütze, C. (1996). *The Empirical Base of linguistics: Grammaticality judgments and linguistic methodology*. The University of Chicago Press.

Schwarz, F. (2007). Processing presupposed content. *Journal of Semantics* 24: 373–416.

Seidenberg, M. S., and MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science* 23: 569–588.

Sekerina, I. A. (2003). Scrambling and processing, dependencies, complexities and constraints. In S. Karimi (ed.), *Word Order and Scrambling*, pp. 301–324. Malden, MA: Blackwell Publishers.

Sells, P. (1984). *Syntax and semantics of resumptive pronouns*. Amherst, MA: University of Massachusetts dissertation.

Sells, P. (1987). Binding resumptive pronouns. *Linguistics and Philosophy* 10: 261–298.

Sharvit, Y. (1999). Resumptive pronouns in relative clauses. *Natural Language and Linguistic Theory* 17: 587–612.

Shieber, S. M., Pereira, F. C. N., and Dalrymple, M. (1999). Interaction of scope and ellipsis. In Shalom Lappin and Elabbas Benmamoun (eds.), *Fragments: Studies in Ellipsis and Gapping*, pp. 8–31. Oxford: Oxford University Press.

Shields, R. (2005). Russian adverbs and Relativized Minimality. In R. Shields (ed.), *LSO (Linguistics Student Organization) Working Papers in Linguistics 5: Proceedings of the WIGL (Workshop in Generative Linguistics) 2005*, pp. 152–167.

Shlonsky, U. (1992). Resumptive pronouns as last resort. *Linguistic Inquiry* 23: 443–468.

Slobin, D.I. (1966). Grammatical transformations and sentence comprehension in childhood and adulthood. *Journal of Verbal Learning and Verbal Behavior* 5(3): 219–227.

Smolensky, P., and Legendre, G. (2006). *The Harmonic Mind*. Cambridge, MA: MIT Press.

Snedeker, J., and Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology* 49: 238–299.

Snyder, W. (2000). An experimental investigation of syntactic satiation effects. *Linguistic Inquiry* 31: 575–582.

Sohn, H. (1980). Theme prominence in Korean. *Korean Linguistics* 2: 2–19.

Solan, Z., Horn, D., Ruppin, E., and Edelman, S. (2005). Unsupervised learning of natural languages. *Proceedings of the National Academy of Sciences* 102: 11629–11634.

Sorace, A. (2000). Gradients in auxiliary selection with intransitive verbs. *Language* 76: 859–890.

Sperber, D., and Wilson, D. (1986). *Relevance: Communication and cognition*. Cambridge, MA: Harvard University Press.

Sperber, D., and Wilson, D. (1995). *Relevance: Communication and cognition*. 2nd edn. Oxford: Blackwell.

Sportiche, D. (1981). Bounding nodes in French. *The Linguistic Review* 1: 219–246.

Sprouse, J. (2007a). *A program for experimental syntax: Finding the relationship between acceptability and grammatical knowledge*. PhD dissertation, University of Maryland.

Sprouse, J. (2007b). Continuous acceptability, categorical grammaticality, and experimental syntax. *Biolinguistics* 1: 118–129.

Sprouse, J. (2008). The differential sensitivity of acceptability judgments to processing effects. *Linguistic Inquiry* 39: 686–694.

Sprouse, J. (2009). Revisiting satiation: Evidence for a response equalization strategy. *Linguistic Inquiry* 40: 329–341.

Sprouse, J. (2011). A test of the cognitive assumptions of magnitude estimation: Commutativity does not hold for acceptability judgments. *Language* 87: 274–288.

Sprouse, J., and Almeida, D. (2012). Power in acceptability judgment experiments and the reliability of data in syntax. Unpublished MS, University of California Irvine and New York University Abu Dhabi.

Sprouse, J., and Almeida, D. (2013). The role of experimental syntax in an integrated cognitive science of language. In K. Grohmann and C. Boeckx (eds.), *The Cambridge Handbook of Biolinguistics*, pp. 181–202. Cambridge University Press.

