Title

Subject encodings & retrieval interference

Author

Arnett, Nathan Vincent

Publication Date

2016

University of California
Santa Cruz

Subject encodings & retrieval interference

A dissertation submitted in partial satisfaction
of the requirements for the degree of

doctor of philosophy

in

Linguistics

by

Nathan V. Arnett

December 2016

Professor Matthew Wagers

Professor Sandra Chung

Professor Adrian Brasoveanu

Tyrus Miller, Vice Provost & Dean of Graduate Studies

Abstract

Title of dissertation:	Subject Encodings & Retrieval Interference
	Nathan Vincent Arnett Doctor of Philosophy, 2016
Dissertation directed by:	Professor Matthew Wagers Department of Linguistics

This dissertation addresses the role of memory processes in real-time language comprehension. A rapidly growing body of research indicates that the memory access required for incremental sentence comprehension utilizes a content-addressable architecture that gives rise to *similarity-based interference* effects, thereby unifying psycholinguistic research with independently-motivated principles of memory and cognitive models (McElree, Foraker, & Dyer, 2003; Van Dyke & Lewis, 2003). I investigate the nature of the information encoded in linguistic representations, using interference effects to diagnose the properties that are used to retrieve constituents from memory. Empirically, the focus is on the retrieval of subject constituents in two situations: a subject separated from its verb by an intervening relative clause, and so-called control dependencies (*Mary promised John to leave*) where the interpretation of an unexpressed infinitive subject depends on a preceding overt subject argument. Experimental manipulations vary the similarity between the target subject and a grammatically-illicit subject along various linguistic dimensions, and the morpho-syntactic complexity of these constituents. I tie the results to computational simulations implemented in the ACT-R theory of cognition (J. R. Anderson & Lebiere, 1998; Lewis & Vasishth, 2005), investigating different retrieval structures and targets. I conclude that subject retrieval is interference-prone, supporting content-addressable architectures, and that subject encodings are retrieved based on highly abstract syntactic properties. The results also indicate that constituent complexity modulates processing difficulty at the retrieval site (Hofmeister, 2011), which I argue to be an elaboration effect.

For Jennifer Marie Ehrhart

My greatest teacher, closest friend, biggest role model

and mother

Acknowledgments

This work would not have been possible without the support of countless teachers, friends, and family.

Matt Wagers taught me almost everything I know about Psycholinguistics. He also taught me the value of a rigorous and empirical approach to teaching experiences (and the lab, of course). Every stage of this work has benefited from his skepticism, insights, and unflinching support. If the contents of this work are, in some way, an acknowledgment of my appreciation for his efforts, then I fear it is still both too short and unsatisfactory.

Sandra Chung was a mentor and role-model throughout my time at UCSC. One class with Sandy was enough to convince me to pursue a doctorate at UCSC, and I have benefited immensely from her stunning command of both fact and theory. She also should probably get credit as copy editor for this dissertation.

Jorge Hankamer, for showing me and so many others what great teaching is, without *telling* me. Jorge's pedagogy has been the genesis, inspiration, and aspiration of my own in ways too numerous to list—at least without the aid of a blackboard.

My fellow graduate students at UCSC, past and present, for their friendship, support, and inspiration. Matt Tucker, Ryan Bennet, and Boris Harizanov took me under their wing early on, showing me both friendship and what it is to be a great linguist. They set a high bar. My cohort-mates, Bern Samko and Oliver Northrup, in addition to being some of the smartest people I have known, were also friends and supportive colleagues. Thanks also to Karl Devries, Scarlett Clothier-Goldschmidt, and literally the entire UCSC linguistics community—who always made me feel welcome and taught me that we do our best work when we know that everyone in the room is our friend.

My friends at Cafe Pergolessi, whose tireless efforts and friendship created the space where much of this dissertation was written: Hiram Coffee, Max Weinreich, Matt Hintze, Dan Strong, Lucy Raisch, Lauren Wysham-Gullo, Celeste Deruisa, Tim Flegal, Hillary Weisert, Chesney Tyrcha, and Patrick Delaney.

My comrades and close friends, for always taking care of me and making sure I had fun outside of linguistics: Lisa Curran, Janina Larenas, Jeb Purucker, and Delio Vasquez. Thanks constantly reminding me, in both word and deed, that a better world is possible, and worth the effort.

My family, for always supporting me and inspiring me to greater efforts: Robert Arnett, Christopher Arnett, Anabelle Arnett, Hal Ehrhart, and Mike Stuart. Thanks also to those who showed me love and support every day, the good and (especially) the bad: Zach Rais-Norman for playing the role of a younger, wiser brother; Sam Henson, for being the most resolute and consistent of friends; Karen Duek and Jared Gampel, who were always willing to listen and offer support. Their friendship is worth more to me than any doctorate. Finally, Natasha Ross, for being my favorite person anywhere, and for keeping me grounded while still managing to remind me that dreams are important.

Contents

List of Figures

Introduction — 1

1.1 The role of memory in sentence processing

This dissertation is concerned with the role of memory in sentence comprehension. More specifically, the focus of this work is on the nature of linguistic information that is used to identify and reactivate constituent encodings in memory in the course of real-time dependency formation.

Understanding a sentence requires more than understanding the meaning of the words it consists of. Sentence comprehension requires resolving grammatical dependencies that link words or phrases to other constituents, such as understanding that the word *cat* is the object of the verbs *fed* (1) and *love* (2).

(1) The cat seems to have been fed ___ .

(2) This is the cat with one eye that all my neighbors love ___ .

Linguistic dependencies present a challenge to the sentence processing system for three reasons. First, as shown above in (1 – 2), these dependencies are routinely *non-adjacent*, in that dependents may be separated by potentially quite large amounts of intervening material. Thus, the possibility of non-adjacent dependencies means that sentence comprehension requires a means of encoding and storing linguistic material while intervening material is processed, so that it may be remembered later when the dependency is resolved.

The second challenge for dependency formation is that the capacity of memory is finite. Only a limited amount of information can be maintained in a state available for immediate processing. While this point is perhaps uncontroversial, at least for human information processing systems, it is also supported by memory research. For instance, the seminal work of George A. Miller (1957) found that the number of items that people can hold in (short-term) memory is 7±2. For the present purposes, setting aside how 'items' correspond to linguistic objects, capacity limitations on memory reinforce the requirement that language processing requires a means to encode, store, and retrieve linguistic material. Were it otherwise, it wouldn't be hard to find the longest possible sentence.

> *How are linguistic objects encoded, and how is this information re-accessed later in the resolution of grammatical dependencies?*

The third challenge is that linguistic dependencies are grammatically constrained. Not just any word or phrase will do. Crucially, these constraints are stated in terms of hierarchical structure – containment relations – implicating the creation of a structured representation in the course of comprehension. Thus, for instance, the morphology on the verb *be* in (3) co-varies with the person and number features of a nominal expression that is structurally adjacent to the verb, but not the linearly adjacent nominal, *bushes* (brackets indicate containment relations).

(3) *Agreement:*
[The cat [in the bushes]] *is/*are* hunting birds.

(4) *Binding:*
Maggie said that Ben often cooks for *himself/*herself*.

Chapter 3 focuses on subject-verb dependencies beyond agreement. Hierarchical structure also constrains dependencies that must be 'local', in some sense. The sentence in (4) illustrates one such case: English reflexive anaphors like *himself* or *herself* co-refer with another inter-sentential element, their antecedent, with which they must match in number and gender. Importantly, this antecedent must be local, in some sense. Here, the antecedent must be contained within the same clause.

Finally, dependencies may be both hierarchically and lexically constrained. Anaphoric Control dependencies (5) triggered by a verb are constrained to hold between an implicit subject ("___") and an argument of the higher clause. In this sense, they are similar to reflexive anaphors.

(5) *Control:*

a. The captain promised the tug driver ___ to cast-off quickly.

b. The captain persuaded the tug driver ___ to cast-off quickly.

However, the choice of antecedent is further constrained by the lexical semantics of the control verb: *promise* requires the antecedent to be the superordinate subject, while *persuade* requires an object antecedent. Still other verbs, such as *beg*, allow either subject or object antecedents. We consider the processing of control dependencies in Chapter 5.

The general conclusion here is that language comprehension requires memory. More specifically, it requires a means of encoding and storing (potentially non-adjacent) dependents, and retrieving them. At a general level, this dissertation is concerned with questions of what is encoded, and how it is retrieved. Minimally, memory must encode the hierarchical structure that informs grammatical constraints. Relatedly, the retrieval mechanism(s) must provide a means of re-activating constituent representations in a way that conforms to grammatical constraints, which requires a means of targeting sub-parts of arbitrary structures (Wagers, 2008).

> *In what way have memory access mechanisms adapted to support rapid, accurate incremental comprehension?*

I address the first question by examining the nature of the information used to identify and discriminate constituent encodings.

1.2 The challenge of interference

Recently, several researchers have presented evidence that the memory access required for sentence comprehension is accomplished by a cue-directed, associative retrieval mechanism (McElree, 2000; McElree, Foraker, & Dyer, 2003; Van Dyke & Lewis, 2003). This mechanism is cue-based, in that information in the retrieval context – say a verb – provides 'cues' about the properties that the target encoding should contain, for instance a plural noun. It is also content-addressable, in that any and all encodings that match (a subset of) the retrieval cues are activated in parallel, without the need to incrementally search through the contents of memory. The associative property of this access means that retrieval is unaffected by the size of the memory space, and fast enough to plausibly support the timecourse of language processing (Lewis, Vasishth, & Van Dyke, 2006). However, it also means that memory access is prone to difficulty when the retrieval cues cannot unambiguously identify a single encoding (Wagers, E. F. Lau, & C. Phillips, 2009). For instance, many of us have had the experience of forgetting where we place our keys. It is not that we don't remember putting down the keys, but rather that we have many such memories, all of which are very similar to each other, making it difficult to discriminate the the most recent memory. This phenomenon, known as *similarity-based interference*, is a well established and quite general property of memory: items are more difficult to remember when they are similar to other elements in memory (M. C. Anderson & Neely, 1996).

The challenge of interference is even more severe in language processing, because linguistic representations are highly self-similar (Wagers & C. Phillips, 2009). Knowledge of language comprises a system of finite representational primitives, features, and a set of combinatoric principles that combine these objects into novel structures (Chomsky, 1957). Thus, for instance, the representation of a sentence will comprise a set of recursive objects built from linguistic features, making them highly self-similar.

1.2.1 We need a theory of retrieval cues

A central goal of this dissertation is to sharpen our understanding of the dimensions of (linguistic) similarity that engender interference. In what follows, we will review much of the evidence for content-addressable retrieval mechanisms in language processing. A key piece of this evidence consists of findings of interference effects in language processing, as predicted by content-addressable retrieval. However, these studies have rarely sought to directly address the nature of the retrieval cues that drive interference effects.

> *What are the dimensions of linguistic similarity that cause constituents to interfere?*

It is worth stressing a point made by Van Dyke & Lewis (2003), that findings of difficulty due to constituent similarity (interference effects) are largely independent of the precise character of retrieval cues—linguistic theory provides a host of potential sources of similarity, not all of which are mutually exclusive. Nonetheless, I believe that theories of content-addressable retrieval require a theory of retrieval cues for at least 3 reasons.

First, sentence processing models based on cue-driven retrieval take working memory constraints to be the primary determinant of difficulty. One outcome of this is that making predictions about comparative difficulty in such architectures requires a specification of the cues used to reactivate constituent encodings.

The second motivation concerns the way in which cue-based retrieval navigates syntactic constraints. Some dependencies are highly sensitive to interference from structurally-illicit constituents, in particular agreement (Wagers, E. F. Lau, & C. Phillips, 2009). On the other hand, other dependencies, such as reflexive anaphors, are either robust to interference effects (Sturt, 2003; B. W. Dillon, 2011), or relatively less interference-prone (Parker, 2014). This has led to questions about the *structure-sensitivity* of retrieval processes, the degree to which retrieval outcomes conform to syntactic constraints that are defined over the phrase-structure. The issue is important because phrase-structure information is relational, by definition: relations like 'subject' or 'object' refer to the position of a constituent in the broader syntactic context, and so cannot hold of an encoding in isolation. Content-addressable retrieval, though, is so-called because elements are retrieved based on their inherent properties ('item-level' information), not their relations to other constituents.

However, for any given dependency, a lack of interference effects cannot provide evidence that retrieval is structured by syntactic constraints, because it may simply be that cues are combined in such a way as to optimally diagnose the correct constituent, without making reference to the phrase structure. In other words, given just the right combination of retrieval cues, the outcomes of retrieval may appear to reference highly relational information, like the global syntactic representation, when the cues are specific enough.

The third reason to develop a theory of retrieval cues is that this is one of the areas where processing theories can inform grammatical theory. Assuming that the retrieval cues are informed by grammatical knowledge (Lewis & Vasishth, 2005), then we can use interference effects to diagnose the grammatical properties that characterize constituents. While holding other factors constant (i.e. the sentence is acceptable), we manipulate the similarity between the retrieval target and another, intervening constituent, probing for processing difficulty at the putative retrieval site. We further assume that difficulty at the retrieval site reflects interference, stemming from constituents sharing a retrieval cue. These cues, then, are the grammatical features of constituents (provided by the grammar), and thus evidence for a given cue can provide support for theories that take such information to be grammatically relevant, and/or ecidence against those theories that don't.

1.3 Why subject retrieval?

This dissertation addresses questions of what is encoded and how it is retrieved by examining the retrieval of subject constituents in a variety of syntactic configurations. The identification of a subject provides a good window into the nature of information used to diagnose constituent encodings for several reasons.

Theorists of language have long recognized that the subject is central to the meaning of a sentence (for a succinct review, see McCloskey, 1997). The subject is one of two primary relations in the logic of Aristotle, for instance. The centrality of subjecthood is largely undiminished in more contemporary theories, whether as a grammatical primitive in theories like LFG or HPSG (Bresnan, 1982; Pollard & Sag, 1994), or as constellation of derived properties in transformational theories (Chomsky, 1981).

For our purposes, subjects are an ideal testing ground because their grammatical properties are both well-understood and largely independent of each other. Thus, for instance, subjects are typically the most prominent syntactic and thematic argument, but may nonetheless bear a variety of semantic roles. These properties are cross-cut by other syntactic properties, such as the case they bear and whether or not they are the controller of agreement. This allows for experimental manipulation of these properties in a targeted way. To the extent that two subject-like constituents are similar in, say, their semantic roles, but distinct in syntactic position, then the presence of interference effects can be used to diagnose the retrieval cues used to identify the grammatically correct subject constituent.

1.4 Outline of the dissertation

This dissertation is organized as follows. We begin by reviewing existing evidence about the nature of memory, with a particular focus on the nature of memory access mechanisms (Chapter 2). The goal here is two-fold. First, to crystallize the domain-general properties of memory which constrain the language processing system. Second, to motivate the view that similarity matters for language processing. The nature of these retrieval mechanisms is the primary means by which the present work on sentence processing engages with research on memory, in particular, and cognition, more generally. Converging lines of research implicate two closely related properties of the system: very little information can be actively maintained for processing, and the rapid retrieval required to support incremental language comprehension proceeds via an associative, cue-driven mechanism. The picture that emerges is that the resolution of dependencies involves rapidly shunting constituent encodings between active and passive storage states, via retrieval, with the result that processing is more difficult when the grammatically-mandated dependent is similar to other constituents along some linguistic dimension. This sets the stage for subsequent chapters, where we focus on precisely what the relevant dimensions of similarity are.

In chapter 3, I review some evidence that increasing the complexity of a subject increases processing difficulty at its verb. This is hypothesized to be due to structural similarity. I show that the relevant dimensions of similarity are syntactic, not thematic, and that the syntactic cues are sufficient to distinguish between nominative subjects of a finite clause and the accusative subjects of a non-finite clause.

In chapter 4, we turn to the processing of empty categories, and in particular implicit or phonologically null subjects of non-root clauses. Empirically, the focus here is on the retrieval required to identify the implicit subject of a controlled infinitive, such as those in (5), above. Conceptually, the focus is on how the predictability of a missing element modulates interference effects, if at all. Two experiments are presented. The first experiment examines the effects of the complexity and syntactic position of intervening arguments on the retrieval of the antecedent of a controlled subject. We see that processing at the retrieval site is facilitated when intervening material contains a complex consituent that is the co-argument of a gap. This general facilitation contrasts with the variable effects of complexity in Chapter 3. A second experiment addresses whether retrieval of a controlled subject's antecedent resists interference due to predictive processes engaged by a lexical control verb. This hypothesis is consistent with the results of Experiment 5, which indicate that unpredictable control dependencies are more interference-prone than predictable control dependencies.

Chapter 5 presents the results of a series of computational simulations designed to evaluate different retrieval structures. In particular, this chapter addresses whether interference effects in the retrieval of a subject at a verb are due to similarity between the retrieval cues and multiple candidate subject encodings, or if difficulty is due to interference in retrieval of the clausal constituent in need of a subject. An additional focus of this chapter is the explanation of the complexity effects observed in previous chapters. I argue that increasing constituent complexity is a form of elaborative processing giving rise to a distinctiveness effect. Under this view, complex constituents require more processing, which helps to highlight differences among otherwise similar encodings, and making it easier to discriminate among retrieval candidates.

Chapter 6 concludes with a discussion of general findings, their relation to theories of sentences processing, and a look to future directions.

Memory access & interference | 2

2.1 Grounding sentence processing in memory

The resolution of grammatical dependencies—like the *head-dependent* relations of (say) a verb head seeking a dependent subject—requires memory of the recent past. Even subjective experience suggests that memory is finite, and there is only so much information that can be maintained for processing. Thus, language processing requires memory to encode and store linguistic constituents, then potentially retrieve them later, as needed. Now imagine we (as the sentence processor) must comprehend the sentence in (6).

(6) The cats that the neighbor plays with daily are happy.

Proceeding incrementally, roughly word-by-word, we arrive at the matrix verb, *are*. At this point, comprehension of the sentence depends on successfully re-activating the subject—we must remember the subject, *the cats*, and integrate it with the verb. This integration is necessary because the interpretation of a given sentence rests on a structured representation. Thus, we must correctly identify *the cats* as the experiencer of *happy*. Such a structured representation is also necessary for grammatical constraints, such as the those governing the agreement morphology on the auxiliary *are*. Adding difficulty to this task, note that not just any nominal constituent will suffice; the correct encoding must match in features like person/number and stand in an appropriate structural relation to the verb. In particular, it must be the head noun of a c-commanding sister to the verb. A central topic psycholinguistic research (including this dissertation) is how this remembering is accomplished.

2.2 Access mechanisms

Broadly speaking, probing memory to identify and re-activate constituent encodings like subjects might be accomplished in two ways. For one, as illustrated in Figure 2.1, a structured search

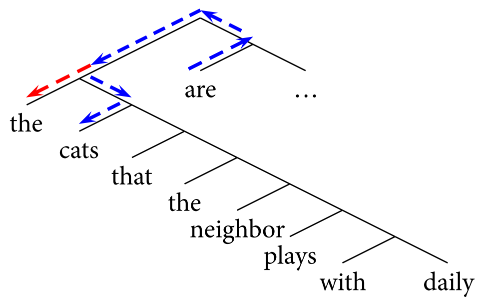


Figure 2.1: Schematized retrieval mechanisms, showing a structured search algorithm (left) and a content-addressable mechanism (right).

might be employed to interatively search the contents of memory until the target memory is found. For example, retrieval might progress backwards through the tree node-by-node, and at each maximal (XP) projection check to see if any of its daughters are morphosyntactically appropriate subjects. Search terminates when the correct daughter is identified. A structured search has a number of advantages. Among them is the fact that such a search algorithm can leverage grammatical knowledge encoded in the phrase-structure representation to make search very accurate.

Alternatively, a second approach would be an associative, content-addressable mechanism (right-hand side of Figure 2.1). A key property of this class of mechanisms is that information in the retrieval context (here, the verb) provides 'cues' about those properties that the target memory should contain, like keywords in a search engine. These cues are matched against the contents of memory at a global level, in parallel. Thus, in (6), the verbal morphology and phrase-structure constraints indicate that the target memory should have the properties [CASE:NOMINATIVE] and [NUM:PLURAL]. Since associative retrieval activates information in encodings directly, avoiding the need for an iterative search process, it has the advantage of being very fast. However, since the outcomes of cue-directed retrieval are any and all encodings matching (a subset of) the retrieval cues, it has the disadvantage of being error-prone. In general, performance of cue-directed retrieval is impaired when the cues used at retrieval cannot uniquely identify the target encoding, a property known as *similarity-based interference*. For instance, in (6), identifying the subject for the auxiliary verb *are* might be more difficult because the sentence contains two subject nouns, namely *the cats* and *the neighbor*.

Initial evidence seemed to favor a structured search mechanism. In a classic series of experiments, S. Sternberg (1966) and S. Sternberg (1969a) measured reaction times (RTs) of participants in a probe-recognition task. Participants studied lists of words or characters, after which they gave a Yes/No recognition judgement as to whether the probe matched an element of the studied list. Reaction times of the response were recorded. Sternberg manipulated the number of items that participants had to store in memory by varying the length of the studied list. Consistent with a structured search mechanism, Sternberg observed that reaction times on the probe-recognition task increased as a linear function of the number of to-be-remembered items. These results are shown in Figure 2.2. S. Sternberg (1975) reviewed this and other evidence, and concluded (somewhat tentatively) that the evidence favored a structured search mechanism for recognition, in particular a serial-comparison ("scanning") exhaustive search mechanism. Interestingly, S. Sternberg (1975) also considered a variant of direct-access mechanisms as an alternative to structured-search, but rejected it based on its inability to capture the linear relation between memory-set size and processing time.

Subsequent researchers, however, have challenged this view (see McElree, 2006, and referenced cited there). One issue is that Sternberg's conclusions were based on evidence from reaction times. The argument against reaction time evidence is based on two points. One, experimental participants can trade speed of processing for accuracy of retrieval. Comprehenders might process core argument relations like subject or object more deeply than non-essential elements like prepositional or relative clause modifiers. This relatively deeper processing could result in longer RTs without implicating a slowdown in retrieval processes. Two, while differences in retrieval speeds will be reflected as a difference in reaction times, a difference in reaction times might reflect distinctions in the underlying strength of a representation or, alternatively, the speed with which it is accessed. Thus, differences in reaction times do not provide evidence about retrieval speeds.

These differences track what Tulving & Pearlstone (1966) term the *availability* and *accessi-*

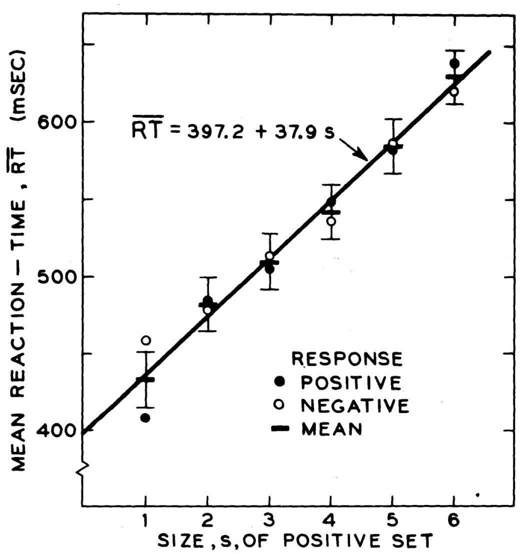


Figure 2.2: Reaction time results from S. Sternberg (1969b). Reaction times on a probe-recognition task (y-axis) were linearly related to the number of to-be-remembered items (x-axis) on a list.

bility of encodings. Availability refers to strength of the trace in memory – unavailable items are not present in memory – while accessibility refers to how easily they can be re-activated from memory. By way of illustration, imagine trying to identify an acquaintance in an old photograph. If the photograph is dirty, identification will be more difficult, but no information has been lost since we can simply wipe the dust away to restore the photograph. In other words, the dirt has reduced the accessibility of the photograph. In contrast, if the photo is sun-bleached and faded, no amount of wiping will help us—the availability of the required information has been reduced.

So how can we distinguish between structured and direct access mechanisms, if reaction times won't help us? What is needed is a way of measuring the speed at which an encoding can influence a process – its *retrieval dynamics* – independent of representational strength. The speed-accuracy trade-off (SAT) procedure (Wickelgren & Corbett, 1977) is one such method, and has provided strong evidence in favor of direct-access mechanisms. The SAT procedure is based on the paired-associate task of Reed (1973) and Reed (1976), and works as follows. Participants are shown the experimental stimulus, usually a symbol (for language experiments, a sentence), followed by a recognition probe. The task is to respond Y/N as to whether the symbol appeared on previously studied lists. In language experiments, the recognition probe is usually a binary acceptability judgement. At various times after the presentation of the probe, a response signal – generally a tone – is given to cue participants to respond Y/N to the probe. The delay between probe and response signal is designed to span the full time-course of retrieval, about 100–3000 ms after probe onset. Participants are required to respond within 100–300 ms of the signal.

By measuring accuracy across different delay periods, the SAT method plots accuracy as a function of processing time. The SAT procedure provides three dependent measures. The SAT *asymptote* (λ) of accuracy measures the long-run probability of retrieval, and provides an estimate of representational strength. The *intercept* is the point at which accuracy departs from chance (about 66%), and is taken to reflect when information initially becomes available for processing. The third dependent variable is the *rate* at which information grows from the intercept to asymptotic accuracy. Accuracy is calculated in terms of d', a measure of discriminability derived from the proportion of correct responses relative to the number of 'false alarms' (positive but incorrect responses). Accuracy at each response delay is fit using the function shown in (2.1), so that accuracy (d') at each time point, t, after the intercept, δ, is a function of asymptotic accuracy (λ) and the rate (β):

$$d'(t) = \lambda(1 - e^{-\beta(t-\delta)}), \text{for } t > \delta, \text{else } 0 \tag{2.1}$$

A set of hypothetical SAT curves is shown in Figure 2.3, taken from McElree (2006). Differences in retrieval speeds will arise due to a difference in either rate or intercept, so that they reach the crucial $1 - 1/e$ proportion of the asymptote at different times. Differences in retrieval dynamics indicate differences in (i) information accrual, if processing is continuous, or (ii) the distribution of finishing times, if processing is discrete (McElree, 2006). Relatedly, information in a state of privileged access will correlate with either a faster rate or an earlier intercept, independent of asymptotic differences. Crucially, as shown in Figure 2.3, the SAT distinguishes between processes that differ in accuracy but not speed (top of figure) from processes that differ in speed but not accuracy (bottom of figure).

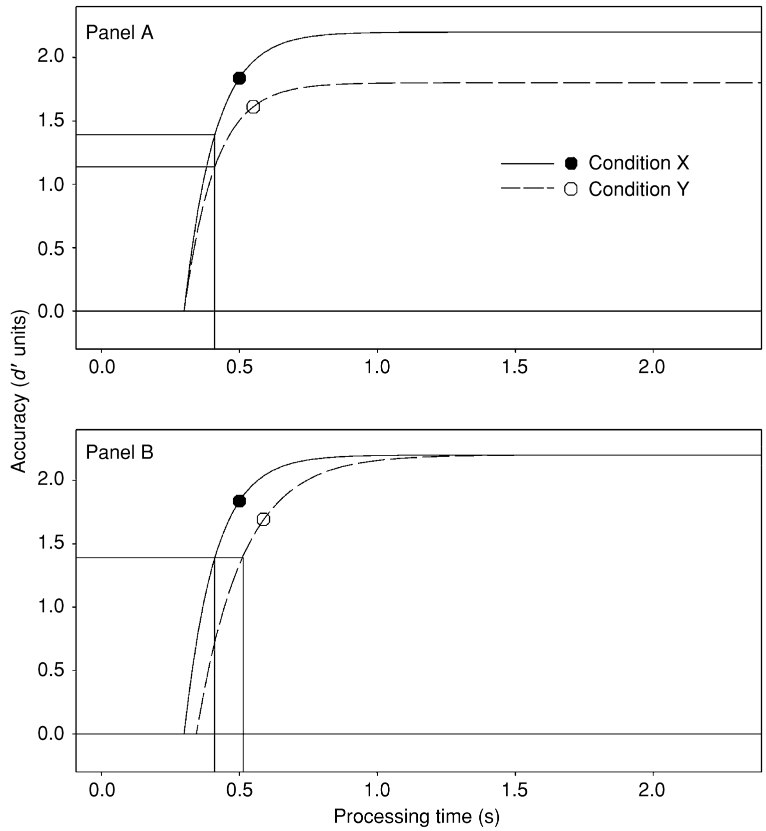


Figure 2.3: Hypothetical SAT functions, from McElree (2006). The top figure shows processes that differ in accuracy but not speed, while the bottom figure shows processes that differ in speed but not accuracy.

McElree & Dosher (1989) note that structured and direct-access mechanisms make distinct predictions about how the shape of the SAT function is effected by (i) memory set-size, and (ii) the serial position (recency) of the probe. If items in memory are accessed by a structured search mechanism, then additional to-be-remembered items will increase the search space, slowing retrieval. Relatedly, more recent items should be found more quickly, so that retrieval slows as the target memory becomes less recent. In contrast, content-addressable retrieval identifies memories via associative resonance between the retrieval cues and the inherent properties of the memories themselves, and so retrieval time remains constant irrespective of memory set-size. However, remembering additional items might decrease the *accuracy* of content-addressable retrieval, if increasing memory set-size also increases the similarity between the to-be-remembered items.

Finally, content-addressable retrieval should also be insensitive to the recency of the probe's associate, since items in the search space are unordered.

2.3 The subject is retrieved at the verb

McElree, Foraker, & Dyer (2003) used an SAT paradigm to examine the effects of material intervening between a subject and its verb on retrieval dynamics. They compared sentences where verb and subject were adjacent (7a), to non-adjacent conditions which increased both the amount of intervening material and the complexity of that material (7b–e). The length manipulation varied the number of intervening clauses, either one (7b–c) or two clauses (7d–e).

(7) *Materials from McElree, Foraker, & Dyer (2003)*

a. The book ripped.

b. The book that the editor admired ripped.

c. The book from the prestigious press that the editor admired ripped.

d. The book that the editor who quit the journal admired ripped.

e. The book that the editor who the receptionist married admired ripped.

The results of McElree, Foraker, & Dyer (2003) are shown in Figure 2.4. Asymptotic accuracy decreased as a function of both distance and complexity—adjacent conditions (7a) were most accurate, while the doubly-embedded object relative clause conditions ('ORC+ORC'; 7e) were least accurate, with accuracy decreasing roughly monotonically in conditions (b–d). Crucially, though, the retrieval dynamics indicated only a three-way distinction. Adjacent conditions were retrieved significantly faster than non-adjacent conditions, but there was no evidence of difference in retrieval speeds for the non-adjacent conditions, with one exception: the ORC+ORC conditions were significantly slower than the other non-adjacent conditions.

The findings of McElree, Foraker, & Dyer (2003) indicate that the distance between dependents matters for sentence processing. More specifically, the contrast in retrieval speeds for adjacent and non-adjacent conditions indicates that the most recent item is accessed much faster than non-recent items. In conjunction with other work (*e.g.* Cowan, 2005) the facilitation from recency is taken to motivate the view that our capacity for maintaining encodings in a state available for immediate processing – the focus of attention – is very limited, with encodings outside this state requiring retrieval. Thus, the facilitation for adjacent dependents indicates that retrieval is required in all but the closest of dependencies.

The fact that re-activation occurs in *constant-time* favors a direct-access mechanism, and is not consistent with a structured search mechanism. On the other hand, the lack of speed differences within the non-adjacent conditions (excepting the ORC+ORC structure, discussed below) indicates that the amount of material in the search space has no effect on retrieval speed, only accuracy. The reduction in accuracy is expected in a content-addressable architecture: more constituents potentially increase their similarity to each other, reducing the retrieval cues' ability to discriminate between candidate encodings. Indeed, the effect of similarity is a key prediction of content-addressable retrieval.

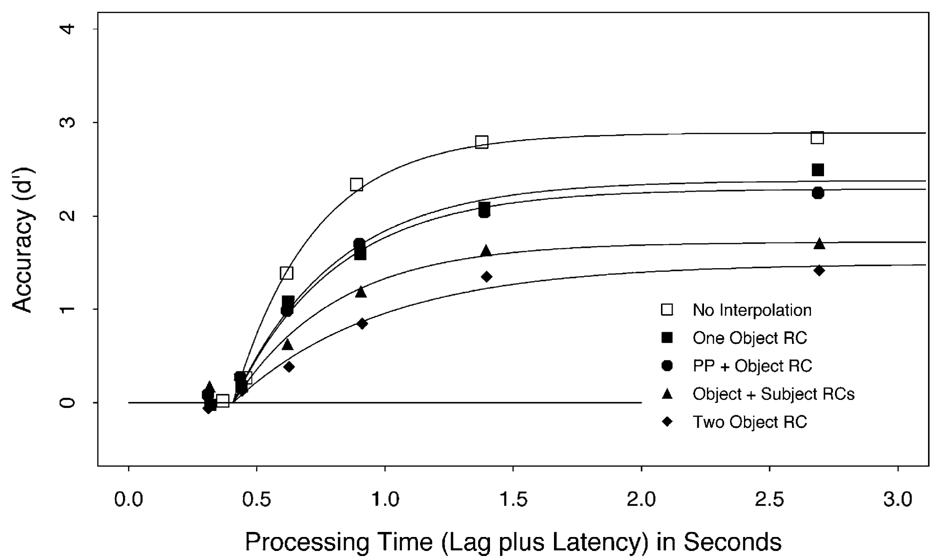


Figure 2.4: SAT results from McElree et al (2003), Experiment 2.

2.4 Similarity matters for language processing

So far, we have seen that one source of difficulty in language comprehension is non-adjacency. The results of McElree, Foraker, & Dyer (2003) indicate that non-adjacency matters because non-adjacent dependents must be retrieved for processing. More generally, studies of retrieval dynamics support the view that memory access is mediated by a content-addressable retrieval mechanism. While associative-match allows for very fast retrieval, it is prone to errors when the cues for retrieval fail to uniquely identify the target encoding, a property known as similarity-based interference. Thus, a key prediction of content-addressable memories is that retrieval is impaired when similarity between the elements in memory reduces the cues' ability to discriminate between potential dependents. In the remainder of this section, we turn to evidence that similarity between encodings matters for language processing, and that this similarity leads to errors in the processing of grammatical dependencies.

Recall that McElree, Foraker, & Dyer (2003) observed the greatest difficulty – lowest accuracy and slowest retrieval speeds – for the ORC+ORC condition (7), repeated as (8), below. Crucially, the ORC+ORC condition was more difficult than the nearly identical ORC+SRC condition, repeated below as (9).

(8) The book [that the editor [who the receptionist married ___] admired ___] ripped.

(9) The book [that the editor [who ___ quit the journal] admired ___] ripped.

This structure instantiates a case of double center-embedding by recursively embedding subject-attached relative clauses. Multiply embedded object-relatives are known to lead to comprehension difficulty (George A Miller & Isard, 1964; Blumenthal, 1966; Marks, 1968; Hakes & Cairns, 1970; Wang, 1970; Blaubergs & Braine, 1974). We return to this issue presently. However, McElree *et al*'s (2003) SRC+ORC condition is also center-embedded, so why should a second ORC, but not SRC, lead to difficulty?

It is a long-standing observation in psycholinguistic research that subject-extracted relatives are easier to process than object-extracted relatives (Hakes, Evans, & Brannon, 1976; Larkin &

Burns, 1977; Hudgins & Cullinan, 1978; Holmes & O'Regan, 1981; Ford, 1983; Jonathan King & Marcel Adam Just, 1991). Interestingly, Bever (1974) notes that multiply embedded object relatives become intuitively easier when the NPs are – in a sense to be determined – dissimilar from each other.

In a series of studies, Peter Gordon and colleagues hypothesized that linguistic memory, like recall/recognition memory more generally, is interference-prone. Gordon, Hendrick, & Marcus Johnson (2001), for instance, examined the effect of similarity on the processing of subject- and object-extracted relatives such as those shown in (10). Gordon *et al* manipulated similarity between NPs by comparing definite description DPs (10) to pronouns and names (11).

(10) a. The teacher that _ questioned *the student* wrote a long science fiction novel during the summer.

b. The teacher that *the student* questioned _ wrote a long science fiction novel during the summer.

(11) a. The teacher that _ questioned *you/Bob* wrote a long science fiction novel during the summer.

b. The teacher that *you/Bob* questioned _ wrote a long science fiction novel during the summer.

Using a self-paced reading methodology, Gordon, Hendrick, & Marcus Johnson (2001) replicated previous findings of a processing asymmetry for extraction type. Object-extracted relatives (10b) were more difficult than subject-extracted relatives (10a). This difficulty was reflected in longer RTs at both the embedded and matrix verbs. ORC sentences also showed lower comprehension accuracy than SRC sentences. More importantly, this difference was drastically reduced or eliminated when the NPs were of different types. When the sentence contained a descriptive NP and a pronoun or name (11), there were no differences between the embedded clause types at the matrix verb.

For our purposes, the important takeaway is that similarity matters for language processing—decreasing similarity between NPs facilitates processing. Indeed, Gordon, Hendrick, & Marcus

Johnson (2001) interpret their results as an interference effect, such that maintaining multiple NPs in memory is more difficult when they are similar to each other.

2.5 Subject retrieval is error-prone

Consistent with content-addressable retrieval, the results of the Gordon, Hendrick, & Marcus Johnson (2001), discussed above, indicate that similarity matters for language processing. We turn now to evidence that re-activation of constituent encodings at a verb is interference-prone, suggesting that dependency formation is also guided by a content-addressable retrieval mechanism. The argument here is that similarity impairs processing at a retrieval site, like a verb, when there are multiple potential dependents matching the cues used at retrieval. Crucially, these errors are driven by *grammatically inappropriate* constituents that would be ruled out if dependency resolution was fully faithful to grammatical constraints. The evidence comes from the processing of agreement relations.

In English the (person/number) agreement morphology of a verb co-varies with its subject—the head of the c-commanding sister DP to the verb. Under certain circumstances, though, the morphology of the verb is erroneously controlled by a grammatically inappropriate noun. For this reason, the effect is called 'agreement attraction':

(12) The key to the cells unsurprising *is* rusty from disuse.

(13) * The key to the cells unsurprisingly *are* rusty from disuse. attraction!

Agreement attraction was initially documented in production studies in English (Bock & C. A. Miller, 1991), but noted observationally much earlier (see Wagers, 2008, and references cited there). It has been replicated in several languages, including Dutch and German (Hartsuiker et al., 2003), French (Franck, Vigliocco, & Janet Nicol, 2002), Spanish (Antón-Méndez, J. L. Nicol, & Garrett, 2002), and Slovenian (Badecker & Kuminiak, 2007). Agreement attraction has also been documented in comprehension studies, where attraction configurations lead to lower processing difficulty (Pearlmutter, Garnsey, & Bock, 1999) or higher rates of acceptability (Clifton, Frazier,

& Deevy, 1999; Wagers, E. F. Lau, & C. Phillips, 2009). It is a well-attested and systematic error that has provided crucial evidence for content-addressable retrieval.

Agreement attraction shows two core asymmetries. First, it shows a markedness asymmetry: a singular subject can 'attract' the plural feature of a nearby noun, but not vice versa. Across production studies, prompts like *the keys to the cabinet...* lead to very few attraction errors (Eberhard, Cutting, & Bock, 2005). In comprehension, plural morphology on a verb can be incorrectly licensed by a grammatically illicit head noun, but there is no reliable attraction for singular attractors:

(14) * The key to the cells unsurprisingly *are* rusty from disuse. attraction!

(15) * The keys to the cell unsurprisingly *is* rusty from disuse. no attraction

The second asymmetry is that agreement attraction is observed only in ungrammatical sentences. In comprehension, agreement attraction can cause an otherwise ungrammatical sentence to be perceived as grammatical, what Wagers, E. F. Lau, & C. Phillips (2009) call an 'illusion of grammaticality'. But there is no corresponding 'illusion of ungrammaticality'; grammatical sentences are not reliably perceived as ungrammatical in the presence of a plural attractor (Wagers, 2008; Wagers, E. F. Lau, & C. Phillips, 2009):

(16) * The key to the cells unsurprisingly are rusty from disuse. attraction!

(17) The key to the cells unsurprisingly is rusty from disuse. no attraction

Initial accounts of agreement attraction were based on the idea that complex subjects contained multiple DPs bearing potentially conflicting number features, which led to the representation of the subject being in some sense 'defective'. This was due to either feature percolation from embedded DPs to the head noun (e.g. Franck, Vigliocco, & Janet Nicol, 2002; Eberhard, Cutting, & Bock, 2005, a.o.), or due to confusability of multiple number features in production processes (Bock & Cutting, 1992; Badecker & Lewis, 2007). The latter proposal gives no account of why agreement attraction should arise in comprehension. On the other hand, the feature percolation account cannot account for the grammaticality asymmetry, since there is no reason to expect that feature percolation should occur only in ungrammatical sentences.

Further evidence against the feature percolation account comes from the comprehension of sentences like (18). In (18b), attraction occurs between a verb inside an embedded relative clause and the RC head noun contained within a dominating clause. The distractor, *drivers*, is thus structurally higher and linearly farther from the verb than the target head noun, *runner* (Wagers, E. F. Lau, & C. Phillips, 2009; see also B. W. Dillon, 2011, B. Dillon et al., 2013).

(18) a. The driver [who the runners *wave* to each morning] always honk cheerfully.

b. * The drivers [who the runner *wave* to each morning] always honk cheerfully. attraction!

The attraction effect from the plural element cannot be due to linear proximity or local-coherence (Tabor, Galantucci, & Richardson, 2004, e.g.), since non-adjacent attractors can exert greater influence than adjacent attractors. For instance, when the head of the subject noun phrase is modified by two PPs, as in (19), speakers produce more attraction errors when the intermediate PP is plural (19a) than when the deepest – and closest – PP is plural (19b; Franck, Vigliocco, & Janet Nicol, 2002):

(19) a. * The threat to the *presidents* of the company...

b. * The threat to the president of the *companies*...

The patterns of attraction in (19) suggest that structural distance, and not linear distance, matters for agreement attraction (Wagers, E. F. Lau, & C. Phillips, 2009).

More recently, agreement attraction has been proposed to be an interference effect, driven by similarity between the distractor and the retrieval cues provided by the verb (Wagers, 2008; Wagers, E. F. Lau, & C. Phillips, 2009). Under this analysis, licensing the agreement features on the verb requires re-activating the subject to check its number features. This re-activation is accomplished via cue-driven retrieval. Assuming that the retrieval cues are provided by a combination of the verbal morphology and grammatical knowledge, then a plural verb gives the cues {NOUN, PLURAL, NOMINATIVE}. This retrieval is schematized below in (20–21).

(20) * the key… …the cabinets… …were…

$$\begin{bmatrix} \text{CAT} : \text{NP} \\ \text{CASE} : \text{NOM} \end{bmatrix} \quad \begin{bmatrix} \text{CAT} : \text{NP} \\ \text{NUM} : \text{PL} \\ \text{CASE} : \text{OBL} \end{bmatrix} \quad \begin{Bmatrix} \text{CAT} : ?\text{NP} \\ \text{NUM} : ?\text{PL} \\ \text{CASE} : ?\text{NOM} \end{Bmatrix}$$

(21) the key… …the cabinets… …is…

$$\begin{bmatrix} \text{CAT} : \text{NP} \\ \text{CASE} : \text{NOM} \end{bmatrix} \quad \begin{bmatrix} \text{CAT} : \text{NP} \\ \text{NUM} : \text{PL} \\ \text{CASE} : \text{OBL} \end{bmatrix} \quad \begin{Bmatrix} \text{CAT} : ?\text{NP} \\ \text{CASE} : ?\text{NOM} \end{Bmatrix}$$

In attraction configurations of the type in (20), neither the grammatical subject, *key*, nor the intervening noun, *cells*, matches all of the verb's retrieval cues. Both match in category, but while the grammatical subject matches the NOMINATIVE case feature, the number cue is matched only by the intervener. This gives rise to a *partial match* scenario, so that in some proportion of trials, retrieval returns the plural intervener, incorrectly licensing the plural morphology on the verb. On the other hand, in grammatical sentences of the type in (21), the distractor still matches the number and category cues, but the grammatical subject fully matches all of the retrieval cues. More generally, in grammatical sentences there will always be an encoding matching all of the retrieval cues, namely the subject, which mitigates the intrusion effect of the partially-matching distractor. Thus, the interference approach nicely captures the grammaticality asymmetry observed for agreement attraction.

The lack of attraction from singular distractors is accounted for by taking the number feature to be 'privative', meaning that only marked plural number is directly encoded as a cue, with singular number serving as a default (Wagers, E. F. Lau, & C. Phillips, 2009).

Let's retrace our steps. I have argued that language processing makes use of memory systems because linguistic dependencies are routinely non-adjacent, in that dependents must be stored and later retrieved. I presented some evidence that retrieval of linguistic material is accomplished via a cue-driven or content-addressable mechanism. Consistent with this view, similarity between nominal constituents modulates processing difficulty (Gordon, Hendrick, & Marcus Johnson, 2001). Furthermore, the processing of agreement dependencies indicates that this retrieval is prone to interference effects when grammatically-inappropriate constituents match the the retrieval cues provided by a verb (Wagers, E. F. Lau, & C. Phillips, 2009).

Subject encodings & retrieval interference[1]

3

3.1 Introduction

3.1.1 *What's at stake*

In Chapter 1, I discussed how the need to resolve non-adjacent dependencies places a number of demands on the sentence processor. For one, since dependents may be separated by potentially quite large amounts of intervening material, the sentence processor must be able to store partial representations of this structure while the intervening material is processed, and later re-activate their sub-parts at a point when the dependency is resolved. Second, linguistic representations need to encode the hierarchical containment relations that underlie some, if not all, constraints on dependency formation. These pressures present a significant challenge for the sentence processor: it must encode recursive structures, and have a way of re-accessing targeted sub-parts of those structures in a way that conforms to grammatical constraints.

In this chapter, I use behavioral reading measures, such as self-paced reading and eye-tracking, to address the issues of what is encoded and how it is retrieved by probing the nature of information used to retrieve a subject at its verb. Along the way, I review some evidence that integrating a subject and its verb is mediated by a content-addressable retrieval mechanism, where information in the retrieval context (*e.g.* a verb) is associatively matched against the contents of memory. Crucially, while this associative retrieval is very fast, the accuracy of retrieval is impaired when the retrieval cues fail to uniquely identify the target encoding, a property known as *similarity-based interference*. The argument for such a mechanism is based on evidence that attaching additional material to a subject leads to difficulty at a subsequent VP, and that a primary determinant of difficulty with these complex subjects is that the additional material increases similarity between potential subject dependents.

Assuming content-addressable retrieval, then, the narrow question addressed in this chapter is how subjects are characterized by the retrieval cues at a verb. The question is important because it directly addresses the theory of retrieval cues, and making predictions about interference in a cue-driven architecture requires precise formulations about the dimensions of similarity that lead constituents to interfere. Based on three reading studies, I aim to convince you that subjects are identified in a way that prioritizes their morphosyntactic properties.

3.1.2 *Overview of the chapter*

The structure of this chapter is as follows. In section 3.2, I summarize a collection of grammatical properties that are symptomatic of subjecthood. Among these properties are thematic properties, syntactic position, and case. However, we will see that none of these properties in isolation uniquely distinguish subjects from other arguments. In section 3.3, I turn to the processing of complex subjects, which have provided key evidence for a content-addressable retrieval mechanism. These studies indicate that increasing the complexity of a subject increases processing difficulty at its verb, and that this difficulty is more acute when the complex subject contains other subjects that are syntactically and/or semantically similar to the target subject. I discuss two issues with these studies. One, the materials confound the presence of an additional subject with the presence of an additional clause. Two, the dimensions of subjecthood that count as 'similar' are unclear, in that the materials do not distinguish between syntactic and thematic subjects. We address these issues in Experiment 1 (§3.4), where we compare similarity between the target subject and an intervening structural or thematic subject, while controlling for the number of clauses.

The results of Experiment 1 indicate that additional subjects do lead to difficulty, even when the depth of embedding is controlled, but only when they are syntactically similar to the target, and not thematically similar. In the remaining sections, I review some evidence implicating that similarity in terms of structural position and/or case properties leads to difficulty (§3.5). In two subsequent experiments (§3.6–3.7), I examine how similarity between a matrix and embedded subject in terms of either case or position modulates difficulty in the processing at the verb.

Here, the results are consistent with those of Experiment 1 in indicating that embedded subjects engender difficulty, but also indicate that the form of this difficulty depends on the degree of morphosyntactic similarity between the embedded subject and the target of retrieval. Section 3.9 concludes.

3.2 Dimensions of subjecthood

In the rest of this chapter, we will develop a theory of retrieval cues by examining the information used at a verb to identify and retrieve its subject. One reason for this is that subject retrieval has provided key evidence for interference effects in language processing. More generally, though, subject relations are at the core of meaning. Even Aristotle based his theory of meaning on the relation between a subject and its predicate. For our purposes, the retrieval of subjects is important because a subject may routinely be separated from its verb – by a relative clause modifier, for example – so that subjects must be retrieved when the verb is encountered. In the remainder of this dissertation, I will assume that this retrieval is content-addressable. The question then becomes, what are the retrieval cues for subjects?

Assuming that the retrieval cues are derived from grammatical knowledge (Lewis & Vasishth, 2005), in this section I sketch a collection of properties that are symptomatic of subjecthood. We will see that 'subject' elements can bear a variety of thematic roles, occur in several domains or syntactic positions, and are associated with different case properties. The broader point is that none of these properties in isolation are diagnostic of subjecthood, and thus would seem to be unreliable retrieval cues. Rather, the subjecthood of a given constituent depends on broader syntactic context. This sets the stage for subsequent sections, where we consider some evidence of interference effects driven by similarity in some, if not all, of these properties. Nonetheless, it is precisely the formal parallels and distinctions among kinds of subjects that make them ideal candidates for diagnosing the retrieval cues. This is because we can selectively manipulate similarity in one or more dimensions while controlling for other differences. This logic underlies the three experiments presented in this chapter.

3.2.1 *Thematic properties of subjects*

From the standpoint of cue-driven subject retrieval, one challenge is that 'subjects' can bear a variety of thematic roles. For instance, the sentences in (22) are all about the same event, namely the breaking of a window, and yet the role of the subject constituent in this event varies. In (22a) *Terry* is the agent of the breaking event, while in (22b) the subject corresponds to the INSTRUMENT of the breaking event, *the rock*. Finally, while in all three sentences *the window* is the THEME, the thematic-role generally associated with the internal argument, in (22c) the subject position is associated with the theme argument.

(22) a. Terry broke the window (with a rock). subj = agent

b. The rock broke the window. subj = cause

c. The window broke. subj = theme

The problem is even more acute when we consider so-called 'psych-predicates'. The standard minimal pair in this class of verbs is *fear* and *scare*, shown in (23). Despite the similarities in meaning between these sentences, the EXPERIENCER, *Robin*, is a subject in (23a), but an object in (23b).

(23) a. Robin fears this old house.

b. This old house scares Robin.

The point, by now, should be clear: the mapping between the argument/event structure of a predicate and the syntactic position of those arguments is not 1:1, at least without additional formal machinery.

Furthermore, even when the subject constituents bear the same thematic role, they need not occupy the same structural position. For example, if we compare a finite (tensed) clause (24a) with a so-called 'event' or 'process' nominalization (24b), the underlined elements are both the most prominent structural and thematic elements in their domain:

(24) a. The enemy destroyed the city.

b. The enemy's destruction of the city (was necessary to save it). (Chomsky, 1970)

More precisely, they occupy the highest structural position, from which they asymmetrically c-command other elements in their domain (Steven Paul Abney, 1987). These properties are summarized in Figure 3.1.

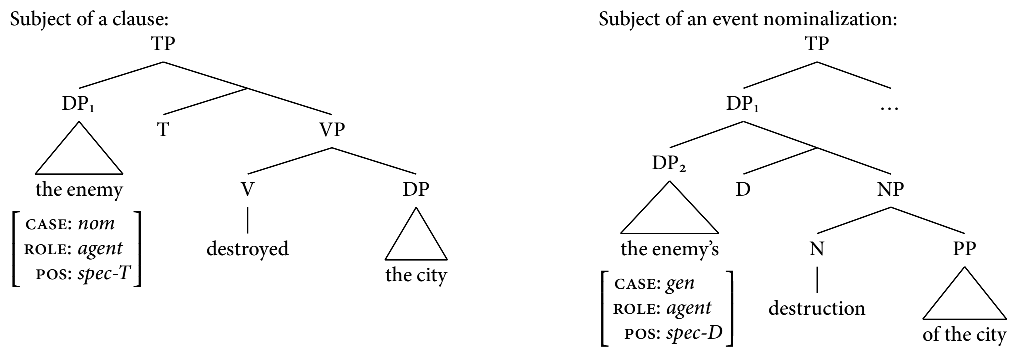


Figure 3.1: Syntactic structure of a clause (left) and an event nominalization (right). The subject-like elements are external arguments linked to the specifier position, but differ in their domain and case.

Furthermore, they are the elements linked to the external argument of the predicate (Grimshaw, 1990). Thus, for example, they asymmetrically bind anaphoric elements like reflexives (25–26), either because of their syntactic prominence (defined in terms of, e.g., c-command; Chomsky, 1981) or because they are the most prominent element of an argument structure list (e.g., o-command; Pollard & Sag, 1994).

(25) a. The enemy destroyed themselves.

b. * Themselves destroyed the enemy.

(26) a. The enemy's destruction of themselves.

b. * Themselves' destruction of the enemy.

The subject of a clause and the possessor of an event nominalization differ in both case and position. The subject of a clause bears NOMINATIVE case and occurs in SPEC-T, while the possessor in (24b) bears GENITIVE case and occurs in SPEC-D. Note, however, that both are the specifier of a functional head, and thus similar in that both are located in SPEC-X.

Another difference between these elements is their optionality. The subject of a (root English declarative) clause is obligatory (27), but the possessor of an event nominalization need not always occur (28b). The object of the nominalization also appears to be optional (28c). However, when the possessor appears without an object, the possessor is obligatorily interpreted as the object.

(27) a. The enemy destroyed the city.

b. * Destroyed the city.

(28) a. The enemy's destruction of the city (was well-documented).

b. The destruction (of the city) was well-documented.

c. The enemy's destruction (by the army) was well-documented.

Despite these differences, Grimshaw (1990) argues at length that event nominalizations nonetheless have argument structure, making a strong formal connection between clauses and complex nominal expressions. In particular, she argues that event nominalizations are ambiguous between forms with argument structure and those without. This ambiguity is the source of the apparent optionality of arguments in these constituents. I will not attempt to summarize all of Grimshaw's arguments here, but one piece of evidence will be significant in Experiment 1.

Grimshaw observes that when the nominalization includes an agent-oriented modifier like *deliberate* or *repeated*, the argument structure reading is obligatory and leads to ungrammaticality when the internal argument is omitted. Indeed, when the possessor is omitted, these constructions are marginal, at best, for most speakers:

(29) a. % The deliberate destruction of the city was well-documented.

b. * The deliberate destruction was well-documented.

While the modifier forces the presence of the internal argument, the possessor element is still optional, as shown in (29b). This is a stark contrast to the subject of a clause. Nonetheless, Grimshaw argues that the possessor is the element linked to the external argument of the predicate. The key piece of evidence for us comes from agent-oriented adjective/adverbs, again. Recall the possessor can occur without a surface object (28c), but must then be interpreted as the logical (underlying) object. However, a subject-oriented adverb in these configurations leads to ungrammaticality:

(30) a. * The enemy's deliberate destruction was well-documented.

b. The enemy's deliberate destruction of the city was well-documented.

Based on this and other evidence, Grimshaw concludes that event nominalizations have argument structure, like verbs, including an external thematic-role which is linked to the possessor element. More generally, these conclusions are emblematic of a long tradition of formal parallels between clauses and nominal expressions (Rosenbaum, 1965; Steven Paul Abney, 1987; Chomsky, 1970, a.o.). The important point is that the possessor of an event nominalization is plausibly a 'subject' of a nominal domain.

3.2.2 *Case & syntactic context*

Finally, subjects that do not vary in thematic role or structural position can nonetheless differ in their syntactic context—their case and/or the properties of their host clause. To see this, consider the sentences in (31). The embedded pronominal subject bears NOMINATIVE case in (31a), but ACCUSATIVE in (31b). For this reason, verbs like *believe*, which take an embedded accusative subject, are called 'Exceptional Case-Marking' (ECM) verbs. Relatedly, the embedded pronoun is the subject of an embedded finite complement clause in (31a; 'S-complement'), but the subject of a non-finite (untensed) clause in (31b). Crucially, though, contemporary syntactic analyses locate both embedded subjects in SPEC-T. The relevant structures are schematized below, in Figure 3.2.

(31) a. Terry believes (that) *she* is intelligent. S-complement

b. Terry believes *her* to be intelligent. ECM

Note that the ECM variants also show a surface string that is almost identical to another class of verbs embedding a non-finite complement, namely (Object) Control verbs like *persuade*:

(32) a. Terry persuaded *her* to leave.

b. Terry_i persuaded her_j [$\text{PRO}_{j/*i}$ to leave]

In (32) the embedded subject is implicit, and obligatorily co-referent with the matrix object of *persuade*. More specifically, the embedded subject is the null pronominal element PRO, anaphorically related to the matrix object (the 'controller') via a Control dependency.

3.3 Complex subject attachment

Content-addressable retrieval is well motivated in the domains of episodic memory and recall/recognition processes, but seems ill-suited for the task of language processing. For one, despite the errors discussed above, our conscious perception of language comprehension suggests that processing is relatively effortless and error-free. On the one hand, there is very strong evidence that memory access is content-addressable, including the access required for language processing. On the other hand, error-free processing suggests that retrieval nonetheless is able to navigate grammatical constraints defined in terms of hierarchical representations (e.g., Binding

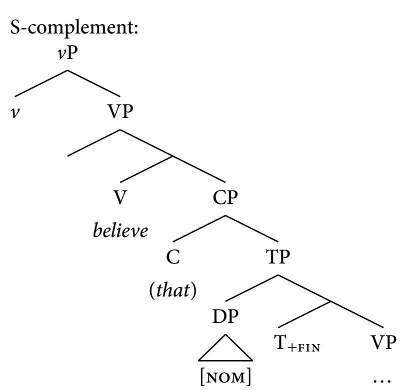


Figure 3.2: Syntactic representions of S-complement (left) and Exceptional Case-Marking (right) verbs.

constraints; see also §1.1). This sets up a tension between the grammatical competence systems and the performance systems in which they are embedded, since the 'content' of encodings are generally assumed to be inherent 'item-level' information, such as the linguistic features PLURAL or CATEGORY, and not relational information like c-command or phrase-structural properties that cannot be encoded on a single constituent. Rather, they encode the relation between two constituents. Furthermore, linguistic representations are combinatoric objects, recursively constructed from a finite featural vocabulary. If these features are also the retrieval cues, then linguistic representations will be highly self-similar, so that the problem of interference is particularly acute for language (Wagers, 2014).

Several questions arise at this point. Do all linguistic dependencies involve cue-driven retrieval, like agreement? If so, how has the system adapted to overcome interference, if at all, in a way that allows it to navigate linguistic constraints? Relatedly, do interference effects lead the parser to construct grammatically illicit dependencies? In this section, I discuss a series of studies that indicate that increasing the complexity of a subject engenders difficulty at its verb. Across these studies, there is general consensus that the source of difficulty is similarity-based interference in the retrieval required to integrate a subject with its verb. Throughout, I focus on the implications of these studies for the theory of retrieval cues, setting the stage for the model presented in Chapter 5. Along the way, however, we will see that the core evidence is undermined by a confound with the depth of embedding. This leaves the conclusion that subject-verb attachment is interference-prone open to an alternative view that makes no reference to interference. The section concludes with an experiment that seeks to extend previous findings of interference in complex subject attachment to materials that (i) control for depth of embedding, and (ii) make distinct predictions about the dimensions of similarity that engender interference in complex subject attachment.

3.3.1 Syntactic similarity matters

There is some evidence that the basic structure-building operations required for dependency resolution are interference-prone. In a series of studies, Julie Van Dyke and colleagues probed the effects of attaching complex subjects on the processing of a subsequent verb, where the subject must be retrieved. Using a self-paced reading methodology, Van Dyke & Lewis (2003, hereafter, V&L) examined the comprehension of subject-verb dependencies in sentences like (33), where a subject-attached relative clause intervenes between subject and verb.

(33) Van Dyke & Lewis (2003)

a. High interference

The secretary forgot that the student who thought that *the exam* was important was standing in the hallway.

b. Low interference

The secretary forgot that the student who was waiting for *the exam* was standing in the hallway.

c. Control

The secretary forgot that the student was standing in the hallway.

V&L hypothesized that integrating a subject with its verb requires retrieval, and that this retrieval is prone to similarity-based interference. Thus, their critical region was the auxiliary of the VP *was standing*, where the subject *the student* is retrieved and integrated. To probe for interference effects, V&L manipulated similarity between the target subject and material inside the complex subject. In their 'high interference' conditions (33a), the potentially interfering element, *the exam*, was the structural subject of the embedded clause. In contrast, for the 'low interference' conditions, the potential interferer occurs in object position. These conditions were compared to control sentences (33c), where the critical VP was adjacent to its subject. V&L also compared these sentences to ambiguous variants, constructed by dropping the complementizer *that*.

I have summarized V&L's RT results in Figure 3.3. In both online and offline measures, the processing at the matrix verb regions, *was standing*, was more difficult – slower RTs, reduced comprehension accuracy – when the interpolated material contained a lexical[2] subject (a), than

[2] Here and thoughout I use the phrase 'lexical subject' to refer to pronunced or overt subjects, and contrast them with other types of subjects that need not be overt (such as the implicit subject of an embedded infinitive clause). Note also that both lexical and implicit subjects may be 'structural subjects' by virtue of their syntactic position.

when it did not (b). The reading time results showed main effects of both ambiguity and interference. Overall, ambiguous sentences were slower than unambiguous sentences. The key result, though, was that in both ambiguous and unambiguous conditions, processing at the verb was more difficult for high interference sentences than low interference sentences. There were no reliable differences between short and low interference conditions, though the effect was trending significant ($p = .09$) for unambiguous sentences. Surprisingly, however, the difference between high interference and short conditions was only reliable for ambiguous sentences.

The comprehension data showed roughly the same patterns of interference: ambiguous sentences were more difficult to understand, though this effect was driven by high interference conditions only, and high interference sentences were more more difficult to comprehend than either low interference or short conditions. As in the reading time data, there were no reliable differences between the short and low interference conditions.

Van Dyke & Lewis (2003) attribute the relative difficulty in high-interference sentences (33b) to structural similarity between the two subject encodings. Assuming that the subject is retrieved at the verb (see above), this similarity between potential subjects leads to difficulty in discriminating the correct subject. Thus, in (33) the verb triggers retrieval of the subject *the student*

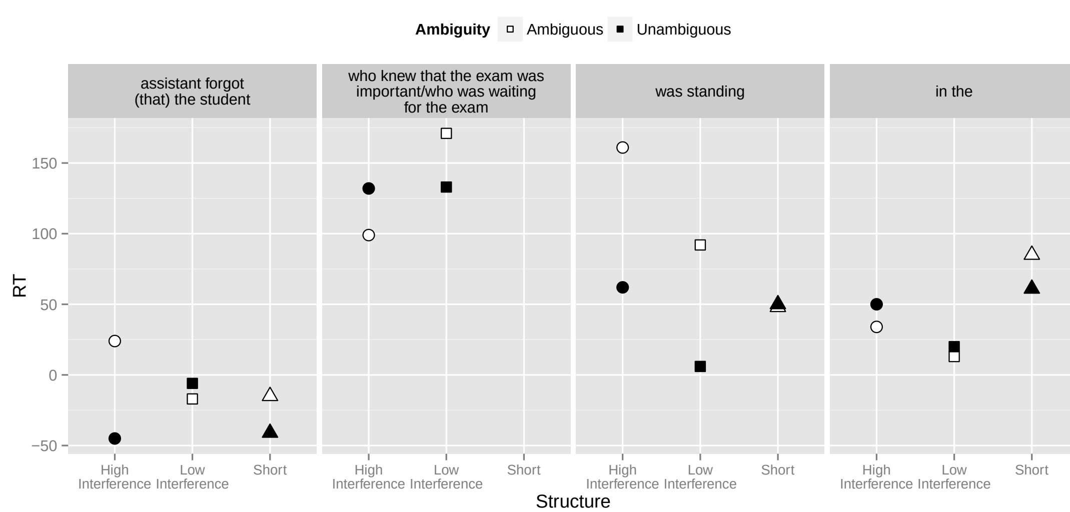


Figure 3.3: Results from Van Dyke & Lewis (2003).

via the retrieval cues. When matched against the contents of memory these cues identify two encodings in (b), *the student* and *the exam*. The relative difficulty of (b) over (a) reflects competition between these two outcomes, which Van Dyke & Lewis (2003) capture with the feature "[+SUBJECT]". More generally, the results of Van Dyke & Lewis (2003) suggest that the retrieval cues are informed by syntactic information present at the retrieval site. Furthermore, the effects in unambiguous sentences suggest that cue-directed retrieval is employed in the basic structure-building operations of syntactic parsing, and not simply repair.

However, the results of Van Dyke & Lewis (2003) are also consistent with another explanation that undermines the crucial contrasts between high and low interference conditions. Note that the high interference conditions contain one more clause than the low interference conditions. The verb *thought* introduces an additional finite clause complement that hosts the interfering subject. It may be, then, that three clauses are simply more difficult than two. The reading time results provided suggestive evidence for this contention: high interference conditions were also slower than the low interference conditions in the intervening region, prior to the critical region. If the contrast between interfering conditions is the result of retrieval interference, as V&L claim, then it is worrying that we see the same interference effect prior to the verb, where the subject is retrieved. However, the difference at the intervening regions was only significant in the raw reading times, not the residual reading times. Thus, as V&L note, the difference is likely due to the presence of an additional word in the high interference conditions. Since the residual reading times correct for effects of word position and character length, the lack of the effect in residual times suggests that the slowdown for high interference conditions at the intervening regions reflects the additional time spent reading the extra word, not interference.

In any case, the materials of Van Dyke & Lewis (2003) confound the similarity manipulation – the additional subject – with the depth of embedding, which undermines conclusions of interference from cue-competition between multiple subject encodings. We return to this issue below, in (SECTION), but first we consider additional evidence that complex subject attachment engenders difficulty at a subsequent VP, and that this difficulty is due to interference from grammatically inappropriate subject encodings.