Sprouse, J., Fukuda, S., Ono, H., and Kluender, R. (2011). Reverse island effects and the backward search for a licensor in multiple wh- questions. *Syntax* 14: 179–203.

Sprouse, J., Wagers, M., and Phillips, C. (2012). A test of the relation between working memory capacity and syntactic island effects. *Language* 88. 82–123

Staab, J. (2007). *Negation in context: Electrophysiological and behavioral investigations of negation effects in discourse processing*. PhD thesis, University of California, San Diego/San Diego State University.

Stabler, E. P. (1994). The finite connectivity of linguistic structures. In C. Clifton, Jr., L. Frazier, and K. Rayner (eds.), *Perspectives on Sentence Processing*, pp. 303–336. Hillsdale, NJ: Erlbaum.

Staub, A. (2007a). The parser doesn't ignore intransitivity, after all. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33: 550–569.

Staub, A. (2007b). The return of the repressed: Abandoned parses facilitate syntactic reanalysis. *Journal of Memory and Language* 57(2): 299–323.

Staum Casasanto, L., Hofmeister, P., and Sag, I. A. (2010). Understanding acceptability judgments: Distinguishing the effects of grammar and processing on acceptability judgments. In S. Ohlsson and R. Catrambone (eds.), *Proceedings of the 32nd*

Annual Conference of the Cognitive Science Society, pp. 224–229. Austin, TX: Cognitive Science Society.

Steedman, M. (2000). *The Syntactic Process*. Cambridge, MA: The MIT Press.

Stepanov, A. (2001a). *Cyclic domains in syntactic theory*. PhD dissertation, University of Connecticut.

Stepanov, A. (2001b). Late adjunction and Minimalist phrase structure. *Syntax* 4: 94–125.

Stepanov, A. (2007). The end of CED: Minimalism and extraction domains. *Syntax* 10: 80–126.

Sternefeld, W. (1985). Deutsch ohne grammatische Funktionen: ein Beitrag zur Rektions und Bindungstheorie. *Linguistische Berichte* 99: 394–439.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review* 64: 153–181.

Stevens, S. S. (1975). *Psychophysics: Introduction to its Perceptual, Neural, and Social Prospects*. New York: John Wiley.

Stevenson, B. (1999). Case study in information processing: Sentence processing. In J. Wiles and T. Dartnall (eds.), *Perspectives on Cognitive Science: Theories, experiments, and foundations*, pp. 277–292. Norwood, NJ: Ablex.

Stowe, L. A. (1986). Parsing WH-constructions: Evidence for on-line gap location. *Language and Cognitive Processes* 3: 227–245.

Stowe, L. A. (1992). The processing implementation of syntactic constraints: The sentence matching debate. In H. Goodluck and M. Rochemont (eds.), *Island Constraints: Theory, acquisition, and processing*, pp. 419–443. Dordrecht: Kluwer.

Stowe, L. A., Tanenhaus, M. K., and Carlson, G. M. (1991). Filling gaps on-line, Use of lexical and semantic information in sentence processing. *Language and Speech*, 34: 319–334.

Strawson, P. (1964). Identifying reference and truth values. *Theoria* 30: 96–118.

Stromswold, K., Caplan, D., Alpert, N., and Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain and Language* 52: 452–473.

Sturgeon, A., Polinsky, M., Gallo, C. G., Kravtchenko, E., Medova, L., and Koula, V. (2010). Subject islands in Czech. Paper presented at FASL-XIX, University of Maryland, May 2010.

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language* 48: 542–562.

Sturt, P., Pickering, M. J., Scheepers, C., and Crocker, M. W. (2001). The preservation of structure in language comprehension: Is reanalysis the last resort? *Journal of Memory and Language* 45(2): 283–307.

Suppes, P. 1974. The semantics of children's language. *American Psychologist* 29: 103–114.

Sussman, R. S., and Sedivy, J. C. (2003). The time-course of processing syntactic dependencies: Evidence from eye-movements. *Language and Cognitive Processes* 18: 143–163.

Swinney, D., Ford, M., Frauenfelder, U., and Bresnan, J. (1988). On the temporal course of gap-filling and antecedent assignment (during sentence comprehension). In B. Grosz, R. Kaplan, M. Macken, and I. Sag (eds.), *Language and Structure Processing*, pp. 12–24. Stanford, CA: CSLI.