3.3.2 Semantic similarity matters

In a related study, Van Dyke (2007) investigated whether subject retrieval is guided by semantic information, using an eye-tracking methodology supplemented with offline comprehension questions. As in Van Dyke & Lewis (2003), syntactic interference was manipulated through the inclusion of an intervening lexical subject. The syntactic similarity manipulation was crossed with a semantic similarity manipulation that varied the appropriateness – *i.e.* animacy – of the intervener as the subject of the matrix verb. Thus, the syntactic interference of Van Dyke & Lewis (2003) will be reflected in a slowdown at the verb for (b,d) relative to (a,c). If semantic similarity also engenders interference, then animate interveners (c,d) will lead to difficulty at the verb above inanimate interveners (a,b).

(34) The pilot remembered that the lady…

a. …who was sitting in *the smelly seat* … LoSyn/LoSem

b. …who was sitting near *the smelly man* … LoSyn/HiSem

c. …who said that *the seat* was smelly … HiSyn/LoSem

d. …who said that the *the man* was smelly … HiSyn/HiSem

… yesterday afternoon *moaned* about a refund for the ticket.

Van Dyke found that the difficulty for multiple subjects was more acute when the intervening subject was also semantically similar to the target subject. Online measures indicated an early effect of syntactic similarity, so that an intervening nominal expression in subject position led to longer first-pass and regression-path times at the critical verb in the LoSem conditions. There was no effect of semantic similarity at the critical region, but semantically similar interveners led to slower regression path times at the spillover region when the intervener was in subject position. The strength of the interference effects was stronger in the offline comprehension questions, which showed main effects of both semantic and syntactic similarity, with the high interference conditions showing lower accuracy. Van Dyke also reported the proportion of trials in which the intervening subject was taken as the subject of the matrix verb. When the intervener was a syntactic subject, it was chosen as the matrix subject almost twice as often as when it was contained inside an adjunct (34c–d versus 34a–b).

The results of Van Dyke (2007) thus replicate the effect of interference from syntactic subjects observed in Van Dyke & Lewis (2003), while extending these effects to semantic similarity, namely animacy. Importantly, the dimensions of semantic similarity are relative to the verb triggering retrieval: *the man*, but not *the seat*, is a semantically appropriate subject of *moan*. We would thus expect the effects to be modulated by the meaning of the predicate that triggers retrieval. For instance, the effect might be reversed for the retrieval of a VP like *was torn*, rather than *moan*.

In a related study, Van Dyke & McElree (2006) provide evidence that the dimensions of similarity are defined with respect to the cues provided by the verb, and not overall similarity between DPs. In other words, interference arise at retrieval, and not encoding. Van Dyke & McElree (2006) examined reading times for sentences like (35).

(35) a. High interference:
It was the boat that the guy who lived by the sea *fixed* in two sunny days.

b. Low interference:
It was the boat that the guy who lived by the sea *sailed* in two sunny days.

Interference was manipulated by carrying the form of the critical verb. In interfering conditions, the verb was *fixed*, while in low-interference conditions the verb was *sailed*. The reading task was crossed with a memory load task (Gordon, Hendrick, & Levine, 2002). In memory load conditions, participants memorized a list of nouns, exemplified in (36). The nouns on this list were selected so that they would match the selectional requirements of the verb in the interfering conditions: tables, sinks, and trucks are *fixable*, but not *sailable*. Thus, interference will be reflected as an interaction between sentence type and memory load, with the greatest difficulty for memory load conditions with a *fixed* verb.

(36) Memory load list:
table-sink-truck

Consistent with interference, the reading times at the matrix verb showed an interaction of verb type and memory load. Reading times at the verb were longer for memory load conditions with *fixed* sentences. Importantly, experimental manipulations did not manipulate similarity between DPs, but held the encoding conditions constant. The interference manipulation only varied the fit of the memory load list to the selectional properties of the critical verb. Thus, the fact that difficulty arose only when elements of the memory list were comprised of potentially good subjects of the verb *fixed* indicates that the difficulty is due to retrieval interference from similarity between the memory list and the cues provided by the verb triggering retrieval.

In summary, the results of both Van Dyke (2007) and Van Dyke & McElree (2006) indicate that the retrieval cues for subjects are informed by the meaning of the verb triggering retrieval, so that semantically similar – but nonetheless grammatically inappropriate – DPs interfere with subject retrieval. These results are also consistent with the results of Van Dyke & Lewis (2003), which indicate that syntactic subjects also interfere. I turn now to some evidence that further supports retrieval cues for semantic and syntactic subjects, but suggests that structural information is prioritized above semantic information.

3.3.3 *The position of the intervener matters*

Van Dyke & McElree (2011) examined how the semantic properties of intervening nouns affected processing at the matrix verb. They reasoned that if semantic cues are used to guide retrieval, then semantically similar, but nonetheless grammatically inappropriate, interveners should engender interference. In their first experiment, the intervener was located in the embedded subject position, while in the second experiment the intervener occurred in object position[3]

(37) a. The attorney who the judge realized had declared that *the motion* was inappropriate compromised.

b. The attorney who the judge realized had declared that *the witness* was inappropriate compromised.

[3] Van Dyke & McElree (2011) also crossed these sentences with a proactive interference manipulation, where the distractor preceded the target (not shown).

Van Dyke & McElree (2011) used both SAT and eye-tracking methodologies to examine retrieval speeds and processing difficulty, respectively. In their first Experiment, participants read sentences such as those in (37). In the SAT results, they observed lower asymptotic accuracy when the animacy of the intervener matched that of the target, so that (37b) was less accurate than (37a). However, consistent with previous work (McElree, 2000; McElree, Foraker, & Dyer, 2003) there was no evidence of differences in the retrieval dynamics, further supporting a content-addressable retrieval mechanism. In the eye-tracking results, the effect of semantic similarity was significant in the total times, with slower total times for animate interveners (37b). Overall, then, the results of Van Dyke & McElree (2011) replicate previous findings of interference from grammatically inappropriate interveners that match the retrieval cues at the verb, and indicate that semantic cues (such as animacy) are among those cues (Van Dyke & McElree, 2006; Van Dyke, 2007).

In a second experiment, Van Dyke & McElree (2011) used the same methodologies to examine the effect of semantic cues when the intervener mismatched the syntactic cues. A sample of the relevant conditions is shown in 38. Note that in this case the intervener occurs in object position, and not subject position.

(38) a. The attorney who the judge realized had rejected *the witness* in the case compromised.

b. The attorney who the judge realized had rejected *the motion* in the case compromised.

As in their first experiment, the SAT results gave no indication of differences in retrieval dynamics. Interestingly, the effect of semantic interference observed in Experiment 1 (see also Van Dyke, 2007) was absent in sentences such as 38. While Van Dyke & McElree (2011) did not set out to directly test the effect of intervener position, and so the crucial contrast is between experiments, the contrasting interference profiles of Van Dyke & McElree (2011) strongly suggest that the structural position of the intervener is a primary determinant interference effects.

(39) [$_{TP}$ The attorney who [$_{TP}$ the judge realized [$_{TP}$ had rejected *the witness* in the case]] compromised.]

(40) [$_{TP}$ The attorney who [$_{TP}$ the judge realized [$_{TP}$ had declared that [$_{TP}$ *the motion* was inappropriate]]] compromised.]

It is perhaps unsurprising that the structural position of the intervener matters, given the relational nature of linguistic constraints. In a content-addressable memory architecture, however, such effects raise a number of critical questions that must be answered if the theory is to make predictions about novel structures.

3.3.4 *The confound with depth of embedding*

The studies by Van Dyke and colleagues, reviewed above, indicate that increasing the complexity of a subject leads to difficulty at the verb, when the subject is retrieved. The relevant constructions are summarized in Table 3.4. Consistent with their claims of interference effects, this difficulty is more acute when the additional material of the subject includes other subjects, which are syntactically and/or semantically similar to the target subject. However, this conclusion is undermined by the fact that in all three experiments, the presence of an additional, embedded subject – the interference manipulation – is confounded with the presence of an additional clause. This is a worry because the depth of embedding has been shown to correlate with both reading comprehension and working memory tasks (Blaubergs & Braine, 1974; Roberts & Gibson, 2002).

Subsequent studies seeking to replicate the findings of Van Dyke and colleagues while con-

SOURCE	SAMPLE CONSTRUCTION
Van Dyke & Lewis (2003)	The secretary forgot that the student who thought that the exam was important was standing in the hallway. *The secretary forgot that the student who was waiting for the exam was standing in the hallway.*
Van Dyke (2007)	The pilot remembered that the lady who said that the man was smelly yesterday afternoon moaned about a refund ticket. *The pilot remembered that the lady who said that the seat was smelly yesterday afternoon moaned about a refund ticket.*
Van Dyke & McElree (2011)	The attorney who the judge realized had declared that the motion was inappropriate compromised. *The attorney who the judge realized had rejected the witness in the case compromised.*

Figure 3.4: Constructions that have been argued to engender interference in subject retrieval. Italicized sentences showed reduced interference.

trolling for depth of embedding have found reduced evidence of interference in online measures. Wagers (2008, Experiment 6) used self-paced reading to measure processing of sentences similar to those in Van Dyke & Lewis (2003) and Van Dyke (2007), where a bi-clausal relative clause introduces an embedded finite clause hosting an additional subject. Syntactic interference was manipulated by comparing overt lexical subjects to *expletive-associate* constructions, where the structural subject is a dummy element, *there*, and the thematic subject, *support*, occurs lower. This innovation allowed for variation in the presence of a lexical intervening subject while holding the number of clauses constant.

(41) a. *Support* was widespread for the candidate.

b. *There* was widespread support for the candidate.

The expletive *there* in (41b) shares a number of properties with the lexical (overt) subject in (41a). Both *there* and *support* in (41) occupy the structural subject position, here assumed to be SPEC-T.

Sentences with intervening subjects were compared to sentences without an intervening clause: either a PP modifying the subject noun or no intervening material. The resulting experimental design was 2×2, manipulating the presence of an intervening subject and the number of intervening clauses.

(42) *Materials from Wagers (2008, Experiment 6)*

The politician was displeased that…

a. Lexical subject:

…the report that *support for her opponent* was widespread…

b. Expletive-associate:

…the report that *there* was widespread support for her opponent…

c. PP modifier

…the report of widespread support for her opponent…

d. Adjacent

…the report…

…was covered on the evening news.

If the effects observed in Van Dyke & Lewis (2003) and Van Dyke (2007) were due to structural similarity between the embedded subject and the target subject, then sentences with a lexical subject in SPEC-T should be more difficult than the expletive-associate sentences. This is because the potentially interfering element in (42a) is more subject-like than the dummy subject of (42b). On the other hand, if processing difficulty at the verb is a function of the number of intervening clauses, then sentences where a clausal complement intervenes between the subject and verb (42) will be more difficult than conditions where either no clause (42d), or a non-clausal PP modifier (42c) intervenes.

The results of Wagers (2008) both challenged and supported the results of Van Dyke and colleagues. In online measures, there were no differences between lexical intervener conditions, but clausal complements showed difficulty in the spillover region, relative to the PP-modifier and adjacent conditions. These patterns add empirical bite to the contention that the results of Van Dyke & Lewis (2003) and Van Dyke (2007) are due to the presence of an additional clause, not an additional subject. The offline comprehension question data, though, largely paralleled the results of Van Dyke (2007) in showing a cline of difficulty: the lexical subject condition was significantly more difficult than either the expletive associate condition or conditions without an intervening clause. Overall, then, much of the core evidence for retrieval interference at a verb due to competition between multiple subject encodings is confounded with depth of embedding, at least in online measures.

Note, though, that the associate itself is also subject-like. For instance, it controls agreement, both morphological (43a) and semantic (43b). It is also the element that receives the external thematic-role (44).

(43) a. There is/*are a cat on the shelf.

b. Outside the courthouse, there gathered a crowd/*a man.

(44) a. There are *children* playing in the garden.

b. # There are *rocks* playing in the garden.

In (43a), the agreement on the verb *be* is controlled by the associate *cat*, while in (43b) the requirements of *gather* for a semantically plural subject are satisfied by the associate. Similarly, when the associate is infelicitous when it cannot bear any thematic-role assigned by the predicate, like the AGENT required for *playing* in (44b). In both cases, the nominal that satisfies the selectional restrictions of the predicate is the associate, which also occurs after the verb. The case requirement of the associate is analyzed as being satisfied via NOMINATIVE case on the expletive, using the formal device of a *chain*[4] (Chomsky, 1986; Chomsky & Lasnik, 1993).

While the expletive and associate are separate lexical items, they are interpreted – formally speaking – as a single element. If we assume that these formal mechanisms mirror the way in which expletive associate pairs are processed in (42), then both the lexical subject and expletive associate conditions contain an additional subject-like element not present in the PP-complement or adjacent conditions. The expletive-associate sentences differ in that the subject properties of case and thematic-role are distributed across two elements, and not manifested on the same constituent. We might thus expect, given the results of the Van Dyke studies above, and Van Dyke & Lewis (2003) in particular, that interference should be greater in the lexical subject condition where a single constituent matches the verb's retrieval cues. The lack of such an effect in the results of Wagers (2008), combined with the difficulty for intervening clauses, and not subjects, significantly undermines the conclusion of Van Dyke and colleagues that their observed difficulty at the verb was due to interference from an embedded subject. In particular, their results are consistent with an alternative view that takes the difficulty to reflect the presence of an additional clause, which might impose storage demands on memory without increasing similarity-based interference.

[4]In more contemporary analyses, both expletive and associate are licensed through an Agree dependency (e.g. Chomsky, 2008) with a functional head, T(ense). Under this view, agreement between the verb and the associate is established by Agree, while the expletive satisfies the requirement for a subject (specifier) and receives NOMINATIVE case from T. The important point for our purposes is that the expletive and associate are linked through the same formal mechanism, with their relation mediated by the functional T head.

3.4 Experiment 1: Structural versus thematic subjects

Experiment 1 had two goals. One, to probe whether interference effects from grammatically inappropriate subjects reflect similarity to the target in terms of argument-structure or phrase structure. Two, to replicate previous findings of interference from embedded subjects using materials that control for depth of embedding.

3.4.1 Design & predictions

We address these issues using sentences such as that shown in (45), by probing for interference at the matrix verb from an embedded subject that intervenes between a verb and its target subject, while manipulating similarity between the intervening subject and the (putative) retrieval cues provided by the verb. We manipulate thematic similarity by comparing sentences such as (45) to variants without a possessor.

(45) The hostess who thought that the (chef's) careful preparation of the blowfish delayed the guest was yelling in the kitchen.

The possessor *chef* in (45) is thematically similar to the target matrix subject in that both are an external argument. Both elements are also configurationally similar, in that they are the most prominent structural element within their domain. They differ in their syntactic context—a nominal for the possessor, and a clause for the target subject—and their case: genitive or nominative, respectively. Thus, the sentence in (45) contains three thematic subjects – *hostess*, *chef*, and *preparation* – but only two structural subjects, namely *hostess* and *preparation*.

While the nominalization is required to make the possessor sufficiently 'subject-like', it also increases the complexity of the intervening subject. There is evidence that increasing a constituent's complexity modulates difficulty in retrieval. Hofmeister (2011) found that increasing the complexity of a to-be-retrieved constituent facilitated processing at the integration site, so that processing at a verb was faster when its subject was complex, e.g. *a communist* versus *an alleged Venezuelan communist*, but complexity had no effect on processing at the intervening regions. In configurations like 45, the complex constituent is not the target of retrieval, and so increasing the complexity of the intervening subject may facilitate (mis)retrieval of the embedded subject, and lead to difficulty when the matrix subject is retrieved.

In order to examine the role of intervener complexity on retrieval, we compared sentences like (45), to sentences like (46), where the nominalization occurs object position. These sentences simply swapped the embedded subject and object, so that lexical items remained constant across conditions, modulo the possessor.

(46) The hostess who thought that the guest delayed the (chef's) careful preparation of the blowfish was yelling in the kitchen.

However, the additional word for the possessor might matter, either because it introduces an additional discourse referent, or because it increases overall similarity between constituents in (e.g.) category. While an overall cost for the possessor will be reflected as increased difficulty in possessor conditions, if the possessor engenders retrieval interference then the difficulty will not arise before the critical verb, when retrieval is required.

The sentences in (45) and (46) all contain an embedded structural subject, either a nominalization or a simple *Det+Noun* constituent. In order to establish interference for the presence of an embedded subject, we also compared the sentences in (45–46), the 'interference conditions', to control sentences such as that shown in (47).

(47) The hostess who seemed to have delayed the careful preparation of the blowfish was yelling in the kitchen.

Control sentences used raising predicates, e.g. *seem*, to leave the embedded subject unexpressed, and co-referent to the matrix (target) subject. Our control sentences are analogous to the 'low interference' conditions of Van Dyke & Lewis (2003) and Van Dyke (2007), but contained an additional clause (*to have delayed*) to control for the number of clauses across experimental conditions. Otherwise, there were only minimal lexical differences between the control condition and the interference conditions. In particular, like the simple intervener sentences in

(46), the control sentences contained a nominalization in object position. This was done to easily match lexical material, and guard against an overall cost for the additional complexity of the nominalization in the interference conditions.

The resulting experimental design was 2 × 2 + 1, with factors for the presence of a possessor (yes/no), the position of the complex nominalization constituent (subject or object) plus the control condition. For all experimental sentences, the critical regions are the matrix verbs, e.g. *was yelling*, where the verb is integrated with the matrix subject.

The predictions are as follows. If embedded lexical subjects engender interference, then the interference conditions will be more difficult than the control conditions, reflected as either a slowdown at the critical regions or reduced comprehension accuracy. Predictions about the source of interference depend on comparisons within the interference conditions. If the cues at the verb characterize subjects in terms of their thematic properties, then similarity between the possessor and target matrix subject engender interference at the verb, and sentences with a possessor will be more difficult than sentences without a possessor. On the other hand, if the verb's cues identify subjects by their structural properties, *i.e.* case or position, then we expect no effect of the possessor. Finally, if complexity affects constituent retrieval, then we expect that greater difficulty at the verb for sentences with a nominalization in subject, but not object position, due to (structural) similarity between the complex nominalization and the target matrix subject.

OSPAN working memory task

Looking ahead to Chapter 3, models of sentence processing based on cue-driven parsing take processing difficulty to reflect working memory constraints, in particular interference effects (Lewis & Vasishth, 2005). One potential prediction of this view is that comprehenders with greater working memory span will be more resistant to interference effects, either because they can maintain more information in active processing states or because of an ability to more durably encode constituent representations in a way that makes them more available at retrieval. Indeed, performance on working memory tasks has been shown to correlate with individuals' performance on a variety of cognitive tasks, including reading comprehension (Daneman & Carpenter, 1980; Jonathan King & Marcel Adam Just, 1991; Roberts & Gibson, 2002).

To examine the effect of working memory performance on subject retrieval, the reading time results were correlated with participants' performance on the Operation Span (OSPAN) working memory task (Unsworth et al., 2005). The OSPAN task is an outgrowth of the Reading Span task of Daneman & Carpenter (1980), in which participants are asked to memorize a list of serially presented symbols (here, alphabetic characters), then repeat the list back in reverse order. The OSPAN task is similar, except that the presentation of the to-be-remembered items is interrupted by an additional distractor task, such as an arithmetic problem, designed to displace the list from focal attention and prevent rehearsal strategies. Experiment 1 used an automated version of the OSPAN task from (Unsworth et al., 2005), who demonstrate that the task accounts for additional variance in models of reaction time studies.

If working memory span correlates with interference effects, then we expect that participants' OSPAN scores will be a significant predictor of their performance on the self-paced reading task, either reading times or comprehension accuracy. Note, however, that this is not a necessary prediction: it may be that the resources required for structure-building operations like subject-attachment are outside the scope of inter-subject variation. Thus, while a significant correlation between WM performance and interference effects would support cue-driven parsing models, the lack of such a correlation does not constitute evidence against these models.

3.4.2 Method

Participants

Participants were 40 native speakers of English from the University of California, Santa Cruz community. All participants provided informed consent. Compensation for participation in the experiment was $15.

Materials

Sample materials for Experiment 1 are shown in Table 3.1. The full materials for all experiments are provided in Appendix A. Experimental materials consisted of 30 item sets arranged in a 2 × 2 + 1 factorial design that crossed the presence of a *possessor* (yes/no) with the *complexity* of the intervening subject (nominalization/simple DP).

All experimental sentences began with a 1-2 word adverbial phrase, e.g. *Somewhat surprisingly*, to eliminate a potential primacy effect for the target subject occurring trial-initially. Following this phrase, the first five words of each experimental item were of the form: *det-N-who-V-that*. The noun is the target subject, and the verb always embedded a clausal complement of the form DP_1-V-DP_2. In complex intervener conditions, the embedded subject DP_1 was an event nominalization and the embedded object DP_2 was a simple *det-noun* sequence. Simple intervener conditions reversed the order of these arguments, with a *det-noun* nominal as the embedded subject and the nominalization the embedded object. Modulo the possessor, the four experimental conditions were controlled to contain the same lexical material. The structure of the nominalization was always of the form *the-(Poss)-Adj-Nominalization-PP*. The nominalization always contained a subject-oriented adjective like *deliberate* to force a reading in which the noun had an eventive argument structure and the possessor is unambiguously thematic (Grimshaw, 1990). The 30 sets

Possessor	Sample sentence
Complex intervener conditions	
Poss	The hostess who thought that the chef's careful preparation of the blowfish delayed the server was yelling…
No Poss	The hostess who thought that the careful preparation of the blowfish delayed the server was yelling…
Simple intervener conditions	
Poss	The hostess who thought that the server delayed the chef's careful preparation of the blowfish was yelling…
No Poss	The hostess who thought that the server delayed the careful preparation of the blowfish was yelling…
Control condition	
	The hostess who was fortunate to have delayed the chef's careful preparation of the blowfish was yelling…

Table 3.1: Sample set of experimental materials for Experiment 1. Not shown: Initial adverbial phrase and sentence-final prepositional phrase.

of five conditions were distributed across five lists in a Latin Square design, and combined with 114 filler sentences of similar length. All of the sentences were grammatical.

Each trial was followed by a comprehension question. These were designed to probe comprehension of the subject-verb relation in the matrix and embedded clauses. A third question type probed the interpretation of the sentence-final verb-PP modifier, to serve as a baseline for measures of interference in comprehension accuracy. Negative versions of each comprehension question were created by substituting the correct matrix/embedded subject with the subject of the embedded/matrix clause, respectively. This avoids participants answering based on a simple recognition judgement of whether the subject noun occurred in the sentence or not.

Procedure

Sentences were presented on an Apple desktop computer using the Linger software (Doug Rohde, MIT) in a self-paced word-by-word moving window paradigm (Marcel A Just, Carpenter, & Woolley, 1982). After each sentence, a yes/no comprehension question was presented in its entirety. Participants were instructed to read each sentence at a natural pace, and to answer each question as accurately as possible. Feedback was provided for incorrect responses only. Order of presentation was pseudo-randomized for each participant in a Latin square design. Each experimental session began with six practice trials.

Analysis

Experimental data was analyzed in the R programming environment for statistical computing (R Core Team, 2016), and modeled in a series of linear mixed-effects models using the *lme4* package (D. Bates et al., 2013). Prior to analyzing the reading times, extreme observations less than 50 ms and greater than 3000 ms were removed. This exclusion affected 0.17% of the data. Six participants were removed for extremely low accuracy (<65%). RTs were analyzed in regions consisting of a single word and aligned prior to analysis, as in (48).

(48) Experiment 1 analysis regions:

Target the$_1$ hostess$_2$

Embedding V who$_3$ thought$_4$ that$_5$

Intervener (*complex*) the$_6$ chef's careful preparation$_9$ of the blowfish

(*simple*) the$_6$ guest$_9$

Embedded VP (*complex interveners*) delayed$_{13}$ the$_{14}$ guest$_{17}$

(*simple interveners*) delayed$_{13}$ the$_{14}$ chef's careful preparation$_{17}$ of the blowfish

Critical was$_{21}$ yelling$_{22}$

Spillover in$_{23}$ the$_{24}$ kitchen$_{25}$

While the ordinal position of the critical region in all conditions varied by only one word (due to the possessor), the embedded verb occurred earlier in simple intervener and control conditions than complex intervener conditions. In order to correct for these differences, and correct for differences in participants' reading rates, statistical analysis of reading times was performed on the residual reading times. This correction has been advocated by (Ferreira & Clifton, 1986) in order to control for the effect of character length on reading times (Rayner, 1977; Rayner, Sereno, & Raney, 1996), while more recently a growing number of researchers have used this method to control for this and other confounding factors (e.g. Van Dyke & Lewis, 2003; Hofmeister, 2011; Jaeger, Fedorenko, & Gibson, 2010). In addition to facilitating comparisons to previous studies, residual reading times have the advantage of allowing results to be reported on a natural scale, namely milliseconds. Residual reading times were computed by first fitting a first-order linear regression equation to the filler sentence reading times for each subject, with predictors for word-length in characters, the position of the trial sentence in the materials list, and the position of the word within the sentence. Since the effect of word-position is non-linear, reflecting differences in reading rates, we used the natural log of word-position. For each participant, the model from filler data was used to predict the RTs for experimental sentences as a function of word-length, trial position, and word position. The residuals of this model were then used as the dependent variable in the reported statistical models.

Outliers were removed by first computing z-scores of residual RTs by region and condition, then removing observations whose absolute z-score exceeded $|3|$ (Grodner & Gibson, 2005). This criteria excluded 1.21% of the data. Residual RTs at each region were analyzed in a series of linear mixed-effects models with fixed-effects for experimental factors. For comparisons within the interference conditions, experimental factors were modeled using deviation coding (.5,-.5), with complex intervening subjects and POSS:YES as the positive coefficients. Comparisons between interference conditions and the control condition were analyzed in a separate maximal model, which used treatment coding (0, 1) with the control condition as the reference level, so that coefficients reflect by-condition differences from the control mean. All models follow the recommendations of Baayen, Davidson, & D. M. Bates (2008) and Barr et al. (2013) in using the maximal random-effects structure (experimental factors nested under participants and items). When the maximal model would not converge, a step-wise, backwards procedure was used: inspecting the model and dropping the random effect term associated with the lowest variance. The significance of experimental coefficients was evaluated by interpreting the t-score of the coefficient as a z-score, with significant coefficients those whose *t/z* exceeded 2 (Gelman & Hill, 2007; Levy, Fedorenko, & Gibson, 2013, see Baayen, Davidson, & D. M. Bates, 2008 for justification).

Comprehension accuracy data was analyzed in a series of logistic regression models (Jaeger, 2008), using the same contrast coding described above for the RT analysis. Unless otherwise noted, all reported effects for both RT and accuracy models were significant at α = .05 or less.

3.4.3 *Results*

Self-paced reading

The reading times results from Experiment 1 are presented in Figure 3.5. Overall, we found that increasing the complexity of the intervening subject increased processing difficulty at the matrix verb regions, but no cost for the possessor.

Critical & spillover regions There were no significant effects or interactions at the matrix auxiliary, and no significant differences between the interference conditions and the control condition. While the auxiliary is the point at which the subject must be retrieved and integrated, by

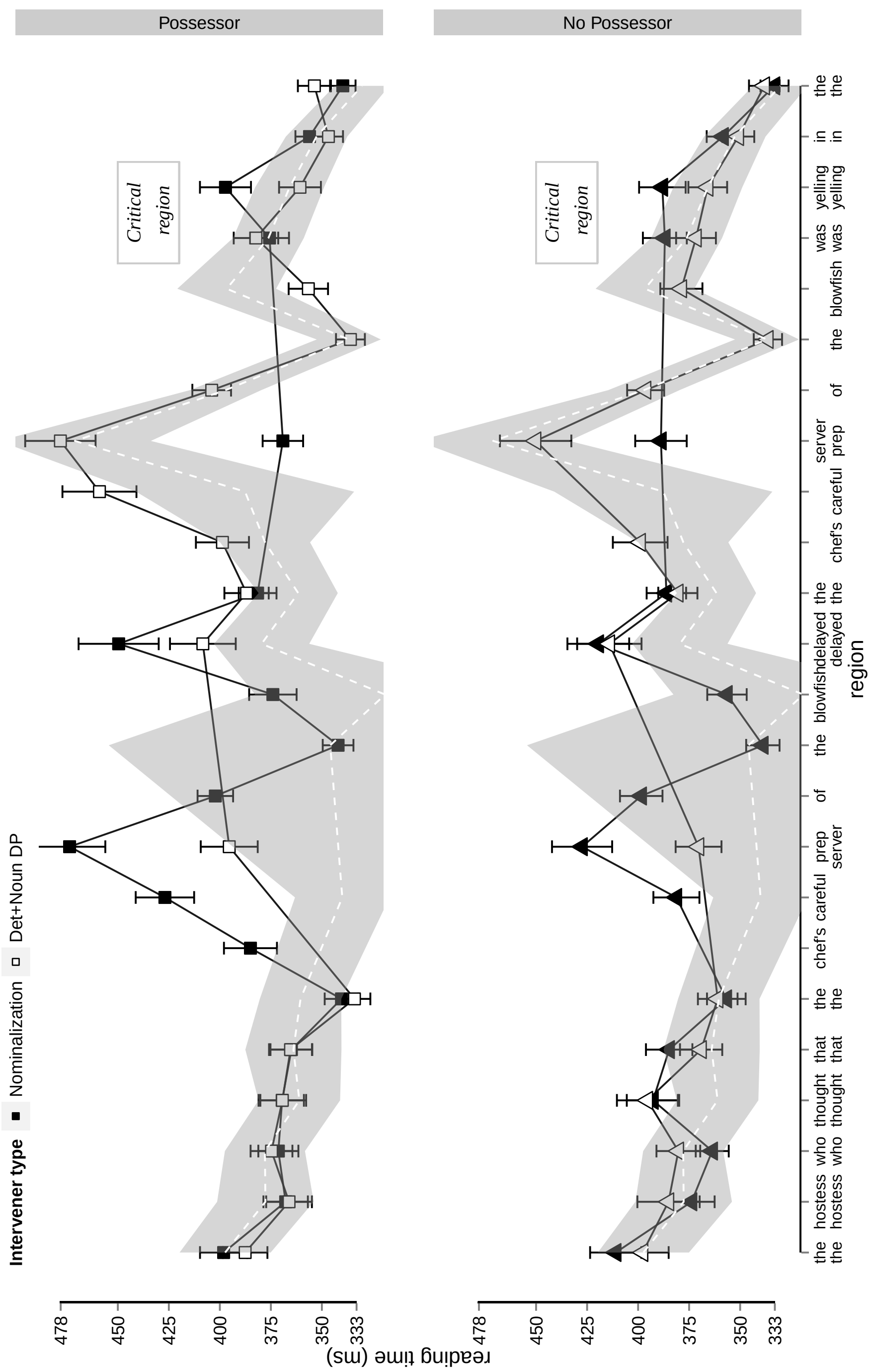


Figure 3.5: Experiment 1 self-paced reading times for all conditions, aggregated over participants and items, showing condition means at each region (symbols) and standard error of the mean (error bars). Control sentences are summarized by the dotted line (means) and the shaded area, which shows the 95% c.i. of the control means for each region.

hypothesis, we suggest that the lack of an effect at the aux region is an artifact of the moving-window display. The auxiliary signals a retrieval, but either readers' forward momentum or the predictability of the incoming VP causes readers to advance to the next word before subject retrieval happens or its effects are felt.

Difficulty arose at the matrix main verb. Sentences containing a nominalization in subject, but not object, position led to a slowdown at the matrix verb, a main effect (β = 20 ms; s.e. = 8 ms). In contrast, there was no significant effect of the possessor, and no interactions. Sentences with a nominalization in subject position were also slower than the control condition, regardless of whether the nominalization contained a possessor (β = 34 ms; s.e. = 16 ms) or not (β = 30 ms; s.e. = 13 ms). There were no interactions at the matrix main verb.

At the spillover region (e.g., *in the kitchen*), there were no significant effects at the preposition head within the interference conditions, and no interactions. At the determiner inside the spillover region, sentences where the nominalization occurred in object position without a possessor were significantly slower than the control condition (β = 30 ms; s.e. = 10 ms). However, the maximal converging model contained only random intercepts by participants, and thus assumed no variation among participants for experimental factors. There were no other differences between the interference and control conditions at the spillover regions.

Embedded subject & VP regions There were significant effects of both complexity and the possessor at the head noun of the embedded subject. Sentences with a nominalization in subject position were reliably slower at this region than sentences with a simple intervener (β = 35 ms; s.e. = 12 ms)[5] The presence of a possessor also led to a slowdown at this region (β = 31 ms; s.e. = 12 ms). The maximal model contained no by-participant random slopes for the complexity manipulation, and thus assumed the effect to be invariant across participants. There was also no by-item random slopes for the interaction term. There were no significant differences between the interference conditions and the control condition at the determiner of the embedded sub-

[5]The presence of a possessor actually *facilitated* processing at the determiner of the embedded subject (β = -21 ms; s.e. = 8 ms). This difference is very likely spurious, however, since the possessor had not yet occurred at this point in the sentence.

ject. (No comparisons were made at the head noun since this position was empty for control sentences.)

There were no significant differences among the interference conditions at the embedded verb of the relative clause (region 13). Comparisons to control sentences, however, revealed a cost for overt subjects. Sentences with an overt complex intervener were reliably slower than control sentences (complex + poss: β = 73 ms; s.e. = 18 ms; complex no poss: β = 40 ms; s.e. = 18 ms). Simple interveners showed a similar slowdown relative to control sentences, though the differences were only trending significant (complex + poss: β = 34 ms; s.e. = 18 ms, t = 1.868; complex no poss: β = 35 ms; s.e. = 18 ms, t = 1.915). The model comparing control sentences to the interference conditions contained only random intercepts, by-participant and by-item, and so pooled the error terms within these groups.

At the determiner of the embedded object, all four interference conditions were reliably slower than the control condition. The greatest difference from control sentences was in possessor sentences with simple interveners (β = 34 ms; s.e. = 11 ms). The smallest difference was between no possessor sentences with simple interveners and the control condition (β = 27 ms; s.e. = 11 ms). At the head noun of the embedded object, sentences containing a nominalization in the embedded subject position were reliably faster than the control condition (Complex + poss: β = -75 ms; s.e. = 21 ms; Complex no poss: β = -55 ms; s.e. = 22 ms) and sentences containing a nominalization in object position, and no differences within the interference conditions. Note that the cost for overt embedded subjects shows up one word after the embedded verb, which is similar to the effects in the critical region.

Comprehension accuracy

Mean comprehension accuracy for each condition is shown in Table 3.2. Grand mean accuracy was 75%. Comprehension accuracy for the interference conditions was significantly lower than the control condition (all p's < .01). There were no significant differences in accuracy within the interference conditions.

	mean	Matrix S-V	Embedded S-V	final PP
Control	87% (3%)	84% (4%)	82% (5%)	94% (3%)
Simple	75% (3%)	79% (5%)	54% (6%)	91% (3%)
Complex	72% (3%)	82% (5%)	51% (6%)	82% (5%)
Simple +Poss	70% (3%)	74% (5%)	60% (6%)	76% (5%)
Complex +Poss	72% (2%)	74% (5%)	57% (6%)	85% (4%)
means	75% (1%)	79% (2%)	61% (3%)	86% (2%)

Table 3.2: Experiment 1 comprehension questions, mean accuracy (and standard error) by condition.

OSPAN working memory task

40 participants performed the OSPAN working memory task in a separate session. The dependent variable was the 'absolute' score of each participant, the sum of all correctly recalled character sets. Mean absolute score was 48 (s.d. = 17).

To evaluate the role of working memory span on interference effects, the OSPAN absolute score was included as a numeric predictor in the maximal models of RTs at the matrix main verb[6]. The OSPAN absolute score was centered prior to analysis, to reduce co-linearity with other predictors. We used the main verb model because this is where we first observed interference effects in the reading times.

This model showed no significant correlation between OSPAN absolute score and RTs (β = -0.33 , s.e. = 0.27; t = -1.217). The effect of intervener complexity remained significant (β = 28 ms; s.e. = 9 ms). The effect of OSPAN absolute score was trending significant in comprehension accuracy models fit to the interference conditions, suggesting that greater memory span was correlated with greater accuracy (β = 0.01 , s.e. = 0). Working memory span was not a significant predictor in models comparing interference conditions to control sentences.

3.4.4 Discussion

The main effect of intervener size suggests that a nominalization in subject, but not object, position leads to difficulty when the verb is initially processed. This is consistent with previous find-

[6]Since the OSPAN score was unique to each participant, there were no random intercepts or slopes associated with the OSPAN predictor (just the fixed effect).

ings (Van Dyke & Lewis, 2003; Van Dyke, 2007), since the verb is the point at which the subject must be retrieved. We also take this to be an interference effect, since it is restricted to configurations where the large intervener is structurally similar to the retrieval target, *i.e.* a subject. Furthermore, comprehension accuracy was reliably lower for experimental conditions – which all included overt subjects – compared to the control condition, which used raising predicates to obviate overt subjects. This is consistent with previous findings of interference from embedded, grammatically inappropriate, lexical subjects (Van Dyke & Lewis, 2003; Van Dyke, 2007). We extend previous results by using materials that did not vary in their depth of embedding. However, the difficulty at the verb was not due to the presence of a possessor, which had no reliable effects on the processing at the verb or subsequent regions. We take the lack of interference at the critical region for possessors to indicate that subject retrieval targets structural properties, and not thematic properties.

Why should complex embedded subjects be more difficult? Assuming the difficulty to arise at retrieval, one potential explanation is that the slowdown for larger interveners is an elaboration effect. Nominalizations are more complex than *det+noun* nominals in terms of argument structure, the nominalizing morphology, and length. While the materials of Experiment 1 did not make distinct predictions about these possibilities, similar effects of complexity have been observed in the processing of filler-gap and subject-verb dependencies, where increasing the complexity of a to-be-retrieved constituent facilitates the processing at the retrieval site (Hofmeister, 2011). Our results differ in that here complexity led to inhibition, not facilitation. One explanation for these contrasting profiles is that in our materials, and unlike Hofmeister (2011), we manipulated the complexity of an intervening element, not the target constituent. If we assume that increasing the complexity of a constituent increases its accessibility at retrieval, then increasing the complexity of the target will facilitate retrieval, while increasing the complexity of non-target constituents will facilitate misretrieval, and inhibit the retrieval of the target. We return to the issue of complexity in the 'General Discussion' section.

How do these results contribute to our understanding of the retrieval cues? If subjects are retrieved based on their thematic properties, *e.g.* [EXTERNAL-ARGUMENT] or [AGENT], then we would expect no distinctions between the subjects of a verb and the subjects of a nominal, contrary to the observed pattern. At the same time, the slowdown for complex embedded subjects, in conjunction with previous results (Van Dyke & Lewis, 2003; Van Dyke, 2007, a.o.), indicates that the retrieval cues do not uniquely identify one constituent as the subject, since a grammatically inappropriate constituent leads to difficulty at the retrieval site, but only when similar to the target. In other words, the retrieval structure at the verb is sufficient to distinguish between the subjects of nominals and the subjects of clauses, but cannot uniquely discriminate between the matrix subject and a structurally similar embedded subject.

3.5 Diagnosing subjects

In Experiment 1, we saw that additional subjects do lead to difficulty at the verb, when complex, even when the number of clauses is held constant. This removes a confound in the constructions examined by Van Dyke and colleagues (§3.3.4). However, there was no interference from additional thematic subjects, indicating that the cues at the verb characterize subjects structurally, and not in terms of thematic properties like *external argument of a predicate*. In other words, the domain of the intervener matters: the retrieval cues are sufficient to distinguish between the subject of a clause and the subject of a nominal.

These patterns are consistent with retrieval cues that target either case or position. The target subject was always the NOMINATIVE subject of a finite verb in SPEC-T. While a complex nominalization in the embedded subject position will match both of these cues, the possessor was a GENITIVE subject of a nominal, occupying SPEC-D. We evaluate these cues in Experiments 2 and 3, but before doing so I consider some evidence that similarity in terms of case and position matter. In reviewing this evidence, we will see that case and position similarity matter, but the evidence is somewhat equivocal about whether this is due to retrieval interference engendered by similarity between the intervener and the cues at the verb.

3.5.1 Case & position

Babyonyshev & Gibson (1999) examined the effects of syntactic similarity in Japanese sentences. Using an offline questionnaire, participants were asked to rate the difficulty of sentences such as those shown below in (49). Since Japanese is a head-final language, sequential complement clauses serve to place multiple subjects in pre-verbal position. Similarity was manipulated by varying the morphosyntactic marker of these subjects: *-ga* for NOMINATIVE structural subjects, and *-wa* for topics.

(49) a. *Wakai kyooju -ga [TA -ga [gakusei -ga*
young professor -NOM [$_{CP}$ teaching assistant -NOM [$_{CP}$ students -NOM
konransita to] sengensita to] utagatta
panicked that] announced that] doubted
'The young professor doubted that the teaching assistant announced that the students panicked.'

b. *Eegakantoku -wa [purodyusaa -ga [kireina joyuu -ga koronda to]*
film director -TOP [$_{CP}$ producer -NOM [$_{CP}$ pretty actress -NOM fell that]
itta to] omotteiru
said that] thinks
'As for the film director, he thinks that the producer said that the pretty actress fell.'

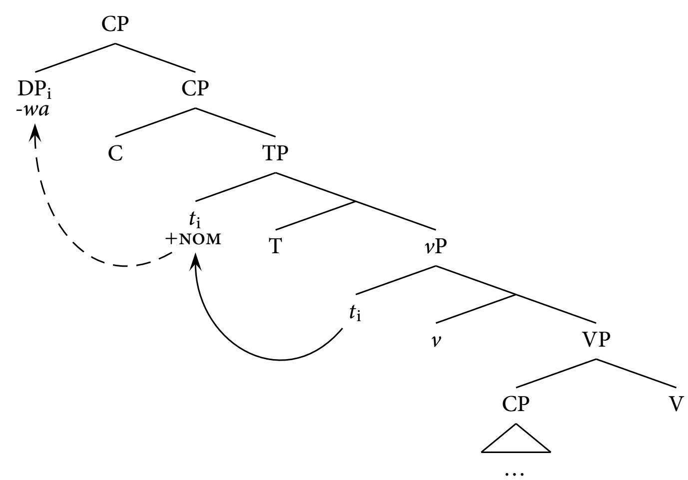
Babyonyshev & Gibson found that sentences containing three nominative *-ga* subjects (49) were perceived as significantly more difficult (rated lower) than sentences such as (49 b), with two nominative subjects and a *-wa* topic.

While the offline measures cannot provide direct evidence of interference, the results are consistent with an cue-driven retrieval approach. If subjects are retrieved at the verb using syntactic cues like case features, i.e. [NOMINATIVE], then such cues will return three matches in (49a) but only two in (49b)[7]. In this way, these results suggest that case similarity between subjects leads to interference effects.

However, the interpretation of this as case interference depends on our analysis of topics. Contemporary transformational approaches would analyze the topic as receiving NOMINATIVE

[7]This is not a necessary conclusion, since *-ga* probably serves several functions. However, the prediction still holds as long as the retrieval cues target some property of *-ga* and not *-wa*.

case in the (matrix) subject position[8], then subsequently undergoing A'-movement to the topic position at the clause periphery (Chomsky, 1995). In other words, both *-ga* and *-wa* constituents are marked [NOM], but differ in their structural position, with *-wa* markers occupying a higher position within the CP layer:



(50)

If we assume this view, then the results of Babyonyshev & Gibson (1999) would suggest that the retrieval cues for subjects are based on structural position, and not (abstract) case. In any case, these results show that the syntactic similarity of subjects, however defined, matters for processing, even when the number of functional subjects and depth of embedding is held constant. Less clear is what the relevant syntactic dimensions of similarity are.

[8]The functions of *-wa*, like *-ga*, are complex. The marker serves at least two functions, what Kuno (1973) calls 'thematic' and 'contrastive'. In the latter case it serves as a 'topic' marker. Importantly, *-wa* can also mark objects, which may then optionally move to a clause initial position. In the interest of focus and scope, I will not attempt to summarize the syntax of these constructions here. Note, though, that the possibility of *-wa* marked objects may make these sentences (locally) ambiguous. As far as I can tell, in Babyonyshev & Gibson's (1999) materials the *-wa* consituent was always a subject, and clause initial. I thus assume that the characterization above is on the right track, if overly simple. Regardless of how we analyze these elements, the results of Babyonyshev & Gibson (1999) are consistent with an interference explanation based on subject similarity. The dimensions of similarity, though, remain unclear.

3.5.2 Subject similarity matters

Further evidence that subjects are identified based on their structural position was found in Japanese by Lewis & Nakayama (2002), who used both acceptability-ratings and self-paced reading to examine sentences like those shown in (51). Lewis & Nakayama (2002) reasoned that if interference is a source of difficulty in multiple embeddings, then difficulty will increase as the number of nouns that must be stored before encountering the verb increases, so that 3 NPs (51a,b) will be easier than 4 NPs (51c,d). More importantly for the present purposes, Lewis & Nakayama (2002) also manipulated the positional similarity of the subject NPs – the number of string-adjacent nouns indexed to the same syntactic position – by varying the word-final marker on the subject NP: *-wa* for topic subjects, and *-ga* for structural subjects. If retrieval of a subject at the verb is guided by the structural position of the target, then retrieval will be more difficult when two identically marked subjects are adjacent (51b,d) than when they are separated by another NP (51a,c), because in the former condition the target can only be identified by its relative position.

(51) Lewis & Nakayama (2002) materials:

a. *Ani-ga sensei-ni onna-no-ko-ga asondeiru-to renrakushita.*
elder.brother-NOM teacher-DAT girl-NOM playing-that notified
'My older brother notified the teacher that the girl was playing.'

b. *Haisha-ga daitooryoo-ga tsuuyaku-o yonda-to oboeteita.*
dentist-NOM president-NOM interpreter-ACC called-that remembered
'The dentist remembered that the President called the interpreter.'

c. *Kyooju-ga shachoo-ni daihyoo-ga kookoosei-o*
professor-NOM president-DAT representative-NOM H.S.-student-ACC
shinsasuru-to yakusokushita
examine-that promised
'The professor promised the company president that the representative would examine the high school student.'

d. *Seito-ga kooshi-ga repootaa-ni sakka-o shookaishita-to kizuita*
student-NOM lecturer-NOM reporter-DAT author-ACC introduced-that noticed
'The student noticed that the lecturer introduced the author to the reporter.'

In both the reading and ratings tasks, Lewis & Nakayama (2002) observed that difficulty increased as a function of both the number of NPs that must be held in memory before the verb and their positional similarity. Maintaining four NPs (51c,d) was more difficult than three NPs (51a,b), and discriminating between adjacent subjects (51b,d) was more difficult than non-adjacent subjects (51a,c). While there were no interactions in the online measures, in the offline rating task participants reported the easiest sentences were those with a ditransitive matrix verb and intransitive embedded verb (51a).

Lewis & Nakayama (2002) note that subjects cannot be identified solely on morpho-syntactic information encoded in the *-wa/-ga* markers alone, since separating these NPs by another NP modulates the difficulty. Overall, then, whatever the precise source of difficulty is in the sentences of Lewis & Nakayama (2002), processing at a verb is more difficult when it requires discriminating between multiple subjects based solely on their relative structural position.

3.5.3 *Case similarity matters*

The previous experiments were taken as evidence that case similarity between subjects engenders interference. I suggested that these results could also be analyzed as reflecting similarity in terms of syntactic position. Other evidence suggests that similarity in case features alone causes constituents to interfere. One piece of evidence for case interference comes from the processing of Russian complement clauses conducted by Fedorenko, Babyonyshev, & Gibson (2004). They examined the effect of case similarity on the reading times at a verb in configurations like those schematized below in (52). The availability of scrambling in Russian allows multiple object DPs to occur in a pre-verbal position. Assuming that integrating the verb requires retrieval of its object, this retrieval is potentially susceptible to interference from the additional object DP inside the participle phrase:

(52) [Participle EmbObj] MatrixObj [Verb] MatrixSubj PP_1 PP_2

One potential retrieval cue for the object is case. Importantly, though, the case of Russian nominals depends on their noun class but case syncretism means that nominals may differ in abstract case while bearing phonologically identical case markers[9]. The accusative and dative case forms of Russian Class I and Class II nominals are shown below in (53), borrowed from Fedorenko, Babyonyshev, & Gibson (2004). Note that the masculine dative and feminine accusative markers are both *-u*:

(53) Russian Class I/II case-markers:

	MASC	FEM
ACC	*-a*	*-u*
DAT	*-u*	*-e*

In a self-paced reading experiment, Fedorenko, Babyonyshev, & Gibson (2004) leveraged the availability of scrambling and case syncretism to examine the effect of case similarity on the processing at the verb, *angered*, in sentences such as (54). In a 2 × 2 design, they manipulated the abstract and morphological case properties of the initial object DP, *violinist*. The second DP was always feminine accusative, so that DP_1 could match in both abstract and morphological case (54a), fully mismatch (54d), or match in only one of their abstract/morphological case features(54b–c):

(54) a. *Uvažavšuju skripačk-**u** pianistk-**u** razozlil dirižer iz*
[respecting violinist-FEM.ACC] pianist-FEM.ACC angered conductor-NOM from
izvestnoj konservatorii posle generalnoj repetitsii.
famous conservatory after final rehearsal
"The conductor from a famous conservatory angered the pianist who respected the violinist after the final rehearsal."

b. …*skripač-**a** pianistk-**u**…* *matching abstract case*
…violinist-MASC.ACC pianist-FEM.ACC

c. …*skripač-**u** pianistk-**u**…* *matching morphological case*
…violinist-MASC.DAT pianist-FEM.ACC

d. …*skripačk-**e** pianistk-**u**…* *no match*
…violinist-FEM.DAT pianist-FEM.ACC

Fedorenko *et al.* found that the processing at the verb was significantly more difficult for sentences where the two object DPs matched in both abstract and morphological case (54a).

[9] Alternatively, we can reject the abstract case characterization and instead view the contrast as arising from DPs bearing the same morphological case features but differing in their phonological spell-out.

There were no significant differences between no match or partially-matching conditions (54b–d). Thus, case similarity leads to difficulty, but only when *both* abstract and morphological case match.

However, Logačev & Vasishth (2012) point out that the experimental manipulation of abstract and morphological case also involves an uncontrolled manipulation of the gender properties of the initial DP. This is a worry because in the sentences where the DPs match in either abstract or morphological case (54b–c), they also mismatch in gender. It may be that this dissmilarity is enough to mask any interference from case properties alone.

3.5.4 Case attraction effects

I turn now to a series of studies on the processing of German sentences which suggests that similarity in terms of case properties alone engenders interference, independent of syntactic position. Here, the evidence indicates that abstract case alone is sufficient to engender interference, but the behavioral signature of this effect is facilitation and not inhibition. The evidence comes from so-called 'case attraction' effects in German, first documented in self-paced reading (Bader, 1994; Schlesewsky, 1996). In case attraction configurations, a case-ambiguous DP is erroneously perceived as bearing the case of a coreferent but syntactically distinct DP.

Importantly for us, case attraction has been observed in the processing of complex subjects. For example, Bader (1997) asked German speakers to rate the acceptability of active voice sentences (55) containing a subject-attached relative clause. Proper names like *Maria* are case-ambiguous between nominative, accusative, and dative case. The case of this DP is disambiguated to NOMINATIVE by the sentence-final auxiliary *hat*, yielding a subject-object (active) word order. Until this point, though, the sentences are ambiguous. The crucial manipulation varied the case of the relative pronoun, which was either unambiguously dative *der* (55a) or ambiguously nominative/accusative *die* (55b):

(55) a. …, *daß Maria, der ich gestern begegnet bin, eine Postkarte geschickt hat.*
that Maria, who.DAT I yesterday met am, a postcard sent has
"…that Maria, who I met yesterday, sent a postcard."

b. …, *daß Maria, die ich gestern gesehen habe, eine Postkarte geschickt hat.*
that Maria, who.NOM/ACC I yesterday seen have a postcard sent has
"…that Maria, who I saw yesterday, sent a postcard."

Bader (1997) found that sentences with a dative relative pronoun (55a) were erroneously judged as unacceptable – and took longer to judge – than comparable sentences with a nom/acc relative pronoun (55b). In contrast, sentences with a nom/acc relative pronoun were not significantly different than the same sentences without a relative clause. In other words, the inclusion of a relative clause led to lower accuracy on the acceptability task, but only when that relative clause contained a dative pronoun, and not a nom/acc relative pronoun. Bader concludes that the subject is perceived as bearing DAT case, which conflicts with the case assigned by the final verb, leading to difficulty.

Bader (1997) also compared the active sentences in (55) to passive variants (56), where the sentence-final auxiliary *wurde* disambiguates the initial DP to accusative case, yielding an object-subject word-order. Unlike their active counterparts, the passive sentences showed a facilitation effect in recognition times for the dative pronoun, relative to nom/acc versions, though the effect was only marginal.

(56) a. …, *daß Maria, der ich gestern begegnet bin, eine Postkarte geschickt wurde.*
that Maria, who.DAT I yesterday met am, a postcard sent was
"…that a postcard was sent to Maria, who I met yesterday."

b. …, *daß Maria, die ich gestern gesehen habe, eine Postkarte geschickt wurde.*
that Maria, who.NOM/ACC I yesterday seen have a postcard sent was
"…that a postcard was sent to Maria, who I saw yesterday."

In summary, Bader (1997) found that a dative relative pronoun that is co-referent to a case ambiguous DP led to difficulty when the DP is disambiguated to nominative, but eases processing when the DP is disambiguated towards accusative. By analogy to agreement attraction, Bader suggests that the case feature of a co-referent DP can be 'attracted' to a case-ambiguous DP. Furthermore, while German word order allows the clause to start with any type of DP, not necessarily the subject, there is a general subject-object preference (see Hemforth & Konieczny, 2000, and references cited there). In the sentences above, the dative case of the relative pronoun is attracted to the head noun of the relative clause, with the result that the entire matrix subject DP is perceived as bearing dative case. A dative DP gives the illusion of a OS word-order, leading to a garden path effect when this expectation is disconfirmed by the S-final auxiliary. In the passive sentences, the dative relative pronoun again causes the case-ambiguous DP to bear dative case, but the S-final auxiliary confirms this expectation, facilitating processing.

Bader (1997) relates case attraction to agreement attraction, and indeed there are a number of similarities, both empirically and analytically. For one, as suggested above, case attraction has been claimed to show a markedness asymmetry: dative case on a relative pronoun can override nominative or accusative, but not vice versa (Bader, Meng, & Bayer, 2000; Bayer, Bader, & Meng, 2001). This effect obtains even when the head noun – the attractor – is not morphologically compatible with dative case (Bader, Meng, & Bayer, 2000).

Furthermore, while case attraction is most commonly documented in grammatical sentences, there is some evidence that case attraction can cause an ungrammatical sentence to be perceived as grammatical, at least on some proportion of trials. This 'illusion of grammaticality' is a hallmark of agreement attraction. In a speeded-acceptability experiment, Bader, Meng, & Bayer (2000) found that ungrammatical sentences were incorrectly judged as acceptable significantly more often when they contained a case-ambiguous DP (57) than when they didn't (58; see also Meng & Bader, 2000). The sentences in (57–58) all contain a case conflict between the case assigned (⇝) by the verb and the case of the initial DP of the embedded clause. The S-final auxiliary *hat* assigns nominative case, so that the (a) sentences are active; *wurde* assigns dative case, so that the (b) sentences are passive. Importantly, only feminine definite determiners are ambiguous (57; *der* = dat/gen, *die* = nom/acc). Masculine definite determiners (58) are unambiguous:

(57) a. * *Ich denke, daß der Chefin das Buch geliefert hat.*
I think that the.F.DAT boss the book delivered has⇝ NOM

b. * *Ich denke, daß die Chefin das Buch geliefert wurde.*
I think that the.F.NOM/ACC boss the book delivered was⇝ DAT

(58) a. * *Ich denke, daß dem Chef das Buch geliefert hat.*
I think that the.M.DAT boss the book delivered has⇝ NOM

b. * *Ich denke, daß der Chef das Buch geliefert wurde.*
I think that the.M.NOM/ACC boss the book delivered was⇝ DAT

Continuing with the similarities to agreement attraction, the locus of case attraction effects was initially taken to be defective grammatical representations, either because the case feature of the relative pronoun may percolate upwards to a c-commanding head noun (Bader & Meng, 1999; Bader, Meng, & Bayer, 2000; Bayer, Bader, & Meng, 2001) or because parsing principles attempt to unify the features of co-referent constituents (Fanselow et al., 1999). Recall that the argument against faulty representations as an explanation for agreement attraction was twofold: it cannot explain the grammaticality asymmetry – there is no 'illusion of ungrammaticality' – and attraction effects occur 'downward' when the distractor element both precedes and c-commands the agreement controller (Wagers, 2008; Wagers, E. F. Lau, & C. Phillips, 2009).

Similarly, case attraction effects are not dependent on a particular linearization of the co-referent DPs. In a self-paced reading experiment, Fanselow et al. (1999) examined sentences such as (59), manipulating the case matrix between a matrix DP (*die Frau*) and the relative pronoun of a subject-attached relative clause:

(59) a. *Das ist die Frau, die glücklicherweise die Soldaten besucht hat, obwohl*
that is the woman.NOM who.NOM fortunately the soldiers.PL visited has.SG although

b. *Das ist die Frau, die glücklicherweise die Soldaten besucht haben, obwohl*
that is the woman.NOM who.ACC fortunately the soldiers.PL visited have.PL although

c. *der Soldat überrascht die Frau, die glücklicherweise die Männer besucht hat,…*
the soldier surprises the woman.ACC who.NOM fortunately the men.PL visited has.SG

d. *der Soldat überrascht die Frau, die glücklicherweise die Männer*
the soldier surprises the woman.ACC who.ACC fortunately the men.PL

besucht haben,
visited have.PL

They found that the processing at the embedded auxiliary was faster when the case of the matrix DP matched that of the relative pronoun (59a,d), than when it didn't (59b,c). Importantly, the processing at the embedded auxiliary is modulated by the case of the c-commanding head noun, so that case attraction is 'downward', contra the predictions of Bader and colleagues. Moreover, none of the sentences in (59) involve dative case, the marked option. The results of Fanselow et al. also provide evidence against the view that case attraction is driven by marked cases overriding unmarked cases. Rather, it seems that all that is necessary for case-attraction is for two co-referent DPs to match in their case properties.

Overall, then, it seems that the processing at a verb is facilitated in configurations where a morphologically ambiguous DP bears the same case as a co-referent DP. That case similarity modulated processing at a verb suggests that the case properties of DPs lead to interference. However, sometimes case similarity inhibits processing (Fedorenko, Babyonyshev, & Gibson, 2004), but in case attraction configurations it actually seems to facilitate processing. We might wonder if the form of the interference effect is due to co-reference in the case attraction effects. Relatedly, it remains unclear what dimension of case – morphological (ambiguity) or abstract – leads to interference. In a related study, Logačev & Vasishth (2012) report the results of two eye-tracking studies designed to test whether co-reference is a necessary condition for case attraction, and whether abstract case alone can drive interference from over-lapping case features. They also provide a model of interference effects that captures both inhibitory and facilitatory interference.

In their first experiment, Logačev & Vasishth used partitive constructions to establish coreference between the pronoun, *ihnen*, and a partitive construction, *von ihnen*, which referred to a sub-part of its conjoined antecedent. In a 2 × 2 design, they manipulated the case of the pronoun and the partitive constituent, both of which varied between ACCUSATIVE and DATIVE. The case ambiguity of the conjoined DP is resolved at the second verb, *greet*.

(60) a. *Dass er [Leo und Tim]$_i$, obwohl er [einen [von ihnen$_i$]]*
that he [Leo and Tim] .ACC/DAT, although he [one.ACC [of them.DAT]]

verabscheut, grüßen sollte …
despises, greet should …

b. *Dass er [Leo und Tim]$_i$, obwohl er [einen [von ihnen$_i$]]*
that he [Leo and Tim] .ACC/DAT, although he [one.ACC [of them.DAT]]
verabscheut, glauben sollte …
despises, believe should …

c. *Dass er [Leo und Tim]$_i$, obwohl er [einem [von ihnen$_i$]]*
that he [Leo and Tim] .ACC/DAT, although he [one.DAT [of them.DAT]]
verabscheut, grüßen sollte …
despises, greet should …

d. *Dass er [Leo und Tim]$_i$, obwohl er [einem [von ihnen$_i$]]*
that he [Leo and Tim] .ACC/DAT, although he [one.DAT [of them.DAT]]
verabscheut, glauben sollte …
despises, believe should …

At the disambiguating verb, eye-tracking measures showed a slowdown in go-past times and more regression out when the case assigned by V1 did not match that of V2 (a,c), an interaction effect[10]. V&L conclude that strict co-reference is not necessary for case attraction, since the partative construction refers only to a subset of the case-ambiguous DP, and not the DP itself.

In their second experiment, V&L examined whether abstract case was enough to drive case attraction, or whether morphological case was a necessary condition. In an eye-tracking experiment, they examined the effect of case match between a matrix argument and an embedded relative pronoun in sentences like (61).

(61) *Der Prinz respektiert die Künstler* …
the prince.NOM respects the artists.ACC

a. *die die Königin erst kürzlich getroffen haben* …
who.NOM/ACC the queen.ACC/NOM only recently met Aux.PL

b. *die die Königin erst kürzlich getroffen hat* …
who.ACC/NOM the queen.NOM/ACC only recently met Aux.SG

… *für ihr herausragendes Talent.*
for their outstanding talent.

[10] Logačev & Vasishth (2012) also collected self-paced reading times on these sentences (though they were modified to match for length before the critical region). The self-paced reading times results were similar, except that the interaction effect showed up at the regions immediately following the critical region, rather than the verb itself.

'The prince respects the artists, who the queen only recently visited/who only recently visited the queen, for their outstanding talent.'

Here a relative clause is attached to the matrix plural object, *the artists*, and contains the case-ambiguous relative pronoun *die*, which may be either nominative or accusative. Logačev & Vasishth assume that nominative case on the first DP, *the prince*, biases the matrix object to accusative case. The case of the relative pronoun is disambiguated by the number agreement of the RC-final auxiliary verb. The plural auxiliary *haben* yields a subject (gap) before object word order, so that the relative pronoun bears nominative. The singular auxiliary *hat* reverses this order, and disambiguates the relative pronoun to accusative. Since the antecedent is always accusative, if case attraction depends on case match between the RC pronoun and its antecedent, then (61a) will be more difficult than (61b).

Difficulty in (61a) might also reflect an expectation for an object before subject word-order being disconfirmed by nominative case on the relative pronoun. To make sure that evidence of difficulty is due to case match and not word-order preferences, Logačev & Vasishth also examined passive sentences such as (62). Here, the relative clause is also attached to *the artists*, but the passive means this argument bears nominative case, and not accusative. Thus, if case attraction is driven by case match between the pronoun and its antecedent, mismatching conditions (61b, 62b) will be more difficult than matching conditions (61a, (62a). On the other hand, if processing at the verb depends on word-order preferences determined by the case of the matrix RC head, then (61a) will be more difficult than (61b), but (62b) will be more difficult than (61a).

(62) *Vom Prinzen respektiert werden die Künstler …*
by the.prince respected Aux.Pass the artists.NOM

a.	*die*	*die*	*Königin*	*erst*	*kürzlich*	*getroffen*	*haben*	…
	who.NOM/ACC	the	queen.ACC/NOM	only	recently	met	Aux.PL	
b.	*die*	*die*	*Königin*	*erst*	*kürzlich*	*getroffen*	*hat*	…
	who.ACC/NOM	the	queen.NOM/ACC	only	recently	met	Aux.SG	

… *für herausragendes Talent.*

for their outstanding talent.

'The artists, who the queen only recently visited/who only recently visited the queen, are respected by the prince for their outstanding talent.'

Logačev & Vasishth showed case match led to facilitation in the processing at the verb. Sentences where the case of the relative pronoun mismatched that of its antecedent (61b,62a) resulted in longer go-past times at the auxiliary, and more regressions out at both the auxiliary and main verb. Mismatching conditions also showed longer regressive re-fixation times at the main verb. They conclude that the interactions at the auxiliary + verb regions indicate that a general SO word-order preference is weakened or even reversed when the antecedent is accusative. More generally, only abstract case was varied, while other properties like the thematic role of the head noun (*artists*) were held constant, so the facilitation effect for case match indicates that abstract case alone is sufficient for case attraction.

Conflicting bindings

Up to this point we have been concerned with evidence for case interference. To the extent that case similarity between the target of retrieval and another constituent leads to interference, we have evidence for a case cue. The results of Fedorenko, Babyonyshev, & Gibson (2004) in Russian participle clauses implicated interference, but only when the target and distractor constituent matched in both morphological and abstract case. When we consider the effects of case attraction in German, though, the results of Logačev & Vasishth (2012) indicate that distractor constituents matching in abstract case alone engender interference.

Logačev & Vasishth, 2012 argue that not all interference effects in language comprehension are cue-driven. Some interference effects are driven by dimensions of similarity that are not plausible retrieval cues—they are not selected by the verb[11]. Logačev & Vasishth argue that the results

[11] Logačev & Vasishth give the results of Van Dyke & McElree (2006) as an example of cue-driven interference (see §2.2). In that experiment, a list of memorized nouns led to interference, but only when they were plausible arguments of the verb: *sink…fixed* versus # *sink…sailed*. Van Dyke & McElree conclude that the 'meaning of NPs' informs the verb's retrieval cues. It is unclear to me that these properties are selected by the verb, but if they were, what would the cues be? +*sailable*? I see little explanatory value in such a cue. This is only a worry, though, if we follow Logačev & Vasishth, 2012 in defining non-cue-driven interference as that that occurs along unselected dimensions.

of Fedorenko, Babyonyshev, & Gibson (2004) are not cue-driven. To them, abstract case is a plausible retrieval cue, but morphological case is not. Therefore, while Fedorenko, Babyonyshev, & Gibson's (2004) results indicate that case similarity matters for processing, the interference effect is not due to multiple DPs matching the retrieval cues. Rather, the effect of case similarity in Russian is due to interference that arises from storing similar DPs in memory. Previous researchers generally refer to these effects as *encoding* interference.