Szabolcsi, A. (2006). Strong vs. weak islands. In M. Everaet, H. van Riemsdijk, R. Goedemans, and B. Hollebrandse (eds.), *The Blackwell Companion to Syntax*, Volume IV, pp. 479–531. Oxford: Blackwell.

Szabolcsi, A., and den Dikken, M. (1999). Islands. *GLOT International* 4: 3–8.

Szabolcsi, A., and den Dikken, M. (2003). Islands. In L. Cheng and R. Sybesma (eds.), *The Second Glot International State-of-the Article Book*, pp. 213–240. Berlin: Mouton de Gruyter.

Szabolcsi, A., and Zwarts, F. (1993). Weak islands and an algebraic semantics of scope taking. *Natural Language Semantics* 1: 235–284.

Takahashi, D. (1994). *Minimality of movement*. PhD Dissertation, University of Connecticut.

Takami, K.-I. (1989). Preposition stranding: Arguments against syntactic analyses and an alternative functional explanation. *Lingua* 76: 299–335.

Tanenhaus, M. K., and Trueswell, J. C. (1995). Sentence comprehension. In J. L. Miller and P. D. Eimas (eds.), *Handbook of Perception and Cognition: Speech, language, and communication*, pp. 217–262. San Diego, CA: Academic Press.

Tanenhaus, M. K., Stowe, L. A., and Carlson, G. N. (1985). The interaction of lexical expectation and pragmatics in parsing filler-gap constructions. *Proceedings of the Seventh Annual Cognitive Science Society Meetings*, pp. 361–365.

Taraldsen, K. T. (1982). Extraction from relative clauses in Norwegian. In E. Engdahl and E. Ejerhed (eds.), *Readings on Unbounded Dependencies in Scandinavian Languages*, pp. 205–221. Stockholm: Almqvist & Wiksell.

Tenenbaum, J., and Griffiths, T. (2001). Generalization, similarity and Bayesian inference. *Behavioral and Brain Sciences* 24: 629–640.

Testelets, J. (2006). Review of S. Karimi, ed. On scrambling. *Voprosy jazykoznanija* 4: 121–128.

Thompson, S. A. (1990). Information flow and 'dative shift' in English. In J. Edmondson, K. Feagin, and P. Mühlhäusler (eds.), *Development and Diversity: Linguistic variation across time and space*, pp. 239–253. Dallas: Summer Institute of Linguistics.

Tokimoto, S. (2009). Island phenomenon in Japanese and working memory: Syntactic constraints independent from working memory constraints. Poster presented at the 22nd annual CUNY Sentence Processing Conference, Davis, CA.

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

Trabasso, T., Rollins, H., and Shaughnessy, E. (1971). Storage and verification stages in processing concepts. *Cognitive Psychology* 2: 239–289.

Traxler, M. J., Morris, R. K., and Seely, R. E. (2002). Processing subject and object relative clauses: evidence from eye movements. *Journal of Memory and Language* 47: 69–90.

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

Traxler, M. J., Williams, R. S., Blozis, S. A., and Morris, R. K. (2005). Working memory, animacy, and verb class in the processing of relative clauses. *Journal of Memory and Language* 53: 204–224.

Tribus, M.(1961). *Thermodynamics and Thermostatics: An Introduction to Energy, Information and States of Matter, with Engineering Applications*. New York: D. Van Nostrand Company Inc.

Trueswell, J. C., Tanenhaus, M. K., and Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19: 528.

Trueswell, J. C., Sekerina, I., Hill, N. M., and Logrip, M. L. (1999). The kindergarten path effect: Studying on-line sentence processing in young children. *Cognition* 73: 89–134.

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

Truswell, R. (2011). *Events, Phrases, and Questions*. Oxford University Press.

Tsai, W.-T. (1994). On nominal islands and LF extraction in Chinese. *Natural Language and Linguistic Theory* 12: 121–175.

Tsai, W.-T. D. (1997). On the absence of island effects. *Tsing Hua Journal of Chinese Studies* 27: 125–149.