Outside of the argument from selectional properties, Logačev & Vasishth give one more argument in favor of non-cue-driven interference. They report the results of an unpublished eye-tracking experiment that indicated that interference does not increase monotonically with similarity. Using German sentences such as those in (63), they manipulated the gender and noun-type (occupation or nationality) of two DPs. Note that since neither property is selected by the verb, neither would qualify as a retrieval cue.

(63) a. *Die Malerin hat die Snacks, die die Bildhauerin …*
the painter.FEM Aux the snacks, which the sculptor.FEM

b. *Die Maler hat die Snacks, die die Bildhauerin …*
the painter.MASC Aux the snacks, which the sculptor.FEM

c. *Die Kroatin hat die Snacks, die die Bildhauerin …*
the Croatian.FEM Aux the snacks, which the sculptor.FEM

d. *Die Kroate hat die Snacks, die die Bildhauerin …*
the Croatian.MASC Aux the snacks, which the sculptor.FEM

… neulich mitbrachte nur widerwillig serviert.
recently brought only reluctantly served.

'The painter/Croatian only reluctantly served the snacks that the sculptor bought recently.'

Logačev & Vasishth found that conditions where one feature mismatched (63b,c) were reliably slower than conditions where both matched (63a) or neither matched (63d). They interpret this effect as inhibition for a mismatching gender feature if noun-type also mismatches, but a facilitation if noun-type does match. Logačev & Vasishth argue that non-cue-driven interference reflects difficulty in re-activating constituents in memory when they match in only some features. They call this form of interference *Conflicting Bindings Retrieval Interference* (CBRI).

The CBRI model of Logačev & Vasishth (2012) is meant to subsume both cue-driven and non-cue-driven interference effects under the same mechanism. In particular, effects previously viewed as encoding interference are recast as a variant of retrieval interference. The upshot of this model is that it captures both inhibitory and facilitory interference effects.

Logačev & Vasishth assume that constituent representations consist of feature bundles, as we have up to this point. However, based on findings in visual processing (e.g. Hommel, 1998), Logačev & Vasishth assume that each feature of an object is linked to all of its other features via pairwise bindings. Thus, a constituent representation in the CBRI model is the totality of these pairwise bindings. The source of interference effects are those features that are shared between the target of retrieval and other constituents, but linked to other, mismatching properties. More precisely, these *conflicting* bindings are defined as below:

Conflicting bindings $=_{Def^N}$ Binding B_1 *conflicts* with binding B_2 if one feature is common to both bindings, and B_1 links this feature to another feature F_1, while B_2 links it to F_2, and F_1 and F_2 are distinct. (Logačev & Vasishth, 2012, p.192)

When two objects match in some feature, all links to other matching features will be non-conflicting. Changing this feature, then, will cause all other links to matching features to become conflicting. These conflicting bindings lead to an interference effect. Since constituents are (dis)similar by virtue of their matching features, increasing the similarity of two dissimilar objects – increasing their conflicting bindings – will lead to an inhibition in processing. On the other hand, increasing the similarity of two already similar elements will yield a reduction in the number of conflicting bindings, thereby facilitating processing. In this way, the CBRI model captures both inhibitory and facilitory interference.

What makes two constituents count as 'similar'? If we had a theory of retrieval cues, the answer would be straightforward. Such a theory, though, does not exist. Logačev & Vasishth note that the featural content of encodings might be viewed as a 'free-parameter' in the CBRI model. They offer a formal proof that the model still makes qualitative predictions about the relative differences between encodings. The effect of this difference, though, turns on whether the two constituents are similar or dissimilar. Logačev & Vasishth seem to take similarity to mean 'co-reference', in that DPs linked to distinct referents are always assumed to be dissimilar.

The implications of the CBRI model are that the form of the interference effect for intervening subjects may arise as either inhibitory – as in Experiment 1 – or facilitatory. The difference turns on whether the experimental manipulations increase similarity between two elements that are independently similar or dissimilar. In Experiments 2–3, we continue our examinations of interference from intervening subjects, while extending the experimentally manipulated dimensions of similarity to case and position. I discuss the predictions of the CBRI account for these materials in the next section. For now, note that the effect of similarity between the retrieval target and the intervener, whether case or position, may vary depending on the overall background similarity between these constituents that is not subject to experimental manipulation (e.g. reference).

3.6 Experiment 2: Case & position cues

3.6.1 Design & predictions

The results of Experiment 1 implicated phrase-structure cues in subject retrieval. We found that increasing the complexity of an intervening nominal only mattered when it was in subject position. In contrast, we found no comparable difficulty for the presence of a possessor subject inside the nominalization, despite its being thematically similar to the target matrix subject. These patterns suggest that the retrieval cues available at a verb characterize subjects in terms of phrase-structure, and not their thematic properties. In English, possessors can be differentiated from subjects of verbs by two possible cues: either case, [GENITIVE] versus [NOMINATIVE], or specific structural position, [SPEC-D] versus [SPEC-T], respectively.

In Experiment 2, we extend the investigation of subject interference to a different aspect of similarity, namely the abstract case features of subjects. In English, the patterns of Experiment 1 are consistent with retrieval cues that target either syntactic position or case, since the subject of a clause would bear the features [SPEC-T, NOMINATIVE] while the possessor would bear [SPEC-D, GENITIVE]. The current experiment manipulated case similarity between a matrix and embedded subject by varying the syntactic properties of the verb which embedded the clause containing the intervening subject. Using structures such as those shown in (64), we compared the nominative subject of a finite clause ('S-complement' sentences; 64a) to the accusative subject of a non-finite clause (64b) using so-called 'exceptional case-marking' (ECM) verbs like *believe*.

(64) a. The explorer who believed that [$_{\text{TP}}$ *the monster* was prowling the ruins] was insane…
b. The explorer who believed [$_{\text{TP}}$ *the monster* to be prowling the ruins] was insane…

Crucially, while the subject of an ECM complement, like *the monster* in (64b), differs from the nominative embedded subject of (64a) in its case feature, both elements occupy the structural subject position of the embedded clause (assumed to be SPEC-T). Thus, the embedded subjects of both structures in (64) are similar in terms of their syntactic position, but the embedded subject of the ECM structure is dissimilar to the matrix subject in its case feature.

To further disassociate case and position, we also compared the subject interveners of S-complement and ECM structures (64) to object control structures, such as that shown below in (65). As in the ECM sentences, the potentially interfering element bears accusative case, but occupies the matrix *object* position, which is linked to the embedded subject position (PRO) via the Control dependency. The intervening element in the object control structures is thus maximally dissimilar to the target matrix subject: they are dissimilar in both case and position.

(65) The explorer who persuaded *the monster* [$_{\text{TP}}$ to prowl the ruins] was insane…

Experiment 2 was also designed to extend the complexity effect of Experiment 1 to noun-noun compounds like *the ancient alien monster*, reasoning that the initial noun will be erroneously parsed as the head of the embedded subject, and reanalyzed as a modifier element when the second member of the compound is encountered. Reanalysis should thus require elaboration of the intervener, resulting in a stronger encoding and potentially increasing the probability of its (mis)retrieval. The complexity manipulation of Experiment 2 is similar to that of Hofmeister (2011), who found that processing at the verb was easier when its subject was a modified-NP like *an alleged Venezuelan communist* compared to a simpler nominal like *a communist*. We extend this finding by manipulating the complexity of an intervening subject, not the retrieval target. Thus, if complexity facilitates retrieval, then we expect complex interveners to increase difficulty at the matrix verb. Furthermore, noun-noun compounds do not contain the thematic structure of the nominalizations used in Experiment 1, so that replicating the complexity effect in Experiment 2 would show that the effect was not due to resolving additional thematic links.

The resulting experimental design was 3 × 2, with factors for embedded *structure* (S-complement, ECM, Object Control) and intervener *complexity* (noun-noun compound, *det* + *noun*). The potential dimensions of similarity are schematized below:

(66) target-NP$_{\text{NOM.SUBJ}}$… [a–c] …Aux$_{\text{NOM?SUBJ?}}$+MainV

a. *S-comp:* believes [$_{\text{CP}}$ (that) [$_{\text{TP}}$ intervener$_{\text{NOM.SUBJ}}$ was V…]]

b. *ECM:* believes [$_{\text{TP}}$ intervener$_{\text{ACC.SUBJ}}$ to V…]

c. *Object Control:* persuaded intervener$_{\text{ACC.OBJ}}$ [$_{\text{CP}}$ [$_{\text{TP}}$ PRO to V…]]

As in Experiment 1, the critical region is the matrix verb region, where the subject is retrieved. If the target of the verb's retrieval is identified by structural position, then the intervening element of both S-comp and ECM conditions will be similar to the target matrix subject. We thus expect that both of the subject intervener conditions will engender more interference than the Object Control sentences. In contrast, if the verb's cues target Case properties, like NOMINATIVE, then only S-complement interveners will match the case retrieval cue—both ECM and Object Control conditions contained ACCUSATIVE interveners. If the verb uses retrieval cues for *both* case and structural position, then S-comp interveners will match both cues, ECM sentences will match only the position cue, and Object Control will match neither cue. We thus expect a cline of difficulty, with the greatest interference in S-comp conditions (full match), followed by ECM conditions (partial match), and Object Control conditions the easiest (mismatch).

For the complexity manipulation, the predictions are twofold. If the subject-size effect of Experiment 1 was due to the additional thematic links of the event nominalization, then we expect this effect to be absent in Experiment 2, since the compounds lack such propositional structure. Furthermore, if the effect of complexity is modulated by the degree of similarity to the target of retrieval, then we expect greater difficulty for complex interveners, but this difficulty should be greatest in the S-complement sentences due to greater similarity with the target matrix subject.

Working memory task

As in Experiment 1, the reading time results were correlated with participants' performance on a working memory task. Here, however, we used the N-Back task, reasoning that the lack of correlation between the OSPAN and RT results of Experiment 1 might suggest that OSPAN measures do not tap the same memory resources as reading comprehension tasks.

In the N-Back task, participants are serially presented with a sequence of symbols. The presentation of these symbols is interrupted with a prompt presenting a probe letter and asking participants to judge whether the probe matches the letter that occurred one, two, or three positions previous. Initial probes target the previous letter (1-back), but as participants correctly answer the prompt the distance is increased to two positions previous (2-Back), then finally three positions previous (3-Back).

3.6.2 *Method*

Participants

Forty-two participants from the University of California, Santa Cruz community received $15 for participation in the experiment. All participants gave informed consent.

Materials

A Sample set of materials for Experiment 2 is shown in Table 3.3. Experimental materials consisted of 36 item sets arranged in a 2 × 3 factorial design, manipulating intervener size (complex/simple) and the properties of the verb that embedded the clause containing the intervener (ECM, S-complement, Object Control). The size manipulation compared simple *det+noun* subjects, *e.g. the monster*, to noun-noun compounds with an adjective modifier, *e.g. the ancient alien monster*. The critical region was the matrix auxiliary+verb sequence, *e.g. was yelling*.

As in Experiment 1, we used sentence-initial adverbial phrases to eliminate potential list-initial facilitation and sentence-final prepositional phrases to avoid wrap-up effects. The first three word regions, including the target matrix subject, were the same for all conditions, *e.g. the explorer who*. The fourth region was the verb embedding the intervener, *e.g. believe* for ECM/S-comp and *persuade* for object control structures. Region 5 was the complementizer *that*, and so only occurred in S-complement conditions. Regions 6-9 comprised the intervener. In simple intervener conditions, this was a singular and definite *the+noun* sequence, while in complex intervener conditions this was an adjective-modified singular and definite noun-noun compound, *i.e. the+adj+noun-noun*. The regions following the intervener comprised the embedded VP region, containing the embedded verb along with necessary tense/auxiliaries, *e.g. was/to (be)+prowl*, as well as the embedded object, *e.g. the ruins*. The embedded region always ended with a prepositional phrase, *e.g. for the expedition*, designed to favor VP-level attachment and provide a buffer between the embedded object and critical regions. The critical region was always of the form *aux+intransitive predicate*, and followed by a prepositional phrase that served as the spillover region.

Simple intervener conditions		
S-COMP	the explorer who believed that	the monster was prowling the ruins for the expedition was insane…
ECM	the explorer who believed	the monster to be prowling the ruins for the expedition was insane…
OBJ CTRL	the explorer who persuaded	the monster to prowl the ruins for the expedition was insane…
Complex intervener conditions		
S-COMP	the explorer who believed that	the ancient alien monster was prowling the ruins for the expedition was insane…
ECM	the explorer who believed	the ancient alien monster to be prowling the ruins for the expedition was insane …
OBJ CTRL	the explorer who persuaded	the ancient alien monster to prowl the ruins for the expedition was insane…

Table 3.3: Sample set of materials for Experiment 2. Intervener constituents are underlined. Not shown: Initial adverbial phrase and sentence-final prepositional phrase.

Procedure

Reading Time data was collected using the *Linger* software, using a procedure identical to that described above for Experiment 1.

Analysis

The analysis of Experiment 2 was almost identical to that described for Experiment 1, but differed in the coding scheme for experimental factors. Extreme observations less than 50 ms and greater than 3000 ms were removed (.002%). Five participants were removed for extremely low comprehension accuracy (<65%). Reading times were analyzed in regions consisting of a single word, as shown below. We report only residual reading times, computed using a procedure identical to that described in Experiment 1. Outliers were removed by computing z-scores for residual RTs, by region and condition, then removing those observations whose z-score exceeded |3| (2%).

(67) Experiment 2 analysis regions:

Target The$_1$ explorer$_2$

Embedding V who$_3$ believed$_4$/persuaded$_4$ (that$_5$)

Distractor the$_6$ (ancient alien) monster$_9$

Embedded VP was$_{10}$ prowling$_{12}$ the$_{13}$ ruins$_{14}$

to$_{10}$ be prowling$_{12}$ the$_{13}$ ruins$_{14}$

to$_{10}$ prowl$_{12}$ the$_{13}$ ruins$_{14}$

Embedded PP for$_{15}$ the$_{16}$ expedition$_{17}$

Critical was$_{18}$ insane$_{19}$

Spillover after$_{20}$ the$_{21}$ journey$_{22}$.

The residual RTs at each region were analyzed in a series of linear mixed-effects models. Each model used the maximal random effects structure in the sense of Barr et al. (2013, see Experiment 1). Coefficient contrasts for the similarity manipulations used a Helmert coding scheme, which compared (i) S-complement to ECM sentences (the 'subject intervener' conditions), and

(ii) object control sentences to the mean of subject intervener conditions. In other words, the Helmert contrasts compared the effect of case within sentences with an intervening subject, and the effect of object interveners to subject interveners.

Contrasts for intervener complexity used deviation coding (.5,-.5), with noun-noun compounds as the positive coefficient. Pairwise comparisons were evaluated using linear mixed-effects models fit to data for the two conditions of interest only, using treatment coding so that coefficient estimates reflect differences in the condition means. To avoid confusion with the experimental models, we report the 95%c.i. of the pairwise coefficient, determined by a Wald test. As in Experiment 1, we interpret significant coefficients as those whose t/z exceeded $|2|$.

Comprehension question accuracy was analyzed using logistic regression models, using the same coding scheme described above for the residual RTs. Since comprehension questions were the same for each item set, and so between items, comprehension models included only random intercepts for items. Unless otherwise noted, all reported coefficients for both RT and comprehension models were significant at $\alpha = .05$ or less.

3.6.3 Results

Reading times

Reading times for all conditions are summarized in Figure 3.6. As in Experiment 1, we found an immediate slowdown at the verb for sentences containing complex interveners. This was followed by a delayed interaction of complexity and structural similarity (*e.g. believe* v. *persuade*), such that the difficulty for complex interveners was attenuated in the ECM conditions.

Critical & Spillover regions.— At the critical auxiliary, there were no significant effects for experimental factors or their interactions. As in Experiment 1, we suggest that the auxiliary initiates subject retrieval, but the effects show up one word downstream as an artifact of the self-paced reading procedure.

At the matrix main verb, sentences with an intervening noun-noun compound were reliably read more slowly than sentences with a simple DP intervener (β = 12 ms, s.e. = 4 ms). The effect of

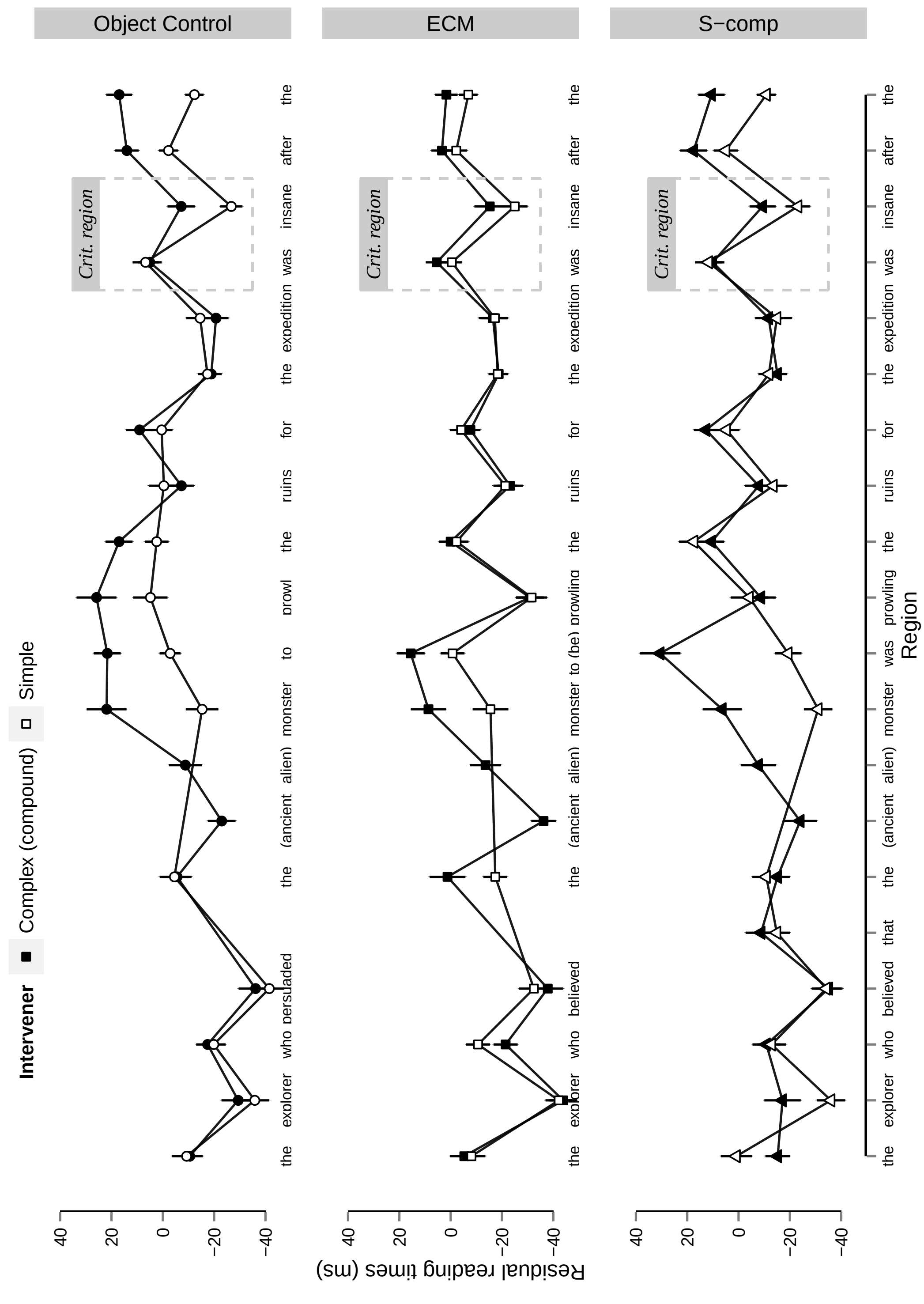


Figure 3.6: Experiment 2 residual reading times, for all conditions, summarizing mean residual reading Times (y-axis) for each experimental region (x-axis). Error bars show one standard error of the mean.

structure, however, was not significant: S-comp sentences were not reliably different from ECM sentences (β = 3 ms, s.e. = 2 ms), and object control sentences were not reliably different from the pooled subject intervener conditions (β = 1 ms, s.e. = 1 ms). There were no interactions. Pairwise comparisons at the matrix main verb revealed that the effect of size was significant only for Object Control structures (95% c.i. [5 ms, 36 ms]) and trending significant for S-comp structures (95% c.i. [0 ms, 27 ms], t/z = 1.97).

The slowdown for complex interveners persisted to the spillover region, but differences emerged between structures. The effect of complexity was significant throughout the spillover region, with a slowdown for complex interveners at the preposition (β = 8 ms, s.e. = 4 ms), the determiner (β = 21 ms, s.e. = 5 ms), and the head noun (β = 23 ms, s.e. = 9 ms). The effect of structure was also significant, such that S-complement structures were read significantly more slowly than ECM structures at the preposition (β = 5 ms, s.e. = 2 ms) and the head noun (β = 10 ms, s.e. = 5 ms), but there were no significant differences between the Object Control structures and subject intervener conditions at any region. The interaction of size and structure was trending significant at the head noun, suggesting that the cost for complex interveners was more acute in Object Control conditions than subject intervener conditions (β = 10 ms, s.e. = 5 ms; t = 1.916). However, the maximal model at the spillover noun did not contain random slopes for the interaction terms, either by-participant or by-item, and thus pooled the error across conditions. Pairwise comparisons revealed that the effect of complexity was not reliable for ECM structures at any spillover region. Complex interveners led to a slowdown for Object Control sentences at the preposition (95% c.i. [1 ms, 32 ms]), determiner (95% c.i. [9 ms, 48 ms]), and head noun (95% c.i. [4 ms, 82 ms]). The slowdown for complex intervener was also significant in S-complement structures, but only at the determiner (95% c.i. [7 ms, 35 ms]). Pairwise comparisons also revealed that the effect of structure at the preposition was driven by differences within the complex intervener conditions, such that complex S-complement sentences were reliably slower than complex ECM sentences (95% c.i. [1 ms, 28 ms]), but there were no reliable differences within simple intervener conditions.

Relative pronoun & intervener regions.— There were no significant effects or interactions in models fit to the relative pronoun region. Visual inspection suggested that there were significant differences among structures, which was confirmed in pairwise comparisons. However, there were no lexical differences between conditions at this point, and the moving-window display precluded any preview effects, so we take these differences to be spurious. Models fit to the the relative clause verb region, i.e. *believe* for S-comp/ECM and *persuade* for Object Control conditions, also did not show any significant effects for experimental factors or their interactions.

At the intervener phrase, i.e. *the (ancient alien) monster*, there were no significant effects or interactions at the initial determiner. There were significant effects of both size and structure at the head noun (*monster*). Complex conditions were read significantly more slowly than simple conditions (β = 32 ms, s.e. = 7 ms). There was also a significant effect of structure, reflecting a slowdown for Object Control sentences relative to the subject intervener conditions (β = 5 ms, s.e. = 2 ms). There were no significant interactions, though we note that the maximal converging model did not allow the slopes of the interaction terms to vary either by-item or by-participant. Pairwise comparisons revealed that the main effect of complexity was driven by a slowdown for complex interveners at the intervener head noun in S-comp conditions (95% c.i. [13 ms, 62 ms]) and Object Control conditions (95% c.i. [17 ms, 60 ms]). The effect of size was not significant in the ECM conditions.

Embedded VP regions & pre-critical regions.— Reading times at the embedded verb phrase showed effects of structure, but no effect of intervener size. At the embedded verb, S-comp conditions were read more slowly than ECM conditions (β = 12 ms, s.e. = 3 ms), and Object Control conditions were read more slowly than subject intervener conditions (β = 11 ms, s.e. = 2 ms). There was no effect of intervener size, and no interactions. The maximal model did not contain random slopes for the interaction terms, either by-participant or by-item. In pairwise comparisons, the effect of intervener complexity was significant only for Object Control sentences (95% c.i. [1 ms, 40 ms]). Pairwise comparisons also revealed significant differences for all structures. Within complex intervener conditions, Object Control sentences were reliably slower than either

S-comp sentences (95% c.i. [15 ms, 54 ms]) or ECM sentences (95% c.i. [32 ms, 79 ms]), while S-comp sentences were reliably slower than ECM sentences (95% c.i. [3 ms, 42 ms]). Within simple intervener conditions, S-comp sentences were reliably slower than ECM sentences (95% c.i. [7 ms, 48 ms]). Simple Object Control sentences were also significantly slower than simple ECM sentences (95% C.I. [21 ms, 53 ms]), but there was no difference between simple S-comp and Object Control sentences.

There were significant effects of structure at the embedded object regions, such that S-comp conditions were read more slowly than ECM conditions at both the determiner (β = 9 ms, s.e. = 2 ms) and head noun (β = 5 ms, s.e. = 3 ms) of the embedded object. There were no other significant effects, but the interaction of size and structure was trending significant at the determiner, suggesting that the effect of size was more acute in Object Control conditions compared to subject intervener conditions (β = 5 ms, s.e. = 3 ms; t = 1.927). The maximal model at both the determiner and head noun did not contain random by-item slopes for the interaction terms, while the maximal model at the head noun did not contain by-participant random slopes for the complexity factor. Pairwise comparisons revealed a slowdown for complex interveners at the determiner, but only for Object Control structures (95% c.i. [1 ms, 29 ms]). Within complex intervener conditions, S-comp sentences were reliably slower than ECM sentences at the embedded object noun (95% c.i. [1 ms, 30 ms]), where Object Control sentences were also reliably slower than ECM sentences (95% c.i. [1 ms, 32 ms]). The slowdown for complex Object Control sentences to complex ECM sentences was trending significant at the determiner (95% c.i. [0 ms, 33 ms], t/z = 1.97). Simple S-comp sentences were reliably slower than simple ECM sentences at the determiner (95% c.i. [4 ms, 34 ms]), while a slowdown for simple Object Control sentences relative to simple ECM sentences was significant at the embedded object head noun (95% c.i. [2 ms, 39 ms]).

The pre-critical regions at the end of the embedded clause, e.g. *for the expedition*, showed fleeting effects of structure, presumably due to spillover from the embedded VP regions. At the preposition head, S-comp conditions were read more slowly than ECM conditions (β = 7 ms, s.e. = 2 ms). This model did not contain by-participant random slopes for the interaction terms. Pairwise comparisons revealed that the effect of structure at the pre-critical preposition was driven by differences among the complex intervener conditions, where both Object Control and S-comp sentences reliably slower than ECM sentences (Object Control 95% c.i. [0 ms, 35 ms]; S-comp 95% c.i. [8 ms, 33 ms]). There were no significant effects of size or interactions at the pre-critical preposition, and no significant effects whatsoever at either the pre-critical determiner or noun.

Comprehension accuracy

Mean accuracy scores on comprehension questions are shown in Table 3.4. Overall, we found significant effects of structure, but no reliable effect of intervener size.

	Structure		
Intervener size	S-comp	ECM	Object Control
Small	80%(2%)	79%(2%)	73%(2%)
Large	77%(2%)	79%(2%)	72%(2%)
	79%(2%)	79%(2%)	73%(2%)

Table 3.4: Experiment 2 comprehension accuracy, condition means (and standard errors).

Grand mean accuracy on comprehension questions was 77%. Overall, Object Control conditions were significantly less accurate than subject intervener conditions (S-comp, ECM; β = -.13 ms, s.e. = .03 ms). There were no effects for intervener complexity or interactions of complexity and structure. Within complex intervener conditions, object control sentences were significantly less accurate than subject intervener sentences (β = -.12 ms, s.e. = .05 ms). In pairwise comparisons, complex Object Control sentences were reliably less accurate than complex ECM sentences (β = -.43 ms, s.e. = .17 ms), but the difference between complex Object Control and complex S-comp was only trending significant (β = -.28 ms, s.e. = .16 ms, p = .09). Within simple intervener conditions, Object Control sentences less accurate than sentences with a subject intervener (β = -.14 ms, s.e. = .05 ms). In pairwise comparisons, complex Object Control were reliably less accurate than both complex S-comp sentences (β = -.45 ms, s.e. = .17 ms) and complex ECM sentences (β = -.39 ms, s.e. = .17 ms).

N-Back working memory task

Participants' performance on the N-Back working memory task was included as a predictor in the RT and comprehension accuracy models discussed above. Performance was measured by computing d-prime accuracy scores for the 3-back portion of the task (hereafter: 'working memory span').

Working memory span was a significant predictor of RTs at the matrix main verb (β = 9 ms, s.e. = 4 ms) and the subsequent preposition of the spillover regions (β = 8 ms, s.e. = 4 ms). At both regions, participants who read more slowly at the critical and spillover regions also showed higher working memory span, suggesting that higher WM span actually made processing at the verb more difficult.

These effects are summarized in Figure 3.7, which shows the relation between N-Back performance and residual RTs at the critical and spillover regions. The slope of the line was determined using a first-order linear regression fit to the mean residual RTs for each condition by participant.

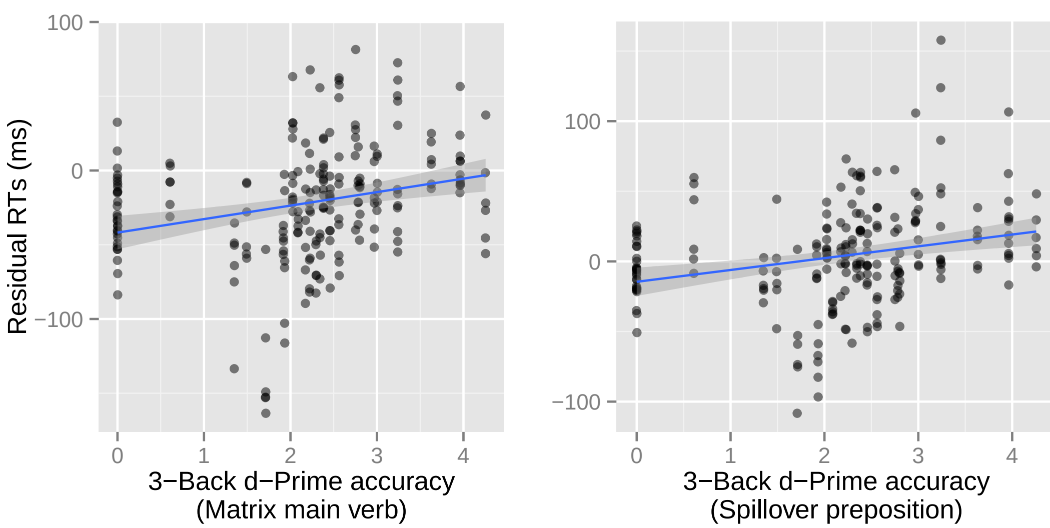


Figure 3.7: Experiment 2 working memory task (N-Back). By-participant residual reading times (y-axis) plotted against d-prime accuracy on the working memory task. The slope of the line shows a first-order linear regression for the relation between WM accuracy and residual RTs at the critical main verb (left) and spillover region (right).

3.6.4 *Experiment 2 discussion*

The results of Experiment 2 suggest that a complex nominal expression leads to difficulty at a subsequent VP. When a noun-noun compound intervened between a verb and its target subject, the reading times at the matrix main verb were significantly longer than when the intervener was a simple *Det+Noun* constituent. Furthermore, while this difficulty arose at the matrix main verb, there was no effect of complexity at the pre-critical regions, indicating that the difficulty arose at the retrieval site, the verb, i.e. retrieval interference. However, the complexity effect was attenuated in ECM structures, with no reliable differences between complex and simple ECM sentences at the intervening head nun, the critical main verb, or the spillover region. Thus, we found that complex interveners lead to difficulty in the retrieval required at a verb, but the strength of this difficulty is modulated by the syntactic properties of the intervener.

As in Experiment 1, complex interveners led to difficulty at the matrix verb regions. The results of Experiment 2 extend the effect of complexity to noun-noun compounds (see Hofmeister, 2011, for a similar manipulation). We return to these issues in the general discussion, but note that the noun-noun compounds used in Experiment 2 lack the internal argument structure of the event nominalizations used in the previous experiment. Thus, the results of Experiment 2 extend the complexity effect to include noun-noun compounds, suggesting that whatever is more difficult about complex constituents, the effect is not due to the need to establish additional thematic bindings.

In contrast to the immediate effect of intervener complexity, we also observed effects of structure at the post-verbal regions of both the embedded and matrix clause. Within the subject intervener conditions, S-complement sentences were consistently more difficult than ECM conditions. We also found that Object Control sentences were reliably more difficult than subject intervener conditions, reflected in longer reading times and lower comprehension accuracy. Consider first the differences among the subject intervener conditions. There were no differences at the pre-critical region, indicating that the difficulty arises at the verb, when the subject is retrieved. Furthermore, whereas the complexity effect for S-complement structures—the same structures used in Experiment 1—persisted to the spillover region, there was no corresponding slowdown for ECM structures. Thus, the retrieval structure at the verb is able to discriminate between the embedded subjects of S-complement and ECM structures.

While both S-complement and ECM interveners are structural subjects, they are distinguished by their case properties: NOMINATIVE for S-complement, and ACCUSATIVE for ECM. Alternatively or additionally, they differ in the syntactic properties of their dominating clause. S-complement verbs embed a finite (tensed) clause, while ECM verbs embed a non-finite (untensed) clause. The target subject was always a nominative subject of a finite clause. Thus, in terms of the retrieval cues, differences within the subject intervener conditions indicate that the embedded subjects of S-comp, but not ECM, structures match the subject cues provided by the verb. The relevant cues, then, would target the nominative subject of a finite clause.

A potential worry in these results is the presence of a local ambiguity in the ECM conditions. Until the disambiguating embedded verb, ECM structures are consistent with continuations where the intervener is parsed as either a matrix object or embedded subject—*i.e.* an 'NP/S' ambiguity (Sturt, M. J. Pickering, & Crocker, 1999). If the intervener is initially parsed as the matrix object, then this lingering misanalysis might give rise to the observed attenuation of the subject-size effect by increasing structural dissimilarity—the intervener would no longer be a structural subject. Experiment 3 was designed to address the role of ambiguity, while using eye-tracking to further probe the timecourse of intervener size and structural similarity.

3.7 Experiment 3: The role of ambiguity

3.7.1 Design & predictions

The reading times of Experiment 2 showed a main effect of intervener size at the main verb, and a delayed effect of structural similarity. These contrasting effects potentially implicate a distinct timecourse of processing at the verb. For instance, it may be that the intervener size effects reflect an impairment in early processing, *e.g.* structure-building, while the delayed effect of similarity reflects later retrievals required for semantic integration.

To better understand the timecourse of processing at the verb, Experiment 3 was designed to replicate the effects of Experiment 2, using an eye-tracking methodology. If the distinct profiles of intervener size and structural similarity are the result of early versus late processing, then this may be reflected in the corresponding reading time measures, discussed below.

Additionally, a key finding of Experiment 2 was an interaction of intervener size and structural similarity in the spillover region, with a slowdown for large interveners in S-complement – but not ECM – structures. One interpretation of this is that the interference caused by an ECM intervener is mitigated by structural dissimilarity. However, it may be that an early effect of similarity in the ECM conditions was masked by persistent effects of ambiguity. To better understand the role of ambiguity in the ECM sentences, the materials of Experiment 3 also included ambiguous versions of the S-complement structures, constructed by removing the complementizer *that*.

3.7.2 Method

Participants

40 participants from the University of California, Santa Cruz community participated in the experiment. All participants gave informed consent, and received course credit or $10 for participating.

Region	2 *target*	3 *embedding*	4 *intervener*	5 *disambiguation*	6 *pre-critical*	7 *critical*	8 *spillover*
s-comp	the explorer	who believed that	the monster	was prowling	the ruins for the expedition	was insane	after the journey.
—amb.	the explorer	who believed	the monster	was prowling	the ruins for the expedition	was insane	after the journey.
ecm	the explorer	who believed	the monster	to be prowling	the ruins for the expedition	was insane	after the journey.
obj ctrl	the explorer	who persuaded	the monster	to prowl	the ruins for the expedition	was insane	after the journey.
×size			the ancient alien-monster				

Table 3.5: Sample materials for Experiment 3, with regions of interest for eye tracking measures.

Materials

A sample set of materials and regions of analysis is provided in Table 3.5. The materials of Experiment 3 are largely identical to those of Experiment 2, with one exception. Since the ECM conditions were ambiguous up to the embedded verb, an additional condition created ambiguous S-complement versions by dropping the complementizer *that*.

Procedure

Eye movements were recorded using an Eyelink 1000 (SR Research Ltd., Ottawa, Canada), which recorded the position of the eye twice every millisecond with a spatial resolution of 0.01 degrees. A chin rest was used to minimize head movement. Calibration was performed at the beginning of the experiment, and throughout when necessary, using a 9-point grid.

Each trial began with a gaze trigger. Upon detection of a stable fixation on this trigger, the full sentence was presented. Participants indicated completion of the trial by pressing a button on a gamepad, after which a comprehension question was presented in its entirety. Answers to the comprehension questions (*yes/no*) were entered by pressing the appropriate button on the gamepad.

Analysis

Artifact rejection was performed using the University of Massachusetts Eye Tracking Lab *Eye Doctor* software. Trials containing blinks or evidence of tracking loss at the critical region were removed. Within trials, blinks were removed, while fixations drastically off the line of text were treated as blinks and removed. An automatic process combined fixations less than 80 ms that occurred within 1 character of a sequentially adjacent fixation, while other fixations less than 80 ms and greater than 1000 ms were removed (Staub, 2011; Kwon & Sturt, 2014).

Regions of interest are annotated in the sample materials shown in Table 3.5. Dependent measures were computed by an automated process, then imported to R for statistical analysis.

In what follows, we report the results in terms of four dependent measures. In interpreting these measures, we follow previous work on interference using eye-tracking (e.g. Logačev & Vasishth, 2012). *First pass times* are the sum of all fixations in a region before exiting to either the left or the right – an estimate of early processing stages, including difficulty in integrating the text during reading (Inhoff, 1984). *Go-past times* are all fixations in a region before exiting to the right, and thus include time spent regressing to previous regions. We take go-past times to reflect difficulty in integrating a word with the preceding context (Clifton, Staub, & Rayner, 2007), and note that previous studies of subject interference have observed interference in go-past times (Van Dyke, 2007; Van Dyke & McElree, 2011) *Total times* are the sum of all fixations in a region, including regressions and re-reading. Finally, *first-pass regressions out* are the observed proportions of regressions (backwards saccades) out of a given region, which we take to reflect processing difficulty at the given region, and in particular difficulty in re-analysis (Frazier & Rayner, 1982a; Meseguer, Carreiras, & Clifton, 2002).

Dependent measures at each region were modeled with a series of linear mixed-effects models, using the `lme4` package in R (D. Bates et al., 2013; R Core Team, 2016). Outliers were removed by computing z-scores by region and condition, excluding those observations greater than $|3|$ standard deviations. As in Experiment 2, experimental factors were coded using Helmert contrasts, comparing: (i) ambiguous S-complement conditions to ECM conditions (the 'ambiguous conditions'); (ii) the pooled RTs from the ambiguous conditions to the unambiguous S-complement sentence (the 'SPEC-T conditions'), and (iii) the pooled SPEC-T conditions to control verb conditions. Contrasts for intervener size used deviation coding (.5, -.5), with large interveners as the positive coefficient. Significant effects were determined as those coefficient estimates with a t/z-score greater than 2 (Gelman & Hill, 2007). Unless otherwise noted, all reported coefficients were significant at $\alpha = .05$.

Structure	Complexity	Region 2 *Target*	Region 3 *who+V*	Region 4 *Intervener*	Region 5 *Emb. V*	Region 6 *Pre-crit*	Region 7 *Critical*	Region 8 *Spillover*
First pass times (ms)								
ECM	Simple	428 (17)	473 (17)	407 (14)	571 (22)	894 (31)	415 (14)	564 (23)
	Complex	405 (14)	430 (16)	998 (40)	553 (21)	850 (30)	398 (12)	589 (26)
S-comp	Simple	388 (14)	463 (17)	411 (15)	451 (17)	908 (35)	394 (12)	548 (22)
—ambig.	Complex	432 (19)	449 (17)	1039 (42)	498 (19)	755 (31)	412 (14)	544 (23)
S-comp	Simple	403 (17)	621 (24)	359 (15)	446 (17)	856 (32)	393 (13)	524 (23)
—unambig.	Complex	391 (13)	639 (27)	902 (43)	450 (16)	822 (31)	391 (13)	552 (24)
Obj. Ctrl	Simple	388 (13)	446 (14)	403 (14)	429 (16)	920 (32)	378 (12)	567 (25)
	Complex	397 (14)	444 (15)	1012 (38)	394 (15)	895 (31)	412 (12)	577 (21)
Regression path/Go-past (ms)								
ECM	Simple	521 (25)	531 (20)	597 (31)	735 (32)	1096 (39)	499 (20)	3922 (215)
	Complex	480 (21)	554 (27)	1463 (56)	688 (32)	993 (30)	442 (17)	4354 (245)
S-comp	Simple	501 (24)	553 (23)	574 (28)	544 (23)	1199 (46)	485 (21)	4439 (256)
—ambig.	Complex	524 (24)	598 (27)	1434 (55)	732 (43)	1213 (63)	567 (28)	4699 (253)
S-comp	Simple	524 (25)	739 (27)	562 (31)	558 (27)	1098 (37)	472 (20)	4550 (256)
—unambig.	Complex	485 (24)	830 (40)	1334 (50)	581 (35)	1022 (36)	522 (28)	5013 (255)
Obj. Ctrl	Simple	463 (20)	528 (21)	507 (23)	568 (30)	1112 (39)	491 (26)	4029 (218)
	Complex	496 (23)	547 (24)	1425 (59)	517 (30)	1056 (39)	496 (20)	4047 (228)
Total times (ms)								
ECM	Simple	1047 (41)	1095 (44)	937 (39)	1172 (46)	1673 (62)	806 (30)	910 (35)
	Complex	1084 (50)	1189 (54)	2176 (85)	1062 (44)	1534 (60)	765 (34)	1008 (44)
S-comp	Simple	1087 (49)	1276 (59)	987 (44)	1079 (48)	1930 (79)	884 (39)	981 (39)
—ambig.	Complex	1136 (47)	1254 (52)	2525 (96)	1175 (46)	1916 (76)	938 (39)	982 (37)
S-comp	Simple	1117 (51)	1697 (72)	914 (44)	994 (44)	1881 (74)	904 (41)	997 (44)
—unambig.	Complex	1079 (49)	1707 (72)	2162 (83)	938 (37)	1769 (65)	912 (41)	1027 (45)
Obj. Ctrl	Simple	1056 (46)	1045 (39)	956 (40)	890 (41)	1761 (66)	826 (37)	996 (46)
	Complex	1056 (47)	1108 (42)	2228 (86)	806 (39)	1587 (59)	833 (35)	933 (34)
Regressions out (%)								
ECM	Simple	13.8 (2.6)	8.9 (2.1)	21.9 (3.1)	18.7 (2.9)	14.7 (2.6)	15.3 (2.7)	91.1 (2.1)
	Complex	10.7 (2.4)	16.9 (2.9)	29.7 (3.5)	12.3 (2.5)	12.8 (2.6)	6.7 (1.9)	91.8 (2.1)
S-comp	Simple	14.4 (2.7)	14 (2.6)	23.6 (3.2)	17.2 (2.9)	21.1 (3.1)	13.9 (2.6)	93.7 (1.9)
—ambig.	Complex	13.1 (2.5)	18.1 (2.9)	25.7 (3.3)	20.8 (3.1)	27.5 (3.4)	18.4 (2.9)	92.7 (2)
S-comp	Simple	14.9 (2.8)	15.2 (2.8)	29.8 (3.5)	18.2 (3)	19.3 (3)	10.6 (2.4)	90.1 (2.3)
—unambig.	Complex	10.1 (2.4)	16.1 (2.9)	26.7 (3.5)	13.8 (2.7)	19.3 (3.1)	17.1 (3.1)	93.7 (1.9)
Obj. Ctrl	Simple	9.8 (2.3)	11 (2.4)	16.3 (2.8)	16.5 (2.9)	15.5 (2.8)	14 (2.7)	93.1 (1.9)
	Complex	13.1 (2.6)	12.3 (2.5)	23.3 (3.2)	14.9 (2.8)	14 (2.6)	12 (2.5)	87.8 (2.5)

Table 3.6: Experiment 3 means (and standard errors) for first-pass, go-past, total times, and proportion of regressions out. Aggregated over participants/items, after removing extreme values (> |3| s.d. by region and condition).

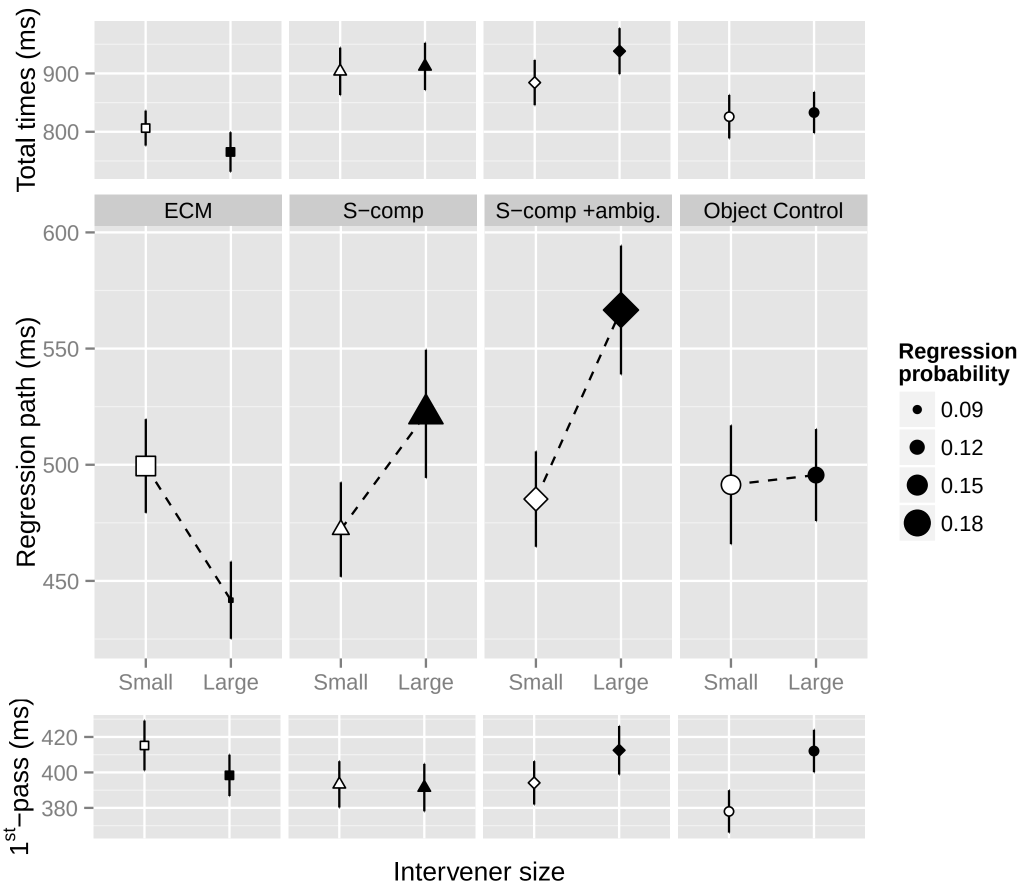


Figure 3.8: Experiment 3 eye-tracking results for the critical region (matrix aux and main verbs), showing condition means in first-pass times (bottom), go-past times (middle) and total times (top). Probability of first-pass regressions out is reflected in the size of the points in go-past times; Error bars show standard error of the mean, aggregated by participants and items.

3.7.3 *Results*

Eye-tracking measures

A summary of Experiment 3 reading time data for all measures is shown in Table 3.6. Reading times showed clear evidence of ambiguity in early measures across the embedded region, while later measures showed an interaction between intervener size and structural similarity.

R7–8: Critical & spillover regions.— Results at the critical region for all measures are summarized in Table 3.8. At the matrix auxiliary and main verb region, there was a significant interaction of intervener size and structure in go-past times, with large interveners leading to significantly slower go-past times in the finite S-complement compared to ECM conditions (β = 62 ms, s.e. = 23 ms). There were no differences within the S-complement conditions. The pattern was more robust in the proportion of regressions out, with finite S-complements showing an overall greater proportion of regressions out than ECM conditions (β = .22, s.e. = .11; z = 1.97). The interaction effect was also present in regressions out, such that large interveners led to significantly more regressions in the S-complement conditions compared to the ECM condition (β = .71, s.e. = .22). S-complement conditions also showed reliably slower total times overall than the ECM condition (β = 58 ms, s.e. = 17 ms). There were no significant effects in first-pass times for experimental factors, and no significant effects in any measure at the spillover region.

R3: Embedding verb & relative pronoun region.— The addition of the disambiguating complementizer *that* led to significant differences in structure at all measures except regressions out. Unambiguous S-complement conditions were slower than ambiguous S-complements in first-pass times (β = 86 ms, s.e. = 13 ms), go-past times (β = 107 ms, s.e. = 15 ms), and total times (β = 225 ms, s.e. = 32 ms). This effect persisted in the pooled comparisons, such that finite S-complement conditions were reliably slower than ECM conditions in first-pass times (β = 47 ms, s.e. = 8 ms), go-past times (β = 70 ms, s.e. = 11 ms), and total times (β = 172 ms, s.e. = 24 ms). Additionally, SPEC-T (ECM+S-complements) intervener conditions were significantly slower than object control conditions in first-pass times (β = 26 ms, s.e. = 10 ms), go-past times (β = 38 ms, s.e. = 11 ms), and total times (β = 120 ms, s.e. = 26 ms). There were no differences at the relativizing region in proportion of regressions out.

R4: Intervener region.— Unsurprisingly, the additional length for large interveners was significant at the embedded subject/controller region in all measures. Complex interveners were significantly slower than small interveners in first-pass times (β = 608 ms, s.e. = 42 ms), go-past times (β = 893 ms, s.e. = 60 ms), total times (β = 1309 ms, s.e. = 88 ms), and led to significantly more regressions out (β = .34, s.e. = .16, z = 2.09). Additionally, SPEC-T intervener conditions showed reliably more regressions out than object control conditions (β = .2, s.e. = .08). Total times also showed significant effects of ambiguity in the finite-S-complement conditions, reflected in slower total times for ambiguous S-complements compared to unambiguous S-complements (β = 89 ms, s.e. = 33 ms).

R5–6: Embedded verb & pre-critical regions.— The results at the embedded verb region paralleled the effects at the critical region, with additional effects of ambiguity. Complex interveners led to significantly faster total times than small interveners (β = -70 ms, s.e. = 35 ms). Within the S-complement conditions, ambiguous conditions showed reliably slower total times than unambiguous conditions (β = 69 ms, s.e. = 23 ms). There was also an interaction of size and structure for the finite S-complement conditions in go-past times, such that ambiguous S-complement conditions with a large intervener were reliably slower than unambiguous S-complements with large interveners (β = 86 ms, s.e. = 36 ms). ECM conditions were reliably slower than S-complement conditions in first-pass times (β = 51 ms, s.e. = 12 ms) and go-past times (β = 54 ms, s.e. = 17 ms). Additionally, and similar to the critical region, analysis of go-past times revealed an interaction of intervener size and structure, reflecting a slowdown for large interveners that was attenuated in the ECM conditions compared to S-complement conditions (β = -72 ms, s.e. = 27). Spec-T intervener conditions were significantly slower than object control conditions in first pass times (β = 50 ms, s.e. = 9 ms), go-past times (β = 60 ms, s.e. = 17 ms), and total times (β = 116 ms, s.e. = 24 ms). There were no significant effects for experimental factors in the proportion of regressions out.

At the embedded object and prepositional modifier region, there were significant effects of both size and structure. Total times for large interveners were reliably faster than small intervener conditions (β = -126 ms, s.e. = 50 ms). Within the S-complement conditions, go-past times for ambiguous S-complement conditions were reliably slower than their ambiguous counterparts (β = 63 ms, s.e. = 22). S-complement conditions were significantly slower than ECM conditions in go-past times (β = 41 ms, s.e. = 16 ms), total times (β = 132 ms, s.e. = 26 ms), and showed a significantly greater proportion of regressions out (β = .29, s.e. = .09). First-pass times for object

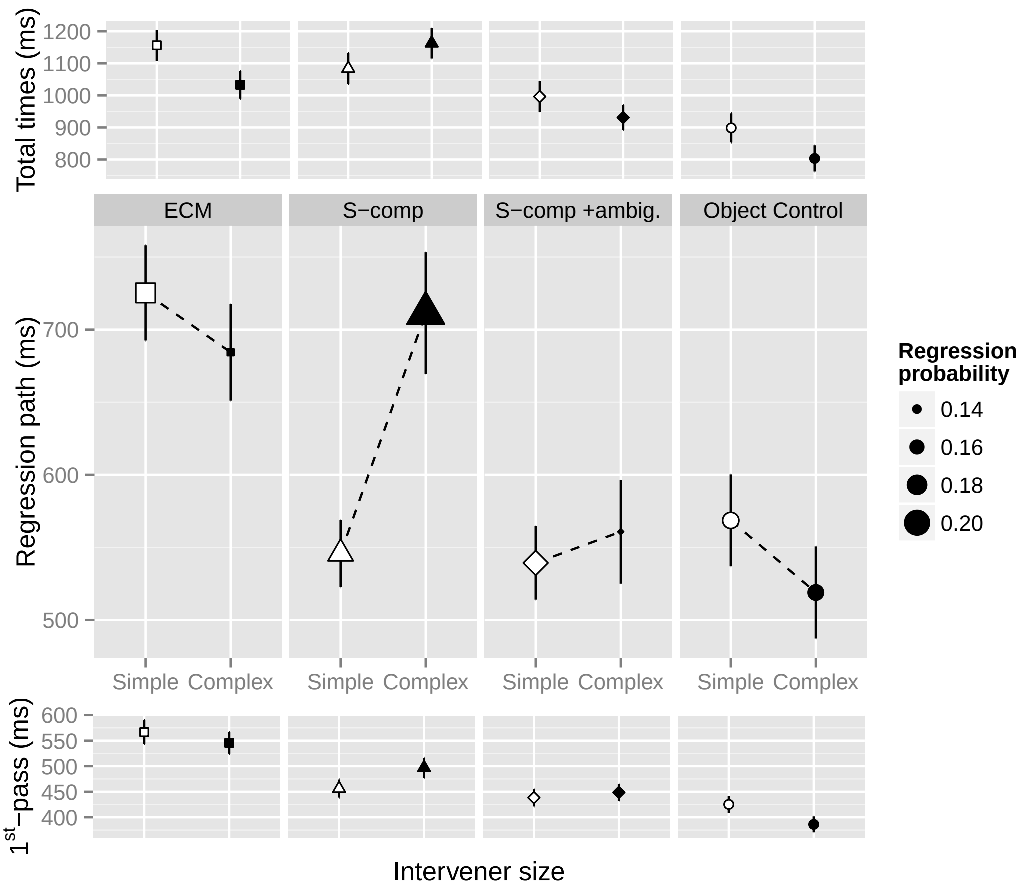


Figure 3.9: Experiment 3 eye-tracking results at the embedded verb region (r5), showing condition means in first-pass times (bottom), go-past times (middle) and total times (top). Probability of first-pass regressions out is reflected in the size of the points in go-past times; Error bars show standard error of the mean, aggregated by participants and items.

control conditions were significantly slower overall than spec-T conditions (β = 31 ms, s.e. = 14 ms).

Comprehension accuracy

Grand mean accuracy on comprehension questions was 76%, with condition means ranging from 72% for large intervener S-complement conditions to 81% for small intervener S-complement conditions. Complex interveners led to significantly lower accuracy in S-complement sentences (β = -.59 ms, s.e. = .27 ms). Otherwise, there were no other effects for experimental factors, and no interactions.

3.7.4 Discussion

Overall, the eye-tracking results of Experiment 3 were consistent with the previous experiments' findings of interference at the matrix verb. Complex interveners led to increased difficulty at the matrix verb, but this difficulty was attenuated in Object Control sentences and reversed in ECM conditions. However, the results of Experiment 3 provide evidence that the interaction of complexity and structural similarity was not due to local ambiguities in the ECM conditions.

At the critical region, S-complement conditions were more difficult when the intervening subject was a noun-noun compound, reflected as a greater proportion of regressions out and slower go-past times. This is consistent with the complexity effect of Experiments 1 & 2, which also used (unambiguous) S-complement conditions. However, in contrast to Experiment 2, there was no corresponding complexity effect for Object Control conditions. Instead, the eye-tracking results of the current experiment for Object Control structures indicated no effect of complexity in either go-past times or regressions out, suggesting that comprehenders are robust to interference effects in these configurations. Similarly, whereas the effect of complexity for ECM interveners was attenuated in Experiment 2, in the current experiment complex ECM sentences showed *faster* go-past times and fewer regressions out than simple intervener ECM sentences.

The differences between Experiments 2 and 3 in both the presence and and form of the interference effect for accusative interveners might reflect the more natural experimental setting for eye-tracking. However, the patterns in regressions out for the Object Control sentences and ECM sentences suggest that comprehenders had no difficulty in re-analysis, as we might expect if the effects in Experiment 2 were due to the moving-window display prohibiting regressing to the string to be re-analyzed. Instead, the reduction in regressions out for complex ECM sentences suggests that complexity reduces the need for re-analysis. It may be, then, that the self-paced reading task of Experiment 2 precluded re-analysis in the simple intervener conditions, which masked the facilitation for complex interveners observed in the current experiment. For Object Control structures, Experiment 3 used a smaller set of Object Control verbs, which were also selected to be relatively more frequent. This may have reduced by-item variability, and reduced any difficulty for integrating the relatively low-frequency Object Control structures. It is somewhat surprising, though, that any frequency effects would show up at the matrix verb, and not at the Object Control verb itself. In any case, we think that there are sufficient reasons to be wary of drawing conclusions from the Object Control results of Experiment 2, and that the results of Experiment 3 are more representative.

Taken together, the consistent patterns of difficulty for complex subjects (not objects) observed in Experiments 1 and 3, in conjunction with the modulation of this effect by the interaction with structural similarity, suggest that the retrieval structure at the verb is minimally able to distinguish between structural subjects and objects, while the interference effect is further modulated by the morphosyntactic properties of the intervening subject. We return to these conclusions in the general discussion.

The results of Experiment 3 also indicate that the complexity effect for ECM conditions was not due to the presence of a local ambiguity. First, as discussed above, at the critical region ambiguous S-complement conditions patterned with their unambiguous counterparts, not the ECM conditions. This suggests that any effect of ambiguity is resolved by the time the matrix verb is encountered. Second, while ECM and ambiguous S-complement conditions were structurally ambiguous up to the embedded verb, at this region only the ECM conditions showed the facilitation effect observed at the critical region: a slowdown for complex interveners in the ambiguous S-complement conditions, but a facilitation in the ECM conditions. On the other hand, we observed a clear cost of ambiguity at the embedded verb, with both ambiguous conditions showing slower go-past and total times than unambiguous conditions. (ECM conditions also showed elevated first-pass times, presumably due to the presence of an additional word in the embedded verb region.) Thus, the interaction of complexity and structure appears to be a property of subject retrieval, and not a lingering effect of ambiguity.

3.8 General discussion: retrieval cues for subjects

This chapter has examined the nature of the information used to re-activate subject encodings at a verb by probing for difficulty in the presence of several kinds of intervening subjects. In three experiments, we manipulated the size of the intervening subject and the syntactic properties of the potentially interfering intervener. By manipulating the similarity between the matrix and embedded subjects, we used interference effects to diagnose how the retrieval cues characterize subjects.

The results suggest that some, but not all, subjects interfere. In Experiment 1 (§3.4), we found that the complexity of the intervening subject matters, in that complex interveners engendered greater difficulty at the matrix verb than simple *Det+Noun* interveners. On the other hand, the results of Experiments 2 and 3 (§3.6–3.7) indicate that the effect of complexity is modulated by the degree of morphosyntactic similarity between the complex intervener and the target subject. Overall, these results indicate that the dimensions of similarity for subject retrieval are morphosyntactic, and not thematic.

3.8.1 Structural versus thematic subjects

The reading time results of Experiment 1 suggest that re-activation of the subject at a verb is guided by phrase-structural, and not thematic, properties. While there was a clear interference effect for complex interveners at the main verb, there was no effect for the possessor, a thematic subject. This pattern suggests that the information in the retrieval context is sufficient to distinguish between the subjects of verbs and the possessor subjects of nominal expressions.

The lack of interference from the possessor may seem surprising in light of previous studies reporting interference from semantically similar intervening subjects (Van Dyke, 2007; see §3.3.2. If possessors are also semantically good subjects, as in (34), why don't they interfere? One answer to this question comes from the results of Van Dyke & McElree (2011), in that the syntactic position of the interfering element matters (see §3.3.3). Van Dyke & McElree suggest that the difference turns on the syntactic position of the intervener: a subject in (34), and an object in (38). These patterns, in conjunction with previous findings of interference from grammatically inappropriate subject encodings (Van Dyke & Lewis, 2003; Van Dyke, 2007), suggest that the retrieval structure at the verb is sufficient to distinguish between, but not within, core argument positions—objects don't interfere with subjects, but multiple subjects lead to difficulty. The more general conclusion is that structural information is prioritized above semantic information, so that syntactic cues gate interference from otherwise semantically similar constituents (see B. Dillon et al., 2013 for similar claims in the processing of reflexive anaphors in complex subjects, and Parker, 2014 for a competing view). Our results align with these conclusions, in that we found that complex interveners lead to difficulty, but only when structurally similar to the target subject.

3.8.2 *Intervener complexity*

The most consistent finding of experiments 1–3 was that the processing at the matrix verb was more difficult when the embedded intervener was complex, either an event nominalization (Experiment 1; §3.4) or a noun-noun compound (Experiments 2 and 3; §3.6–3.7). We take this to be an interference effect, for two reasons. First, the reading times results of Experiment 1 indicated a slowdown for a nominalization in subject, but not object, position, despite occurring in the same span of words. Second, comprehension question accuracy in Experiment 1 was significantly lower for overt subjects compared to the control condition, which used raising predicates without overt subjects. These patterns suggest that the complexity of the intervening constituent matters, but only when the intervener is similar to the target constituent, and in particular both are subjects.

The slowdown for complex embedded subjects is similar to previous studies showing interference effects in the processing of complex subjects, where experimental materials also used embedded finite clause complements, e.g. complements selected by *thought*, to introduce additional subjects (Gordon, Hendrick, & Marcus Johnson, 2001; Van Dyke & Lewis, 2003; Van Dyke, 2007; Van Dyke & McElree, 2011). We extend these findings to materials that do not vary in their depth of embedding.

However, why complex interveners – when matching – lead to difficulty remains an open question. Nominalizations are more complex than simple *det+noun* nominals in terms of thematic structure, morphology, and length. One potential explanation, consistent with these observations, is that the additional complexity requires further processing which results in a stronger representation of the embedded subject constituent. If complex constituents are associated with stronger representations, then the features of the complex intervener might disrupt those encoded on the target, a form of encoding interference (e.g. feature over-writing or conflicting bindings). Alternatively, complexity may modulate difficulty because attaching additional material to a head increases its availability and facilitates (mis)retrieval, i.e. retrieval interference. The results of the experiments presented here seem to favor the retrieval interference approach. For one, the effect of complexity arose at the verb, when the subject is retrieved. Two, the effect is modulated by the degree of similarity between the complex constituent and the target of retrieval, the matrix subject. Finally, in both Experiments 2 and 3, the encoding conditions for the intervener remained largely constant, so that the observed differences are more likely due to retrieval. Note, however, that latency measures such as eye-tracking and self-paced reading cannot distinguish between the strength of an encoding in memory and the speed of its access, both of which can modulate the accuracy of retrieval (see McElree, 2006, and references cited there). Thus, additional research is required to distinguish between encoding and retrieval interference effects, using methodologies such as the Speed-Accuracy Trade-off procedure to estimate retrieval dynamics.

The complexity effects observed here are consistent with previous studies showing that complexity modulates difficulty at the retrieval site. For instance, Hofmeister (2011) showed that filler-gap sentences were sensitive to the syntactic and semantic complexity of the filler. Complex constituents, including noun-noun compounds, led to longer reading times at the head of the filler, but facilitated processing at the gap site, with no effect on intervening regions. Hofmeister concludes that these effects implicate representational strength and similarity-based interference. In particular, Hofmeister argues that complexity modulates interference because attaching additional material to a constituent requires repeatedly re-activating that constituent's head, resulting in a stronger encoding for the complex constituent, an elaboration effect. Indeed, working within a cue-driven parsing model, Lewis & Vasishth (2005) assume that modifying a constituent requires either re-activating a predicted head or retrieving the to-be-modified constituent, strengthening the encoding (its activation levels). For example, adjectival modification re-activates the head of the modified noun. The takeaway of this approach is that elaboration of a constituent has the effect of increasing the associative strength of the features bound to that constituent, increasing the probability that those features will be unique, i.e. distinct. Increasing the distinctiveness of features would modulate interference by either reducing similarity at encoding, or increasing the associative match between the features of the complex constituent and the retrieval cues.

3.8.3 *Structural properties of subjecthood*

The results of Experiments 2 & 3 support the conclusion that subjects are identified based on their case properties, and not their syntactic position alone. Complex interveners led to greater difficulty in finite S-complement structures – longer RTs, go-past times, and more regressions out – but for ECM structures this difficulty was attenuated in self-paced reading (Experiment 2) and reversed in eye-tracking measures (Experiment 3), with faster go-past times and fewer regressions out. In Object Control structures, complex interveners led to a slowdown in Experiment 2, but had no effect in Experiment 3. These contrasting profiles might reflect task differences, such as the relatively more natural experimental task in eye-tracking, or that Experiment 3 used fewer Object Control verbs overall, with relatively higher frequency. In other words, we found strong, consistent evidence that complex subject interveners engendered interference in S-complement sentences; consistent evidence that complex interveners in ECM sentences either engender no interference or lead to facilitation; and we found inconsistent evidence for object control interveners.

The subject/object distinction is consistent with the findings of Van Dyke & McElree (2011), and indicates that phrase-structural information in the retrieval context guides re-activation of the subject. On the other hand, the contrasting effect of complexity for S-complement and ECM structures suggests that the the form of the effect is sensitive to the syntactic properties of the intervener. In particular, interference is sensitive to the case similarity between intervener and target, even when case is not overtly encoded in the surface form of the noun.