Tsimpli, I.-M. (1999). Null operators, clitics and identification: A comparison between Greek and English. In A. Alexiadou, G. Horrocks, and M. Stavrou (eds.), *Studies in Greek Syntax*, Volume 43 of *Studies in Natural Language and Linguistic Theory*, pp. 241–262. Dordrecht: Kluwer Academic Publishers.

Tullving, E. (1983). *Elements of Episodic Memory*. Oxford University Press.

Tulving, E., and Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review* 80: 352–373.

Ueno, M., and Garnsey, S. M. (2008). An ERP study of the processing of subject and object relative clauses in Japanese. *Language and Cognitive Processes* 23: 646–688.

Uriagereka, J. (1999). Multiple Spell-out. In S. Epstein and N. Hornstein (eds.), *Working Minimalism*, pp. 251–282. Cambridge, MA: MIT Press.

Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition* 40: 21–81.

Vallabha, G. K., McClelland, J. L., Pons, F., Werker, J. F., and Amano, S. (2007). Unsupervised learning of vowel categories from infant-directed speech. *Proceedings of the National Academy of Sciences* 104: 13273–13278.

Vallduví, E. (1993). *Information Packaging: A survey*. HCRC/RP 44. Edinburgh: Human Communication Research Center (HCRC).

Van Dyke, J. A. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33(2): 407–30.

Van Dyke, J. A., and Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A retrieval interference theory of recovery from misanalyzed ambiguities. *Journal of Memory and Language* 49(3): 285–316.

Van Dyke, J. A. and McElree, B. (2006). Retrieval interference in sentence processing. *Journal of Memory and Language* 55(2): 157–166.

van Gompel, R., and Liversedge, S. (2003). The influence of morphological information on cataphoric pronoun assignment. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29: 128–139.

van Gompel, R. P. G. V., Pickering, M. J., and Traxler, M. J. (2001). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language* 45: 225–258.

Van Hoek, K. (1995). Conceptual reference points: A Cognitive Grammar account of pronominal anaphora constraints. *Language* 71: 310–340.

Van Valin, R. D. Jr. (1998). The acquisition of wh-questions and the mechanisms of language acquisition. In M. Tomasello (ed.), *The New Psychology of Language: Cognitive and functional approaches to language structure*, pp. 221–249. Hillsdale, NJ: Lawrence Erlbaum.

Vasishth, S., Brüssow, S., Lewis, R., and Drenhaus, H. (2008). Processing polarity: How the ungrammatical intrudes on the grammatical. *Cognitive Science* 32: 685–712.

Vasishth, S., and Lewis, R. L. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Sciences* 29: 375–419.

Vasishth, S., Lewis, R., Kern, S., and Suckow, K. (2010). Short-term forgetting in sentence comprehension: Crosslinguistic evidence from verb-final structures. *Language and Cognitive Processes* 25(4): 533–567.

Verhagen, A. (2006). On subjectivity and "long distance Wh-movement." In A. Athanasiadou, C. Canakis, and B. Cornillie, (eds.), *Subjectification: Various paths to subjectivity*, pp. 323–346. New York: Mouton de Gruyter.

Wagers, M. (2008). *The structure of memory meets memory for structure in linguistic cognition*. Doctoral dissertation, University of Maryland.

Wagers, M. W., Lau, E. F., and Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language* 61: 206–237.

Wagers, M., and McElree, B. (2011). Memory for linguistic features, evidence from the dynamics of agreement. Unpublished MS, University of California Santa Cruz and New York University.

Wagers, M. W., and Phillips, C. (2009). Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics* 45: 395–433.

Wagers, M. W., and Phillips, C. (2012). Going the distance: memory and decision making in active dependency construction. Unpublished MS, University of California Santa Cruz and University of Maryland.

Wang, H., and Mintz, T. (2008). A dynamic learning model for categorizing words using frames. In H. Chan, H. Jacob, and E. Kapia (eds.), *BUCLD 32 Proceedings*, pp. 525–536. Somerville, MA: Cascadilla Press.