How can we account for the contrasting interference profiles for subject interveners? Assume that complexity has the effect of increasing the uniqueness of the features bound to the complex intervener, modulating the degree of cue-match between the intervener and the retrieval structure of the verb. Consistent with this view, the reading times results of Experiment 3 showed no differences between structures in simple intervener conditions—structural differences arose only for complex intervener conditions. If the features of the intervener fully match those of the retrieval target, as in S-complement structures, then complexity will strengthen the cue-match between the retrieval structure and the intervener, inhibiting subject retrieval due to cue-overload. In other words, complexity will make similarity between the intervener and target more acute. In Object Control structures, though, the intervener is an accusative object, and thus maximally dissimilar to the target nominative subject. Increasing the uniqueness of the intervener's features, then, will have no effect, since there is no competition from the intervener. Now consider ECM structures, where the intervener is similar to the target in its structural position (subject), but distinct in its case properties. In this configuration, complexity will increase the uniqueness of the accusative case feature, so a nominative case feature will become more diagnostic of the target subject. Thus, complexity has the effect of making the accusative subjects of ECM complements more distinct by highlighting their dissimilarity to the target.

3.8.4 Diagnosing subjects

The interactions of complexity and structural similarity indicate that abstract morphosyntactic information in the retrieval context is used to re-activate subject encodings. More specifically, the three-way contrast of structural similarity indicates that the cues provided by a finite verb identify the target subject in terms of both case and position. If only structural subjecthood mattered, then we should have observed inhibitory interference from both S-complement and ECM interveners, which both match the target in their structural position. Similarly, if only the case properties of the target subject mattered, then a stronger case cue would have led to inhibition in S-complement structures, as we observed, but would have made no distinction between the accusative interveners of ECM and Object Control structures, contrary to our results. Instead, we found that subject interveners interfere, while object interveners don't, but the form of the interference effect depends on the syntactic properties of the intervening subject. Thus, structural subjecthood (the position cue) seems to be a necessary condition for interference, while the case feature of the intervener modulates this similarity: inhibitory interference when it matches the target, but facilitatory interference for partially-matching interveners.

Is a retrieval cue for case plausible, even when case properties are not morphologically reflected, as in our materials? The results of Fedorenko, Babyonyshev, & Gibson (2004) and Logačev & Vasishth (2012), discussed in Section 3.5.3, suggest that it is. In particular, the results of Fedorenko, Babyonyshev, & Gibson (2004) showed a slowdown at the verb when both preceding object NPs matched in both abstract and morphological case, an interaction effect. This pattern is consistent with the results of our experiments, if we assume that in the absence of morphological case (there were no pronouns in our materials) retrieval relies on abstract case alone. Further evidence comes from Logačev & Vasishth (2012), who showed that similarity in abstract case alone was sufficient to engender interference in the processing of German sentences.

In summary, the results of the current study replicated previous findings of interference from embedded subjects, but demonstrates this effect to be modulated by both the complexity and syntactic properties of the intervening subject. In all three experiments, the S-complement conditions patterned together, with a slowdown for at the main verb when the embedded subject was complex, either an event nominalization or a noun-noun compound. In contrast, complex intervening subjects in ECM structures facilitated processing at the main verb, while the object control structures seemed robust to interference effects. Taken together, the patterns of interference for grammatically inappropriate subjects, but not objects, supports a retrieval cue for the syntactic position of the intervener. The modulation of this effect in S-complement and ECM sentences supports a retrieval cue for case, while the contrast between accusative subjects (ECM) and accusative objects (Object Control) argues against the view that retrieval relies on case alone.

Retrieval cues for both case and position are consistent with the grammatical requirements of English root clauses: a tensed verb or auxiliary requires a nominative subject. Assuming these cues are supplied by the verb, we expect that the interference profile would be different if the critical word for retrieval had different requirements. For example, the infinitive "to" requires its subject to not be in a tensed clause[12]. The experiments reported here were not designed to test this prediction, but the embedded clauses of the ECM and Object Control sentences of Experiments 2–3 provide some supporting evidence. In Experiment 2, ECM sentences showed faster processing at the embedded verb and subsequent word than S-complement sentences. Since S-complement and ECM sentences differed only in the embedded verb (tensed versus untensed, respectively) this could suggest that the cues provided by the embedded infinitive of the ECM sentences mitigate interference from the matrix nominative subject. In Experiment 3, ECM sentences were slower overall than S-complement sentences, but in contrast to the facilitation for complex interveners observed at the matrix tensed verb, there was no corresponding effect of complexity. Furthermore, the Object Control sentences of Experiment 3 showed faster processing at the embedded VP than either ECM or S-complement sentences, though in Experiment 2 this effect was presumably masked by a general slowdown throughout the embedded regions. Overall, then, the contrasting interference profiles between untensed embedded verbs and tensed matrix verbs give credence to the idea that the cues provided by the verb are based on that verb's grammatical requirements. In our materials, though, the effects at the embedded verb were potentially prone to spillover effects from processing complex interveners, and the embedded clauses were not matched for length at this region.

3.9 Conclusions about subject retrieval

The above experiments converge with a rapidly growing body of research indicating that similarity-based interference is a crucial determinant of difficulty in the comprehension of linguistic dependencies. Interference effects are a key prediction of models utilizing content-addressable retrieval, and so the current study also provides evidence that sentence-processing is mediated by

[12] I am grateful to an anonymous reviewer at the *Journal of Memory & Language* for pointing this out.

$$\left[\begin{array}{rl} \text{CATEGORY}^{1}: & DP \\ \text{NUMBER}^{2}: & \textit{Plural} \text{ or } \textit{NULL} \\ \text{POSITION}^{3}: & \textit{spec-T} \\ \text{CASE}^{4}: & \textit{nom}, \textit{acc}, \textit{dat} \\ \text{GENDER}^{5}: & \textit{masc}, \textit{fem} \\ \text{ANIMACY}^{6}: & \pm\textit{animate} \\ \text{X-BAR}^{7}: & \left[\begin{array}{rl} \text{HEAD}: & X_i \\ \text{COMP}: & YP_j \\ \text{SPEC}: & ZP_k \\ \text{MODIFIER}: & \textit{Adj.}, RC_m \end{array}\right] \end{array}\right]$$

1. Gordon, Hendrick, & Johnson (2001)
2. Wagers, Lau, & Phillips (2009)
3. R. L. Lewis & Nakayama (2002), Van Dyke (2007), Van Dyke & R. L. Lewis (2003), and Van Dyke & McElree (2011)
4. Experiments 1–3; Case attraction effects (e.g. Bader, 1997; Bader & Meng, 1999; Logačev & Vasishth, 2012, a.o.)
5. Badecker & R. Lewis (2007)
6. Van Dyke (2007)
7. Assumed following R. L. Lewis & Vasishth (2005)

Figure 3.10: Summary of cues implicated in subject retrieval. Showing a sample subject encoding (left) and selected sources for evidence (right). See preceding sections, and references cited there.

a cue-driven retrieval mechanism operating over content-addressable memory representations. The above studies extend this research by demonstrating that interference effects in subject retrieval are driven by structural information provided by the retrieval context. More specifically, our results indicate that the relevant dimensions of structural similarity (for subjects) involve both structural position and case properties. At the very least, our results indicate that the retrieval cues should distinguish between (i) subjects of a verb and subjects of a nominal; (ii) syntactic subjects and syntactic objects; and (iii) nominative subjects of a finite clause and accusative subjects of a non-finite clause.

Processing implicit subjects 4

4.1 Introduction

4.1.1 Issues in processing implicit subjects

In Chapter 3, we focused on subject-verb dependencies, which are routinely non-adjacent. Consistent with content-addressable retrieval, we saw that subject retrieval is sensitive to the properties of a structurally illicit intervening subject. In particular, we observed difficulty at the matrix verb in sentences containing a complex embedded subject, either an event nominalization or a noun-noun compound. More importantly, the complexity effect interacted with the degree of syntactic similarity between the intervener and the target nominative subject. Processing was inhibited when the intervener matched the target in both case and structural position, but it was *facilitated* in ECM sentences, where the intervener matched the target in position, but differed in case. I argued that this evidence favored a retrieval structure targeting both case and position.

A key component of this reasoning was that intervener complexity seemed to have no effect on the processing of sentences containing an (object) control dependency, where the intervener was distinguished from the target in both case and position. This motivated the position cue to create a context of similarity between subject encodings, so that elaborating a complex subject in the ECM sentences increased the diagnosticity of the case cue. However, the logic was undercut by a somewhat mixed picture of the processing in the control sentences. In Experiment 2, control sentences showed an immediate and persistent effect of complexity at the critical region, relative to ECM sentences. On the other hand, in Experiment 3 the processing profile at the critical and spillover regions appeared robust to interference, with no effect of complexity. In other words, object control sentences patterned with S-complement sentences in Experiment 2, but showed a distinct profile in Experiment 3.

In this chapter we extend our examination of dependency formation to the identification and interpretation of null subjects. The motivation here is twofold. One, we seek to resolve the contrasting interference profiles for the object control sentences of Experiments 2 and 3. Two, we probe for subject interference effects under a different retrieval context, namely that required to interpret an implicit (phonologically null) subject argument. In English, null subjects most commonly occur inside infinitive clauses:

(68) a. The cats seem ___ to be gathering outside.

b. The cats want ___ to go outside.

The infinitive complements of (68) and the adjunct in (69) both involve an empty argument position, which – lacking inherent semantic content – is anaphorically dependent on another element (the *antecedent*) for interpretation. In an effort to be theoretically-neutral, I will refer to these empty argument positions as *gaps*. In (68a), the subject is assigned its thematic role inside the infinitive, and linked to its surface position via movement ('raising'). The sentence in (68b) illustrates a case of *control*, where the lexical properties of the verb *want* dictate that the implicit subject of the infinitive is interpreted as co-referent to the matrix subject. Whereas raising links a single thematic argument with two syntactic positions, control is an anaphoric dependency that links a single overt argument with with the semantic roles assigned by two different predicates.

However, not all implicit subjects are dictated by verbs like *seem* and *want*. Adjunct clauses such as that headed by *before* in (69) also contain an implicit subject, whose antecedent may be grammatically constrained, despite the optionality of the adjunct. Here, the implicit subject must be interpreted as co-referent with the matrix subject:

(69) The cats should eat dinner (in order ___ to take the medicine).

In both (68) and (69), the antecedence relation is signalled by the infinitive marker *to*, which is thus the putative retrieval site. The gaps differ in their predictability; in the complement control sentences (68b) the presence of an implicit subject is predictable from the control verb, whereas in (69) the optionality of the infinitive makes the implicit subject only one possible continuation.

The comprehension of control dependencies can be functionally decomposed into two steps. First, the presence of a gap must be recognized. While the presence of a gap can often be predicted from, e.g., a preceding control verb, the possibility of modifiers, aposatives, etc., means that the ordinal position of the gap cannot be predicted. Second, the gap must be interpreted by identifying its antecedent. This may involve information from several sources, including lexical access, grammatical constraints on co-reference processes, and/or plausibility.

Control dependencies are important for processing theories for at least three reasons. For one, successful comprehension of a sentence requires the system to apprehend information that is encoded explicitly, in (say) overt constituents, but also implicitly, in the form of gaps (Mauner, Tanenhaus, & Carlson, 1995). A routine example comes from question sentences such as (70), where the fronted WH-phrase *which park* must be interpreted as the object of the embedded prepositional phrase. Any model of sentence processing must account for how and when this information is interpreted by the system.

Furthermore, the location of such elements may be implicit, but in some cases their very existence is also implicit. The sentences in (71) show that the antecedent of a gap may itself be an implicit argument, namely the unexpressed agent of a passive.

(70) Which park did you say that we should go for a hike in ___ ?

(71) a. The wheel was spun ___ to win a prize.

b. * The wheel spun ___ to win a prize.

More pertinent to our concerns, the implicit subject of a controlled infinitive lacks any inherent semantic content, and instead depends on another notional argument for its interpretation. By now-familiar logic, the reactivation of a gap's antecedent presumably requires retrieval, at least when the dependents are non-adjacent. Properly identifying the correct antecedent, though, depends on the parser's ability to utilize information from a variety of sources. In particular, the antecedent of an implicit subject is often determined by the lexical semantics of the control verb. Verbs like *promise* or *persuade* require either a subject or object antecedent[1], respectively, while verbs like *beg* allow either subject or object antecedents:

[1] Note that these requirements hold only for active verbs. The situation is famously different for passive forms:

(i) The captain was promised ___ to be allowed to leave quickly.

(ii) The passenger persuaded the captain ___ to be allowed ___ to leave quickly.

See, e.g., Ladusaw & Dowty (1988) and Manzini (1983).

(72) a. The passenger$_i$ *promised* the captain$_j$ ___ to leave quickly.

b. The passenger$_i$ *persuaded* the captain$_j$ ___ to leave quickly.

c. The passenger$_i$ *begged* the captain$_j$ ___ to leave quickly.

Furthermore, Frazier, Clifton, & Randall (1983) observe that correctly identifying an antecedent often depends on information beyond that preceding the gap. Their examples are shown below, in (73). The verb *want* is an object-control verb, at least in its ditransitive form. Assuming that the parser can immediately identify the object gap in the embedded relative clause, then lexical access alone should suffice to correctly identify the antecedent. However, in the case of (73 b), this analysis will turn out to be incorrect, but only after reaching the preposition, several words downstream from the gap.

(73) a. Mary$_i$ is one student who the teacher$_j$ wanted ___ ___ to talk to the principal.

b. Mary$_i$ is one student who the teacher$_j$ wanted ___ to talk to the principal about ___ .

While the issue here involves correctly resolving the filler-gap dependency of the relative clause, the example serves to illustrate that comprehending a control dependency may involve information encoded in multiple levels of representation. Thus, in extending our examination to the retrieval required to identify PRO's antecedent, we also extend our purview to questions of how various information sources – e.g. lexical or statistical – inform the retrieval structure.

The final issue that we address in our examination of control dependencies is the role of predictive mechanisms in the identification of an implicit subject. If identifying the antecedent requires retrieval, and this is accomplished via a content-addressable mechanism, then interpreting PRO is potentially interference-prone. However, as the above discussion implied, the presence of a gap and/or the identity of its antecedent is often predictable from preceding input. Control verbs like *promise* mandate an infinitive argument with an implicit subject ('complement control'). On the other hand, the adjoined infinitive ('adjunct control') in (75) is optional, and so any requirement to retrieve its subject is unpredictable, even when the identity of its antecedent is grammatically-constrained:

(74) *Complement control; Predictable gap*

Shackleton promised the crew ___ to get them home safely.

(75) *Adjunct control; Unpredictable gap*

Shackleton drilled the crew ___ to get them home safely.

Thus, a comparison of different control dependencies can provide insight into how expectations generated in the course of parsing affect the retrieval of the antecedent for an implicit subject. Predictive processes might influence interference effects either because predictable antecedents do not require retrieval, or because they provide additional retrieval cues about the identity of the antecedent, mitigating interference from other consitutents.

4.1.2 Overview of the chapter

The ultimate aim of this chapter is to examine how interference effects are modulated by parser expectations, if at all. In order to get there, though, we first need to fix some assumptions about the grammatical properties of control dependencies, and review previous findings about the kinds of information used to identify the antecedent of an implicit subject. This chapter is structured as follows.

In §4.2, I begin by reviewing the evidence that 'empty categories', including a controlled implicit subject, are psychologically real. Here, the evidence takes two forms. One, end-of-sentence tasks indicate that sentences with a gap show greater reactivation of their antecedents than comparable sentences without a gap. Two, on-line evidence indicates that the presence of a gap leads to additional processing costs at or around the gap site. This is important because if an implicit subject reactivates its antecedent, then presumably this requires retrieval of that antecedent. If this retrieval is content-addressable, then it should give rise to interference effects.

In section 4.4 I review previous work on the processing of control dependencies. We see that the recency of the antecedent matters, a pattern generally symptomatic of memory access/retrieval. At the same time, other evidence suggests that control dependencies are processed 'actively', with expectations about the antecedent generated in advance of any direct evidence from the gap itself. In particular, detection of the gap is guided by lexical information in the control verb. The predictability of the gap serves to de-emphasize the importance of cues based on non-syntactic information, like plausibility, though recency remains a facilitating force. Throughout this section, we also see evidence that a gap is detected very quickly, as diagnosed by (im)plausibility or gender-mismatch effects. Finally, we consider a few studies that have found evidence that the processing at the gap-site is sensitive to the presence of grammatically inappropriate antecedents, consistent with interference effects.

The final sections of this chapter present a series of experiments probing interference in the recovery of PRO's antecedent. These experiments extend our consideration of subject retrieval to implicit subjects – controlled PRO – thereby disassociating the retrieval of a subject from other processes triggered by a verb. In Experiment 4 (§2.7), I consider the role of structural information and intervener complexity in the recovery of PRO's antecedent. The results here are somewhat at odds with the results of previous experiments in a general facilitation for complex interveners. More importantly, we find no evidence of interference driven by structural similarity at the PRO gap. In Experiment 5 (§2.8), I consider the possibility that the lack of interference at PRO is due to predictive processes engaged at the control verb. The experiment examines the effect of a plausibility manipulation of PRO's antecedent and an intervener DP on the processing of lexical control verbs and adjunct control.

4.2 Gaps reactivate their antecedents

I turn now to evidence that empty categories such as PRO and NP-trace are 'psychologically real' in the sense that they modulate processing difficulty. The first piece of evidence comes from end-of-sentence tasks, such as probe recognition. These studies indicate that sentences containing some species of gap show evidence of greater re-activation than similar sentences that contain no gaps. This conclusion is also supported by online studies, which provide more direct evidence that processing a gapped sentence requires resources over and above those required for non-gapped sentences.

The goal of this discussion is to motivate the claim that interpreting a gap, in particular PRO, requires the reactivation of its antecedent. If gaps reactivate their antecedents, then this may require retrieval, which in turn may be interference-prone. We will see that there is good evidence for the psychological reality of empty categories, especially in the case of filler gap questions (WH-trace). However, the picture of PRO that emerges from these studies is less clear. While the evidence generally supports the conclusion that PRO reactivates its antecedent, it also indicates that the processing correlates of this reactivation are less acute than those for NP-trace.

4.2.1 Gaps prime their antecedents

There is now fairly substantial evidence that linguistic gaps engender greater reactivation of their antecedents than comparable sentences without an anaphoric dependency. Much of this evidence comes from probe recognition studies. In this task, participants first read a sentence, after which a probe word is presented. The probe is commonly presented at the end of the sentence, but some variants of the task strategically interrupt the experimental trial to present the probe in order to measure how priming effects unfold in the course of the sentence. The probe is usually a word, either drawn from the previous sentence or an unrelated foil. Participants respond yes or no as to whether they recognize the probe word as having occurred. Probe recognition provides two dependent measures. One, participants' accuracy on the recognition judgement, and two, the latency of their response. Reactivation is reflected in either greater accuracy or faster responses.

Bever & McElree (1988) used a probe recognition task to examine whether gaps reactivate their antecedents. They compared sentences containing gaps to non-anaphoric control sentences without gaps (76). They also included sentences with a pronoun – anaphoric but overt – though they compared it only to the non-anaphoric sentences. Gapped sentences (78) contained an empty argument position created by NP-raising, subject control (PRO), or so-called *tough*-movement constructions:

(76) *Non-anaphoric control*
The astute lawyer who faced the female judge hated the long speech during the trial.

(77) *Pronoun sentence*
The astute lawyer who faced the female judge hoped he would speak during the trial.

(78) Gapped sentences

a. *NP-raising*
The astute lawyer who faced the female judge was certain *t* to argue during the trial.

b. *PRO gap*
The astute lawyer who faced the female judge strongly hoped [PRO] to argue during the trial.

c. Tough-*movement*
The astute lawyer was hard for the judge to control *t* during the very long trial.

The probe for correct trials was the adjective of the antecedent DP, *astute*. If interpreting a gap involves access to its antecedent, then the gapped sentences (78) are expected to show evidence of greater reactivation – either better accuracy or faster responses – than the non-anaphoric control sentences. Relatedly, support for the formal distinctions between the empty categories in (78) would be reflected in significant differences in reactivation among the gapped sentences.

Bever & McElree's results confirmed both of these predictions. Sentences containing a gap showed faster recognition times than non-anaphoric controls, indicating reactivation of the gap's antecedent. Pronoun sentences showed both faster recognition times and greater accuracy than non-anaphoric sentences, suggesting that the facilitation for gapped sentences is not due to (just) the gap itself, but rather its anaphoric requirement for an antecedent. Finally, they also found that NP-raising constructions were significantly more accurate, and marginally faster, than PRO sentences. This last finding is taken to reflect formal distinctions between empty categories, in particular gaps created by movement (NP-raising) and control (PRO). Bever & McElree suggest further that in terms of processing, the antecedent of NP-raising is 'uniquely determined' by phrase-structure, while the relation between PRO and its antecedent is more indirect, and involves co-indexation processes.

These results are generally seen as evidence for the psychological reality of linguistic gaps, in that gaps reactivate their antecedents. One reason to be skeptical of these results, though, is that Bever & McElree (1988) presented the probe only at the end of the sentence. It may be, then, that the observed reactivation is due to semantic or conceptual processing, and not syntactic parsing.

4.2.2 *Recency of probe to gap matters*

Subsequent studies provide more direct evidence that the recency of the antecedent matters. In a related probe-recognition study, McElree & Bever (1989) examined reactivation effects in sentences nearly identical to those of Bever & McElree (1988). Unlike that study, McElree & Bever also manipulated the position of the probe by adding a test point immediately after the gap. I illustrate this below for NP-raising, PRO, and non-anaphor sentences. The probe presentation points are marked with "*".

(79) NP-raising

The stern judge who met with the defense is sure ___ to * argue about the appeal *.

(80) PRO gap

The stern judge who met with the defense adamantly refused ___ to * argue about the appeal *.

(81) No gap

The stern judge who met with the defense flatly rejected the * arguments for an appeal. *

As in the previous study, McElree & Bever found that recognition times for sentences with a gap were faster overall than for non-anaphoric controls. In particular, both NP-trace and PRO showed greater reactivation than non-anaphoric sentences. However, their results extend those of Bever & McElree (1988) in two ways. One, the recency of the antecedent to the probe presentation matters: responses were slower and less accurate at the sentence-final probe position. In contrast, across experiments they observed no differences at the probe position immediately following the gap site.

A second important difference concerns the results for NP-trace and PRO. NP-trace generally showed greater reactivation than PRO. In McElree & Bever's first experiment, sentences containing an NP-trace showed significantly greater accuracy and marginally faster response times than PRO sentences, replicating Bever & McElree (1988). In their second experiment, NP-raising again showed faster response times than both PRO and non-anaphoric controls. Crucially, though, the difference between PRO sentences and non-anaphoric sentences was not reliable. Thus, a controlled PRO sometimes shows greater activation than non-anaphoric sentences, and sometimes not.

More direct evidence for the psychological reality of empty categories for NP-trace and PRO comes from neuro-imaging studies. Featherston et al. (2000) conducted an ERP experiment on the processing of German sentences. They compared sentences containing subject-control, raising, and simple transitives.

(82) [*Der Sheriff*]$_i$ *hoffte* [*als die Witwe plötzlich in das Zimmer kam*] [$_{IP}$ *PRO* [*den Täter*] *endlich verurteilen zu können*]
[the sheriff]$_i$ hoped [as the widow suddenly into the room came] [$_{IP}$ PRO$_i$ [the offender] at-last sentence to can]
"The sheriff hoped, as the widow suddenly came into the room, to be able to sentence the offender at last."

(83) [*Der Sheriff*]$_i$ *schien* [*als die Witwe plötzlich in das Zimmer kam*] [$_{IP}$ t_i [*den Täter*] *endlich verurteilen zu können*]
[the sheriff]$_i$ seemed [as the widow suddenly into the room came] [$_{IP}$ PRO$_i$ [the offender] at-last sentence to can]
"The sheriff seemed, as the widow suddenly came into the room, to be able to sentence the offender at last."

(84) [*Der Sheriff*]$_i$ *verurteilte* [*als die Witwe plötzlich in das Zimmer kam*] [$_{IP}$ t_i [*den Täter*] *endlich im Scheinwerferlicht*]
[the sheriff]$_i$ sentenced [as the widow suddenly into the room came] [$_{IP}$ PRO$_i$ [the offender] at.last in the.spotlight]
"The sheriff sentenced, as the widow suddenly came into the room, the offender at last in the spotlight."

Consistent with the existence of empty categories, Featherston et al. found that both raising and control elicited greater positivities than simple transitive sentences. This effect arose immediately after the gap position, and not at the verb where thematic roles are assigned. This is contrary to the predictions of so-called 'direct-association' models (e.g. M. Pickering & Barry, 1991), where thematic links for dependents are established directly, when the verb is encountered. Instead, Featherston et al. suggest that the parser constructs the (VP-internal) argument gap as soon as the input provides sufficient evidence. Finally, extending the results of Bever & McElree (1988) and McElree & Bever (1989), raising sentences also showed a significantly greater positivity than control sentences at the critical object position (immediately after gap).

Overall, then, end-of-sentence tasks indicate greater reactivation for sentences containing an anaphoric dependency than similar sentences with such a dependency. In particular, participants are faster at recognizing a probe—and sometimes more accurate—when the probe is the antecedent of a gap or pronoun. These effects are consistent with the view that interpreting PRO requires reactivation of its antecedent from memory. However, end-of-sentence tasks cannot provide evidence about the timecourse of these reactivation processes. Indeed, the findings of McElree & Bever (1989) indicate that the recency of the gap to probe-presentation matters: Sentence-final probes show evidence of reactivation, but probes presented immediately after the gap don't. Thus, while supporting the 'psychological reality' of (PRO) gaps, the evidence summarized thus far does not directly bear on questions of when and how a PRO gap is identified and linked to its antecedent, or the nature of the representations involved.

More generally, Janet Nicol & Swinney (1989) point out that end-of-sentence tasks are not a good index of reactivation, because any observed differences in relative activation could be due to either reactivation of the antecedent (as desired) or simply reactivation that arises only from the memory search triggered by the probe recognition task itself. Furthermore, Janet Nicol & Swinney also point out that end-of-sentence tasks cannot provide evidence as to whether the candidate set of reactivated antecedents is sensitive to grammatical constraints. Reactivation of an anaphor's antecedent is consistent with three possibilities: (i) reactivation of *all* preceding DPs; (ii) reactivation of *all and only* those preceding DPs that stand in the appropriate structural configuration to the anaphor; and (iii) reactivation of *only* the grammatically appropriate antecedent DP. Assuming that the antecedent of a gap is reactivated via retrieval, these alternatives can be re-framed as questions about the retrieval cues that are used to identify the target antecedent (see also Chapter 3).

Fortunately, more direct, on-line evidence also supports the psychological reality of gaps, but I will defer discussion of these studies until the next section. Here I focus on a series of (unpublished) on-line studies about reactivation of PRO's antecedent, as reported in Janet Nicol & Swinney (1989). Overall, these studies support the view that the set of reactivated constituents is restricted to those not directly ruled out by grammatical constraints.

Janet Nicol & Lee Osterhout (1988) (reported in Janet Nicol & Swinney, 1989) examined the role of structural constraints on antecedent reactivation in sentences containing a controlled PRO. In a cross-modal priming experiment, they presented sentences containing an embedded ditransitive object control verb with an object gap, *invite*, to sentences containing an embedded transitive control verb like *plan*. The purpose of this manipulation was to vary the linear position of the grammatical antecedent: the matrix object in (85), and the matrix subject in (86).

(85) There is the actress$_i$ that$_i$ the dentist$_j$ from the new medical center in town * had invited t_i PRO$_i$ to * go to * the party.

(86) There is the actress$_i$ that$_i$ the dentist$_j$ from the new medical center in town * had planned PRO$_j$ to * go to * the party with ___ .

The results for the ditransitive sentences (85) were consistent with activation for the correct antecedent only. There was significant reactivation after the embedded verb for *the actress*, while *the dentist* showed no significant priming effects at any probe position. However, in the transitive sentences (*plan*; 86) the activation profiles were the same. In particular, there was no significant priming of the correct antecedent, *the dentist*. Janet Nicol & Swinney (1989) suggest that these patterns indicate that a trace is posited as the direct object of a verb, so that the reactivation of *the actress* is a passive reflection of the WH-trace occurring in the embedded object position. They point out that this explanation also raises a potential worry about the results of Janet Nicol & Lee Osterhout (1988): it is unclear whether the reactivation of *the actress* is due to the presence of PRO or simply the adjacent NP-trace. Thus, the crucial evidence for reactivation of PRO's antecedent will come from constructions where PRO is not adjacent to another (co-referent) gap.

A related cross-modal priming experiment provides evidence of this form (L Osterhout & J Nicol, 1988; again reported in Janet Nicol & Swinney, 1989). L Osterhout & J Nicol presented sentences consisting of a matrix clause containing two potential antecedents, and an infinitive complement clause containing a controlled PRO subject. The position of the correct antecedent was manipulated by the voice of the matrix verb: active verbs (87a) taking an object antecedent, and passive verbs (87a) taking a subject antecedent. Reactivation of potential antecedents was probed at five separate points: immediately after the infinitive, then again at 500 ms, 1000 ms, 1200 ms, and 1500 ms.

(87) *Materials from L Osterhout & J Nicol (1988)*

a. The actress$_i$ invited the dentist$_j$ from the new medical center PRO$_j$ to *1 go to the pa*rty at the * mayor's house.

b. The actress$_i$ was invited by the dentist$_j$ from the new medical center PRO$_i$ to * go to the pa*rty at the * mayor's house.

L Osterhout & J Nicol (1988) found no significant priming for either antecedent at the first two probe points. This is consistent with the findings of McElree & Bever (1989), and again suggests that the recency of probe to gap matters. Significant priming arose only at later positions. At the third probe position, the most recent potential antecedent, *the dentist* – and only the most recent antecedent – showed significant priming effects. By the fourth probe position, significant priming arose only for the grammatically correct antecedent. Finally, at the fifth probe position, only the most distant antecedent, *the actress*, was reactivated. In other words, the results of L Osterhout & J Nicol (1988) are consistent with the view that only the grammatically-appropriate antecedents of PRO are reactivated. At the same time, this targeted reactivation is delayed, and occurs only about 1000 ms after the gap is located.

4.3 Previous work on processing control dependencies

In the following section, I review previous work on the processing of control. Empirically, the focus of this section is on measures of processing effects in sentences containing a controlled PRO subject, either inside a complement clause required by a control predicate (complement control) or inside an optional phrase (adjunct control).

I approach these topics in an effort to answer some basic questions about the processing of control dependencies. To wit: How is the antecedent of PRO identified and re-activated from memory? Is retrieval required, and is it content-addressable (and interference-prone)? If retrieval is required, then we would expect the processing of control to show characteristics of memory access. In particular, the time that intervenes between the encoding of PRO's antecedent and the memory test-point—for instance, the gap—should modulate the accuracy of retrieval processes. In other words, if establishing co-reference between PRO and its antecedent requires retrieval, then the recency of the antecedent should matter.

By the end of this section, I hope to convince you that resolving the identity of PRO's antecedent *does* require retrieval. The question then becomes about the nature of information used to diagnose PRO's antecedent. Here, the evidence strongly indicates that antecedent re-activation is guided by lexical properties of (say) a control verb, so that comprehenders have already made a commitment about the identity of PRO by the time the gap is reached. On the other hand, we will see that some studies have found evidence of interference effects. In particular, intrusion effects suggest that the parser considers potential antecedents that would otherwise be ruled out by the grammar, especially when these distractors are similar to the desired antecedent. The presence of both interference effects and predictive processes opens up the question addressed in the Experiment 5: do we see interference only when expectations are violated?

4.3.1 Recent antecedents are easier than distant antecedents

The Most Recent Filler Strategy

Frazier, Clifton, & Randall (1983) argue that the processes that link a filler to its gap initially ignore grammatical constraints, and instead adopt a 'most recent filler' strategy, where most recently preceding DP is assigned as the antecedent of a gap. This predicts that sentences where the most recent DP is the gap's filler will be easier than comparable sentences where the controlling DP is 'distant', with an incorrect DP intervening between filler and gap.

(88) *The Most Recent Filler Strategy:*

During language comprehension a detected gap is initially and quickly taken to be co-indexed with the most recent potential filler. (Frazier, Clifton, & Randall, 1983, p. 196)

To test this hypothesis, Frazier, Clifton, & Randall used a 'stops making sense' task with a word-by-word presentation. Experimental sentences included two gaps: the subject of an infinitive complement clause, and the object of a later PP-adjunct. In 'recent filler' sentences, the filler of the first gap was the subject of the relative clause, *the little child*, while in 'distant filler' sentences the filler of the first gap was the relative pronoun, which preceded the RC subject:

(89) *Unambiguous sentences:*
   a. *Recent filler*
      Everyone liked the woman$_i$ (who) the little child$_j$ forced ___ i/*j PRO to sing those stupid French songs (after Christmas).
   b. *Distant filler*
      Everyone liked the woman$_i$ (who) the little child$_j$ started PRO to sing those stupid French songs for ___ (after Christmas).

The recent/distant manipulation was crossed with an ambiguity manipulation. Ambiguous sentences replaced the control verb with a predicate like *beg*, which is (locally) consistent with either subject or object control:

(90) *Ambiguous sentences:*
   a. *Recent filler*
      Everyone liked the woman$_i$ (who) the little child$_j$ begged ___ PRO to sing those stupid French songs (after Christmas).
   b. *Distant filler*
      Everyone liked the woman$_i$ (who) the little child$_j$ begged ___ PRO to sing those stupid French songs for ___ (after Christmas).