Wanner, E., and Maratsos, M. (1978). An ATN approach to comprehension. In M. Halle, J. Bresnan, and G. A. Miller (eds.), *Linguistic Theory and Psychological Reality*, pp. 119–161. Cambridge, MA: MIT Press.

Warren, T., and Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition* 85: 79–112.

Wason, P. C. (1961). Response to affirmative and negative binary statements. *British Journal of Psychology* 52: 133–142.

Wason, P. C. (1965). The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavior* 4: 7–11.

Wasow, T. (1972). *Anaphoric Relations in English*. Cambridge, MA: MIT dissertation.

Wasow, T. (2002). *Postverbal Behavior*: Stanford, CA: Center for Study of Language and Information.

Wasow, T., and Arnold, J. (2005). *Intuitions in linguistic argumentation*. *Lingua* 115: 1481–1496.

Webelhuth, G. (1989). *Syntactic Saturation Phenomena and the Modern Germanic Languages*. University of Massachusetts, GLSA.

Weinberg, A. (1992). Parameters in the theory of sentence processing: Minimal commitment theory goes east. *Journal of Psycholinguistic Research* 22(3): 339–364.

Wellwood, A., Pancheva, R., Hacquard, V., and Phillips, C. (unpublished). Deconstructing a comparative illusion. Unpublished MS, University of Maryland and University of Southern California.

Weskott, T., and Fanselow, G. (2011). On the informativity of different measures of linguistic acceptability. *Language* 87: 249–273.

Wexler, K., and Culicover, P. (1981). *Formal Principles of Language Acquisition*. Cambridge, MA: MIT Press.

Wickelgren, W. A. (1976). Network strength theory of storage and retrieval dynamics. *Psychological Review* 83: 466–478.

Wilcox, R. (1997). *Introduction to Robust Estimation and Hypothesis Testing*. New York: Academic Press.

Wurmbrand, S. (2004). No TP-fronting meets Nearly Headless Nick. Unpublished MS, University of Connecticut.

Wurmbrand, S. (2006). Licensing Case. *Journal of Germanic Linguistics* 18(3): 175–234.

Xiang, M., Dillon, B. W., and Phillips, C. (2009). Illusory licensing effects across dependency types: ERP evidence. *Brain and Language* 108: 40–55.

Xiang, M., Heestand, D., and Polinsky, M. (2008). Resumptive pronouns: (non)-rescuing function? Poster presented at the 21st Annual Meeting of the CUNY conference on Human Sentence Processing. University of North Carolina, Chapel Hill.

Xu, F., and Tenenbaum, J. (2007). Word learning as Bayesian inference. *Psychological Review* 114: 245–272.

Yang, C. (2002). *Knowledge and Learning in Natural Language*. Oxford University Press.

Yngve, V. (1961). The depth hypothesis. In R. Jakobson (ed.), *Structure of Language and its Mathematic Aspects*, pp. 130–138. Providence, RI: American Mathematical Society.

Yoshida, M. (2006). *Constraints and mechanisms in long-distance dependency formation*. PhD dissertation, University of Maryland.

Yoshida, M., Aoshima, S., and Phillips, C. (2004). Relative clause prediction in Japanese. Talk at the 17th Annual CUNY Conference on Human Sentence Processing, University of Maryland.

Yoshida, M., Dickey, M. W., and Sturt, P. (in press a). Predictive processing of syntactic structure: sluicing and ellipsis in real-time sentence processing. *Language and Cognitive Processes*. DOI: 10.1080/01690965.2011.622905

Yoshida, M., Kazanina, N., Pablos, L., and Sturt, P. (in press b). On the origin of islands. *Language and Cognitive Processes*.

Zaksas, D., and Pasternak, T. (2006). Directional signals in the prefrontal cortex and in area MT during a working memory for visual motion task. *Journal of Neuroscience* 26(45): 11726–11742.

Zukowski, A., and Larsen, J. (2004). The production of sentences that we fill their gaps. Poster presented at the CUNY Sentence Processing Conference, University of Maryland.

Zukowski, A., and Larsen, J. (2011). *Wanna* contraction in children: Retesting and revising the developmental facts. *Language Acquisition* 18: 211–241.