Since the parser initially ignores grammatical info, by hypothesis, the processing of ambiguous sentences should be the same as the processing of the unambiguous variants, where lexical control info would otherwise serve to determine the location of the filler. The results showed the pattern predicted by the Most Recent Filler Strategy. Distant filler sentences yielded slower response times and a reduced percentage of 'got it' responses than recent filler sentences, independent of ambiguity. Frazier et al. attribute the difference to re-analysis in the distant filler sentences: the parser initially assigns the most recent NP as the filler, forcing a revision when the gap inside the PP disconfirms this hypothesis.

Importantly, there were no significant differences between ambiguous and unambiguous sentences in either response times or the percentage of 'got it' responses—unambiguous sentences were processed as if they were ambiguous. Frazier, Clifton, & Randall explain the difficulty for distant filler sentences as a reflection of the Most Recent Filler Strategy: the most recent DP *the child* is assigned as the filler of the gap, but this commitment must be revised in the distant filler sentences when disconfirmed by the second gap. In other words, non-recent fillers lead to a garden-path effect. Under this view, the parser initially focuses on purely syntactic information, ignoring (e.g.) lexical information, relying instead on recency of potential antecedents[2].

MRFS only in ambiguous sentences

This view was later challenged by several studies indicating that lexical control verb information does play a role, with the MRFS being relegated to an ambiguity resolution strategy. In a self-paced reading task, Crain & Fodor (1985) examined the reading of sentences similar to those of Frazier, Clifton, & Randall (1983), with recent fillers (object control) and distant fillers (subject control), while extdning the comparison to declarative control sentences. Declarative sentences control for effects of frequency, plausibility, etc..

(91) *Unambiguous questions*
- a. *Recent filler:*
  Who could the little child have started ___ to sing for ___ ?
- b. *Distant filler:*
  Who could the little child have forced ___ ___ to sing for Cheryl?

(92) *Declarative controls*
- a. *Recent filler:*
  The little child could have started ___ to sing for Cheryl.
- b. *Distant filler:*
  The little child could have forced us ___ to sing for Cheryl.

[2]Subsequent researchers seem to have routinely mis-characterized this proposal as claiming that *only* the recency of a filler matters. The actual proposal is somewhat more nuanced:

Note that this formulation seems like an early version of activation-based approaches to memory. Thus, we can also reformulate the MRFS as saying that it is easier to re-activate those antecedents associated with stronger traces, with residual activation of recent antecedents giving them relatively more activation.

Following Frazier, Clifton, & Randall (1983), Crain & Fodor also included ambiguous versions of recent and distant gap sentences. As above, these were constructed by using verbs like *beg* which are consistent with either subject or object control. Here again the most recent filler strategy predicts no effect of ambiguity, since the most recent filler should always be assigned to the first gap.

(93) *Locally ambiguous conditions*

a. *Recent filler question:*
Who could the little child have begged ___ to sing for ___ ?

b. *Distant filler question:*
Who could the little child have begged ___ ___ to sing for Cheryl?

c. *Recent filler declarative control:*
The little child could have begged ___ to sing for Cheryl.

d. *Distant filler declarative control:*
The little child could have begged us ___ to sing for Cheryl.

Crain & Fodor extended the design of Frazier, Clifton, & Randall (1983) by controlling for the two adjacent gaps in the distant filler sentences. 'Late-gap' sentences contained a lexical DP in object position, instead of a gap. Since the filler for the first gap in these sentences is the most recent potential filler, the most recent filler strategy predicts that these sentences should be easier than distant filler sentences.

(94) *Late WH-gap conditions*

a. *Unambiguous:*
Who could the little child have forced us ___ to sing NP for ___ PP?

b. *Locally ambiguous:*
Who could the little child have begged us ___ to sing NP for ___ PP?

Crain & Fodor reported their results in terms of an 'increase score', determined by computing the z-scores for each region × participant, then subtracting the z-score for the previous region from the region of interest. These 'increase scores' showed significant effects of ambiguity at both positions, broadly supporting the claim that the initial analysis of the gaps was correct for unambiguous sentences, contra the Most Recent Filler Strategy, but ambiguous setnences showed the garden-path effect predicted by the MRFS.

The results at the first gap position showed that unambiguous distant filler sentences were significantly harder to read than their ambiguous counterparts. This is consistent with a garden-path effect in the ambiguous conditions, due to the initial gap being associated with the most recent (but incorrect) filler. On the other hand, the differential processing of unambiguous sentences suggests that in these sentences grammatical constraints help to correctly identify the antecedent of both gaps. Furthermore, there were no significant differences between distant and recent filler sentences in the unambiguous sentences, which Crain & Fodor see as weakening the evidence that distant-filler sentences were assigned a recent-filler analysis. It seems plausible instead that the primary determinant of difficulty in these constructions is the predictability of the gap's antecedent, and not its position (distant subject or recent object), so that the unpredictable antecedents in the ambiguous conditions engage the most recent filler strategy as an ambiguity resolution strategy.

At the second gap position, the results indicated the expected garden-path effect, but only in the ambiguous conditions. Ambiguous distant-filler sentences were significantly harder than their declarative controls, while there were no significant differences between unambiguous distant-filler sentences and the declarative controls. Independent of the declarative controls, there was no significant effect of ambiguity in the distant-filler sentences at the second gap position. The expected effect did arise at the subsequent adverbial region, though, so that ambiguous sentences were significantly harder than declarative sentences, but unambiguous sentences were not. These results are consistent with a garden-path effect at the later gap position, presumably due to the gap being incorrectly associated with the recent filler, but only in ambiguous sentences where the antecedent could be either subject or object of the higher clause.

Whether the most recent filler strategy is a heuristic or a principle of the parsing architecture, note that recency effects are a hallmark of memory retrieval. Thus, one view of these results is that gap-filling processes respect grammatical constraints when the material preceding the gap is sufficient to uniquely identify its antecedent. When faced with multiple potential antecedents, though, as in the ambiguous conditions, antecedent identification is more shallow, simply binding the candidate closest to the gap. Since the memory trace for a recent antecedent presumably retains some residual activation, a possible explanation for this recency preference is that recent fillers are easier to retrieve.

However, if we are to predict garden-path effects in these constructions, we need some notion of what counts as ambiguous. In the above experiments, ambiguity was manipulated by varying the lexical properties of a control verb, e.g. *beg*. I turn now to more direct evidence that this information does serve to guide antecedent selection, and that processing at the gap site is sensitive to semantic properties of the grammatical controller.

4.3.2 *Lexical information guides antecedent selection*

It seems that the co-reference processes involved in identifying PRO's antecedent are not as myopic as the Most Recent Filler Stategy would predict. Rather, the evidence suggests that lexical information, when available, can be used to generate expectations about upcoming gaps and guide the identification of PRO's antecedent. However, recency can still play a role, especially when the material preceding the gap does not uniquely identify a grammatical antecedent.

Plausability of the grammatical antecedent matters

Boland, Tanenhaus, & Garnsey (1990) present the results of two experiments that used the stops-making-sense task to examine when the lexical information encoded in a control verb becomes available. Using sentences like those in (95), they manipulated the plausability between a matrix argument and the implicit PRO subject of an infinitive complement: *outlaws*, but not *horses*, can surrender.

(95) *Declarative sentences:*

a. The cowboy signalled the outlaw ___ to surrender to the authorities.

b. # The cowboy signalled the horse ___ to surrender to the authorities.

Boland, Tanenhaus, & Garnsey hypothesized that the gap-filling processes are guided by lexical information, e.g. that *signal* is an object-control verb. If identifying PRO's antecedent at the gap is constrained by this lexical information, then comprehenders should be immediately sensitive to the plausibility of the matrix object as the implicit subject, so that implausible sentences are more difficult than plausible sentences.

While a plausibility effect would indicate that expectations about the antecedent are generated prior to the gap site, the sentences in (95) cannot provide direct evidence against the Most Recent Filler Strategy—the object controller is always the most recent potential filler. In order to vary the distance between filler and gap, the plausibility manipulation was crossed with a sentence-type manipulation, comparing the declaratives of (95) to interrogative sentences like those in (96). To the extent that the MRFS can account for plausibility effects, it predicts a plausibility effect when the implausible filler is also the most recent.

(96) *WH-question sentences:*

a. Which outlaw did the cowboy signal ___ to surrender to the authorities?

b. # Which horse did the cowboy signal ___ to surrender to the authorities?

Consistent with lexically-driven parsing, the results of Boland, Tanenhaus, & Garnsey showed that implausible fillers led to significantly more stops-making-sense responses in both declarative and interrogative sentences. This effect arose at the infinitive verb—the earliest gap-detection point—in both the stops-making-sense task and self-paced reading.

A second experiment manipulated the control properties of the verbs, comparing subject and object control crossed with the plausibility of a WH-filler. Object control question sentences from Experiment 1 were compared to subject control sentences like those below:

(97) a. Which outlaw did the cowboy refuse to surrender to the authorities?

b. # Which horse did the cowboy refuse to surrender to the authorities?

Implausible fillers again led to 'no' responses at the infinitive verb following the gap, but only for object control sentences, indicating that control information was already available at the gap site. There was no difference in the timing of plausibility effects between WH-questions and declarative sentences. The RTs for positive responses, though, showed a slowdown for implausible object control sentences only, at the word immediately following the infinitive verb. Since the implausible fillers were only implausible in object control (*horses* cannot surrender, but *cowboys* can) Boland, Tanenhaus, & Garnsey conclude that the verb's control info was used to guide the identification of PRO's antecedent. The slowdown in RTs for object control sentences is taken to suggest that object control is just more difficult than subject control, though the reasons are left open.

Additional evidence that lexical information in the control predicate is used to predict upcoming gaps comes from Boland, Tanenhaus, Garnsey, & Carlson (1995), who again used the stops-making-sense task with self-paced reading to further probe the expectations generated by lexical information in a control verb. Their experiments manipulated the plausibility of a fronted WH-phrase in constructions involving three types of verbs: simple transitives like *visit*, 'Dative ditransitives' like *donate*, and object control verbs like *remind*:

(98) a. Simple transitives:
   Which prize did the salesman visit ___ while in the city?
  b. Dative ditransitives:
   Which charity did the executive donate ___ after meeting the deadline?
  c. Object control:
   Which movie did your brother remind ___ to watch the show?

Again, plausibility effects serve to diagnose the point at which lexical information becomes available (*prizes* cannot be visited). Importantly, though, dative ditransitives and control verbs both provide an additional gap site for the WH-phrase beyond the internal argument (complement) of the simple transitive verbs. Thus, if lexical information serves to constrain the expectations about potential gap sites for the WH-filler, then plausibility effects will be sensitive to the sentence-type manipulation. The additional gap site should delay plausibility effects past the direct-object gap, since the placement of the filler cannot be confirmed at the verb.

The results of Boland, Tanenhaus, Garnsey, & Carlson confirmed these predictions. For sentences with simple transitive verbs, implausible fillers led to 'stops making sense' judgements at the verb, consistent with the active-filler strategy. For sentences with object control verbs, though, the sentences were judged to make sense as long as the implausible filler could be interpreted as one of the verb's arguments, not just as the direct object. Boland, Tanenhaus, Garnsey, & Carlson conclude that the bias to interpret the implausible filler as the direct object (a filled-gap effect) was blocked when the verb provided additional argument positions. The more general claim is that these results favor contraint-based lexical models of comprehension, in that all that is necessary is that comprehenders have immediate access to the argument structure of the predicate.

The gender-mismatch effect

In addition to the rapid sensitivity to the plausibility of construing PRO as coindexed with its potential antecedents, other studies indicate that the identity of a controlled PRO is established quickly enough to evaluate the well-formedness of syntactic dependencies like agreement (mismatches) at the gap site.

For instance, Demestre, Meltzer, et al. (1999) recorded the ERPs of participants listening to Spanish sentences with controlled infinitive complements. The matrix clause always contained two arguments, while the controller of the implicit subject (PRO) was determined by the control verb: subject control for verbs like *quiere* 'want', object control for verbs like *aconsejado* 'advise'. The detection of a gender-agreement violation was used to diagnose whether the parser had correctly established co-reference between PRO and the grammatical controller. While holding constant both the number of matrix arguments and the gender of an embedded adjective, they varied the gender of the grammatical controller:

(99) a. *Pedro quiere ser rico en un futuro próximo.*
Peter.M wants PRO to be rich in the future near.
'Pedro wants to be rich in the near future.'

b. * *Maria quiere ser rico en un futuro próximo.*
Mary.F wants PRO to be rich in the future near.
'Maria wants to be rich in the near future.'

(100) a. *Pedro ha aconsejado a María ser educada con los trabajadores.*
Peter.M has advised Mary.F PRO to be more polite.F with the employees.

b. * *María ha aconsejado a Pedro ser educada con los trabajadores.*
Mary.F has advised Peter.M PRO to be more polite.F with the employees.

Agreement violations were detected quickly, within a few hundred milliseconds of the gap, as evidenced by an N400 and P600 effect in the (ungrammatical) gender-mismatch conditions relative to the matching conditions.

In a subsequent study, Demestre & García-Albea (2007) present ERP evidence from Spanish sentences indicating that agreement mismatches between an adjective of a controlled infinitive clause and its grammatical antecedent (as determined by the control verb) are detected within a few hundred milliseconds. A 2 × 2 experimental design manipulated (i) ditransitive versions of subject and object control verbs, so that the matrix clause contained two potential antecedents, *Peter* or *Mary*; and (ii) the gender agreement between the grammatical controller selected by the verb and a predicate adjective inside the infinitive clause (*estricto* or *ecucada*).

(101) a. *Pedro ha prometido a María ser estricto con los alumnos.*
Peter.M$_i$ has promised Mary.F$_j$ PRO$_i$ to be strict.M with the students.
"Peter/Mary has promised Mary/Peter to be strict with the students."

b. * *María$_i$ ha prometido a Pedro$_j$ ser estricto con los alumnos.*
Mary.F$_j$ has promised Peter.M$_i$ PRO$_j$ to be strict.M with the students

(102) a. *Pedro ha aconsejado a María ___ ser ecucada con la gente.*
Peter.M$_i$ has advised Mary.F$_j$ PRO$_i$ to be polite.F with the people
"Peter/Mary has advised Mary/Peter to be polite with people."

b. * *Pedro ha aconsejado a María ser ecucada con la gente.*
Peter.M$_i$ has advised Mary.F$_j$ PRO$_i$ to be polite.F with the people

Results showed greater positivities for ungrammatical conditions compared to their grammatical counterparts, occurring 500–900 ms after the adjective. Importantly, this effect arose in both subject and object control sentences. This provides further evidence that lexical information in the control predicate guides antecedent selection, since if something like the MRFS were at play then we would expect an early effect of gender mismatch only when the most recent (object) filler mismatched the gender of the adjective.

Interactions of receny & lexical info

Overall, then, initial findings that recent antecedents are easier than distant ones have been challenged by evidence that non-recent antecedents are also considered very quickly, at least when lexically determined. While these competing sources of information are usually considered as alternative explanations, there is evidence that both recency and lexical information are implicated in the re-activation of PRO's antecedent. The view that emerges from these studies is that lexical information is given priority, but in the absence of an antecedent that is strongly constrained by preceding material, non-syntactic information like plausibility or recency can facilitate processing at a PRO-gap.

One such piece of evidence comes from a series of experiments on processing of controlled PRO in Spanish infinitives reported in Moises Betancort, Carreiras, & Acuña-Fariña (2006). Their stimuli were similar to those of Demestre & García-Albea (2007), but used an eye-tracking methodology instead of ERPs. Moises Betancort, Carreiras, & Acuña-Fariña were interested in the role of both thematic information and recency effects in the processing of controlled PRO, and the timecourse of same. In order to determine when readers had made a commitment to PRO's antecedent, experimental sentences contained an adjective inside an infinitive clause, which must match the antecedent of PRO in its gender features.

A sample set of materials used by Moises Betancort, Carreiras, & Acuña-Fariña (2006) are shown in (103). Each sentence consisted of a matrix clause with two arguments (proper names, to signal gender) and a subordinate infinitive clause embedded as the complement to a subject control verb like *promise*. Experimental manipulations varied the gender of the matrix arguments (masc/fem) and the gender of the adjective inside the infinitive:

(103) a. *María prometió a Pedro ___ ser bastante cauta con los comentarios.*
Mary$_i$ promised Peter$_j$ PRO$_{i/*j}$ to be quite cautious.fem with her comments

b. * *María prometió a Pedro ___ ser bastante cauto con los comentarios.*
Mary$_i$ promised Peter$_j$ PRO$_{i/*j}$ to be quite cautious.masc with his comments

c. *María exigió a Pedro ___ ser bastante cauto con los*
Mary$_i$ demanded from Peter$_j$ PRO$_{j/*i}$ to be quite cautious.masc with his
comentarios.
comments
'Mary demanded from Peter that he be quite cautious with his comments.'

d. * *María exigió a Pedro ___ ser bastante cauta con los*
Mary$_i$ demanded from Peter$_j$ PRO$_{j/*i}$ to be quite cautious.fem with her
comentarios.
comments
'Mary demanded from Peter that she be quite cautious with her comments.'

Consistent with the use of lexical information in the control verb, object and subject control sentences were processed differently, even before the gap. Object control sentences were faster than subject control sentences at the matrix direct object region in first-pass, regression-path, and total times. Comprehenders also showed rapid sensitivity to gender agreement between the grammatical controller and PRO. A gender-mismatch effect arose in regression-path times at the PP region, with a slowdown for mismatched sentences[3]. There was also a significant slowdown for object control sentences in total times at the second NP region. The gender mismatch effect was reflected in longer total times for mismatched sentences at the direct object, the adjective, and the PP regions.

In their second experiment, Moises Betancort, Carreiras, & Acuña-Fariña (2006) examined the processing of the controlled PRO of an infinitival adjunct, where there is no control verb to signal the upcoming gap. While both matrix arguments could serve as the controller of PRO, the prepositional head creates a strong preference for the antecedent: *por* introduces an purpose clause with a preference for subject control, while *para* introduces a reason clause with a preference for object control.

(104) a. *Yolanda$_i$ se casó con Jorge$_j$ para PRO1$_i$ tener dinero y*
Yolanda.fem married George.masc in order to have money and
*PRO2$_i$ ser heredera/*heredero de una fortuna.*
be the heir.fem/*masc to a fortune.
'Yolanda married George in order to have money and inherit a fortune.'

[3] A gender-mismatch effect also arose the the PRO-infinitive region, immediately preceding the adjective. While the effect was only significant in the by-participant analysis, the effect is plausibly a preview effect from the following adjective.

b. *Yolanda$_i$ se casó con Jorge$_j$ por PRO1$_j$ tener dinero y*
Yolanda.fem married George.masc in order to have money and
*PRO2$_j$ ser heredero/*heredera de una fortuna.*
be the heir.masc/*fem to a fortune.
'Yolanda married George because he has/had money and is/was the heir to a fortune.'

As in the first experiment, object control sentences were generally easier than subject control sentences. In contrast, though, these differences arose only in later measures: regression-path, re-reading and total times for subject control were slower than object control at the first PRO region. The pattern was reversed at the second PRO region, with subject control showing faster regression path times. Mismatched sentences showed a slowdown in regression path times at the PP region, and in total times at the matrix object, both PRO regions, NP3 (*dinero*), the adjective, and the sentence-final PP.

Overall, the studies of Betancort et al. support the view that lexical information in the control predicate guides antecedent selection. The differential processing of subject and object control—even before the gap—is not predicted by the MRFS. If recency is the only factor, all sentences should be processed as object control sentences. However, when there is no control verb to constrain the choice of antecedent, as in (104), recency does matter: subject control sentences with a distant filler were always harder than object control sentences. Thus, both lexical information and recency play a role in the identification of PRO's antecedent.

The general pattern of these studies is that agreement mismatches are detected more quickly for predictable antecedents, generally taken to indicate that comprehenders have already made a commitment about PRO's antecedent by the time the gap is identified. Similarly, both recency and plausibility of the antecedent can facilitate processing—recent, plausible antecedents are easier—but their effects are attenuated or even eliminated when the preceding context places strong lexical or syntactic contraints on the identity of the antecedent. Taken together, the rapid sensitivity to grammatical violations and its interaction with the predictability of the PRO gap suggests that the parser engages predictive mechanisms that generate expectations about the presence of PRO and the identity of its antecedent.

4.3.3 Predictability matters: the role of ambiguity

Indeed, several studies have found evidence that the predictability of PRO is a determinant of difficulty at the gap-site. Moisés Betancort, Meseguer, & Carreiras (2004) used an eye-tracking methodology to examine the role of plausibility and its timecourse in the processing of Spanish sentences containing controlled PRO. In their first experiment, participants were asked to read sentences with an embedded gerund, which allows the controller to be either the subject or object of the matrix clause. The plausibility of the matrix arguments was manipulated, so that one argument was a more plausible subject of the gerund: *judges* are more plausible agents for *the sentencing* than *lawyers*. In order to evaluate the timecourse of comprehenders' commitment to PRO's antecedent, the agreement morphology of the gerund's adjective was manipulated, so that it either matched or mismatched the gender of the plausible antecedent. The rationale here is that if comprehenders have already made a commitment to the plausible antecedent at the PRO gap, then they should be immediately sensitive to the gender match between the adjective and the plausible antecedent.

(105) *Matched conditions*

a. *La abogada oyó al juez ___ sentenciando muy*
The lawyer.fem$_i$ heard the judge.masc$_j$ PRO$_{j/*i}$ sentencing very
seguro al acusado de homicidio.
confident.masc to-the accused of homicide.

b. *El juez oyó a la abogada ___ sentenciando muy*
The judge.masc$_j$ heard the lawyer.fem$_i$ PRO$_{i/*j}$ sentencing very
seguro al acusado de homicidio.
confident.masc to-the accused of homicide.

(106) *Mis-matched conditions*

a. *La abogada oyó al juez ___ sentenciando muy*
The lawyer.fem$_i$ heard the judge.masc$_j$ PRO$_{j/*i}$ sentencing very
segura al acusado de homicidio.
confident.fem to-the accused of homicide.

b. *El juez oyó a la abogada ___ sentenciando muy*
The judge.masc$_j$ heard the lawyer.fem$_i$ PRO$_{i/*j}$ sentencing very
segura al acusado de homicidio.
confident.fem to-the accused of homicide.

'The lawyer heard from the judge the PRO sentencing of the (one) accused of homicide.'

Initial processing at the gerund, as reflected by first-pass reading times, was easier when the plausible antecedent was recent (object control) than when it was distant (subject control). A gender mismatch effect also arose in total times at the gerund, irrespective of controller recency, which Moisés Betancort, Meseguer, & Carreiras take to indicate that readers had already made commitments about the identity of PRO's antecedent. Their results also indicated that both recency and plausibility play a role in controller selection: regression-path times at both the adjective and prepositional-phrase regions showed an interaction effect, with plausible object sentences easier than plausible subject sentences.

However, the timecourse of the plausibility and recency effects was sensitive to how much the choice of antecedent was constrained by preceding material. A second experiment again used the gender-mismatch effect to diagnose when readers had made a commitment to PRO's antecedent, but used controlled infinitival complements to perception verbs like *see*. In these sentences, and in contrast to the first experiment, there is only one grammatical choice of controller: the matrix object.

(107) Plausible—Match

a. *Al juez lo oyó la abogada PRO*
To-the judge.ACC.MASC heard the lawyer.NOM.MASC PRO
sentenciar muy seguro al acusado de homicidio.
to sentence.INF very confident.MASC to-the accused of homicide.

(108) Implausible—Match

a. *A la abogada la oyó el juez PRO sentenciar*
To the lawyer.ACC.FEM heard the judge.NOM.MASC PRO to sentence.INF
muy segura al acusado de homicidio.
very confident.FEM to-the accused of homicide.

Consistent with the view that comprehenders attempt to resolve PRO's identity as early as possible, Moisés Betancort, Meseguer, & Carreiras again observed a gender-mismatch effect at the adjective, with mismatching sentences read more slowly in first-pass, regression path, and total times. However, whereas the complement control sentences of the first experiment showed a very early effect of plausibility, in sentences like (107) the effect was not observed in first-pass times, but only later measures (total times). Thus, it seems that both plausibility and grammatical constraints play a role in the selection of PRO's antecedent from the earliest stages of processing, so that comprehenders have already made a commitment about the controller by the time the gap is processed. The role of plausibility, though, is delayed in structures where grammatical constraints mandate a unique antecedent. This is consistent with Boland et al's evidence that lexical information about a control verb guides gap-filling processes. The claim is further strengthened by Betancort et al's evidence that the predictability of PRO's antecedent modulates plausibility effects.

These findings indicate that the predictability of a PRO gap matters. Agreement mismatches are detected more quickly when PRO's antecedent is known in advance. Furthermore, while both recency and plausibility can facilitate processing at the gap site, their effects are attenuated or even eliminated for predictable antecedents. Two issues arise at this point. One, if recency effects suggest memory retrieval, then lack of recency effects for predictable antecedents might indicate that, at least in some constructions, identifying PRO's antecedent does not require retrieval (at the gap). Relatedly, the role of lexical information and the fact that grammatical violations are detected more quickly when the antecedent is known in advance suggest that control dependencies can engage predictive mechanisms. Predictive processes here mean that the parser is proactive in resolving control dependencies, and attempts to bind an antecedent to an implicit subject position in advance of any direct ('bottom-up') evidence in the input—what Wagers (2014) calls 'active' dependency formation.

Are control dependencies 'active'?

More direct evidence for predictive processing of control comes from Kwon & Sturt (2014), who examined the question of whether control dependencies are processed actively, and how such processing might interact with other principles of ambiguity resolution. In two eye-tracking experiments, they examined the processing of sentences containing an obligatory PRO, and an optional adjunct clause where PRO is not required.

In their first experiment, Kwon & Sturt (2014) examined the processing of nominal control sentences, manipulating the antecedent of PRO by varying the type of control nominal. In *giver*-control, (109), the antecedent of PRO is the (possessor subject) DP receiving the 'giver' θ-role, while in *recipient*-control the controller is an internal argument that is the 'receiver' of the event. Thus, in (110) only the *giver* antecedent is known in advance of the gap site, and can be resolved immediately, while the recipient argument (*the kids*) follows the gap. Active dependency formation then should trigger a forward search for the unresolved antecedent, predicting a processing cost in *recipient* control sentences relative to the *giver* sentences.

(109) *Giver control:*

Before Andrew's$_i$ refusal PRO$_{i/*j}$ to wash(,) the kids$_j$ came over to the house.

(110) *Recipient control:*

Before Andrew's$_i$ order PRO$_{*i/j}$ to wash(,) the kids$_j$ came over to the house.

To further diagnose active processing, Kwon & Sturt also manipulated the ambiguity of the above sentences, accomplished by removing the comma. This introduces an NP/S ambiguity at first DP of the root clause, *the kids*. Kwon & Sturt assume that the parser follows the principle of Late Closure, so that the ambiguous DP will intitially be parsed as the object of the infinitive adjunct.

Late closure Delay closing constituents. Prefer to attach new material to existing nodes (Frazier & Rayner, 1982b).

If active processing guided by lexical control information is strong enough to override the preference to attach the ambiguous DP as the object, then this should lead to a garden-path effect in *giver*-control, but not *recipient*-control, since in the latter grammatical constraints rule out *the kids* as a potential antecedent of PRO.

Kwon & Sturt found easier processing in the *giver*-control sentences, when the antecedent is known in advance of the gap. In particular, *recipient*-control conditions showed a significant slowdown (and marginally more regressions out) than *giver*-control at the infinitive verb, immediately after the gap. The following regions showed an interaction of control-type and ambiguity, with a garden-path effect for both control types in all measures by the root verb, *came*, but the effect was smaller in the *recipient*-control conditions. The relative difficulty for a post-gap antecedent in the *recipient*-control sentences is taken to reflect a cost for the forward search triggered by the unresolved antecedent. On the other hand, the main effects of ambiguity, coupled with the interaction that attenuated the GP-effect in *recipient*-control, are taken to show that 'control information is not a strong enough cue to prevent mis-analysis, but can aid in the recovery process'.

Kwon & Sturt (2014) also compared *giver*-control sentences to adjunct control structures (112), reasoning that if active processing is used for obligatory dependencies, then the unresolved antecedent of the adjunct PRO should guide the recovery of PRO's antecedent and rule out an analysis where *the kids* is parsed as the object of *wash*. Thus, active processing predicts a garden-path effect in the ambiguous *giver*-control condition but not the ambiguous adjunct control condition.

(111) *Giver*-control

Before Andrew's$_i$ failure PRO$_{i/*j}$ to wash(,) the kids$_j$ came over to the house.

(112) Adjunct control

Before PRO$_{*i/j}$ failing PRO$_{*i/j}$ to wash(,) the kids$_j$ came over to the house.

Their results again indicated that the predictability of PRO's antecedent could aid in recovery from a mis-analysis, even if predictive processes are not strong enough to completely override general parsing preferences. *Giver*-control sentences were more difficult than adjunct control sentences at the ambiguous DP in all measures but regressions out. Both control types showed a garden-path effect at the *to+infinitive* region in regressions out and total times, and was significant in all measures at every subsequent region. Importantly, though, an interaction effect indicated that the cost of ambiguity was greater for *giver*-control sentences than adjunct control sentences.

4.3.4 *PRO requires retrieval: the argument from interference*

I turn now to two studies that suggest that resolving PRO requires cue-driven retrieval of its antecedent. The argument here comes from interference: the presence of grammatically-inappropriate distractors that are similar to the target antecedent modulates processing at or around the gap site.

The first piece of evidence comes from the experiments presented in Sturt & Kwon (2015). Here, the findings indicate that agreement mismatches between a reflexive anaphor and its antecedent are detected later when mediated by either raising or control dependencies that also contain additional distractor constituents.

To establish a baseline, they first examined sentences where reflexive anaphor relations were mediated by raising and control dependencies (113). Here, the matrix DP, *John*, serves as the antecedent for the reflexive, *himself*, but is linked to the antecedent via either a control nominal (*agreement*) or a raising predicate (*seems*). For both raising and control structures, processing at the relfexive showed a slowdown for a mismatching antecedent. The effect was significant in all measures, and importantly, in first-pass times. The early detection of agreement mismatches indicates that comprehenders had made a commitment to an antecedent by the time the reflexive was reached.

(113) a. *Control match/mismatch:*
I was surprised at John's *agreement* to be kind to *himself/*herself* appropriately and very sincerely.
b. *Raising match/mismatch:*
It was surprising that John *seemed* to be kind to *himself/*herself* appropriately and very sincerely.

More importantly for our concerns, Sturt & Kwon (2015) also examined the impact of gender match between the reflexive and intervening, illicit distractor consituents. A sample set of their materials is shown in (114). In a 2 × 2 design, they manipulated the gender of the target matrix subject (match/mismatch; *John* or *Mary*) crossed with the gender of an intervening subject (match/mismatch; *Tom* or *Amy*). If the antecedent is identified by a cue-directed retrieval mechanism, then the processing at the reflexive will be sensitive to the degree of similarity between the distractor constituent and the cues provided at the retrieval site (the reflexive, by hypothesis).

(114) a. *Accessible-match, inaccessible match:*
John's agreement with Tom to be kind to himself was surprising to everyone.
b. *Accessible-match, inaccessible mismatch:*
John's agreement with Amy to be kind to himself was surprising to everyone.
c. *Accessible-mismatch, inaccessible match:*
Mary's agreement with Tom to be kind to himself was surprising to everyone.
d. *Accessible-mismatch, inaccessible mismatch:*
Mary's agreement with Amy to be kind to himself was surprising to everyone.

Their results confirmed predictions of interference, and importantly, showed the profile of an intrusion effect. At the reflexive region, mismatching targets led to a slowdown in go-past, total times, and second-pass (re-reading) times. Crucially, though, within mismatching target conditions, a matching distractor led to significant *facilitation* of processing at the reflexive at the spillover region. At this point, sentences with with a mismatching target but matching distractor were processed on par with (acceptable) sentences with a matching target.

As Sturt & Kwon note, this is consistent with an interference approach: grammatically inappropriate constituents can modulate the processing at the retrieval site, when they are similar to the features of the desired antecedent. Importantly, though, in cases of a mismatching target and matching distractor, processing at the reflexive was facilitated, so that an otherwise unacceptable sentence was processed on par with acceptable sentences—a clear intrusion effect.

The second piece of evidence for cue-driven retrieval of PRO's antecedent comes from Parker (2014), who investigated interference in the licensing of controlled PRO in a temporal adjunct gerund clause. To address interference effects, Parker used an intrusion design (Sturt, 2003), manipulating the grammatical properties of both the grammatical antecedent and an intervening DP. However, while intrusion designs commonly manipulate the gender and/or number of target and intervener, the gerunds used by Parker to introduce a controlled PRO do not typically encode this information (in English), or place grammatical restrictions based on these features. The antecedent of PRO in these constructions is not unconstrained, though: the antecedent must be the subject of the superordinate clause. Crucially, Parker follows (Kawasaki, 1993) in assuming that temporal adjunct PRO is subject to an animacy constraint, so that the antecedent must be an *animate* subject of the higher clause . The animacy preference was confirmed in an acceptability study. Inanimate controllers were still rated as 'relatively acceptable', but there was a significant preference for animate controllers.

Parker (2014) manipulated the animacy of the target subject DP, the target antecedent, and an intervening, structurally irrelevant DP contained inside an object relative clause. A sample set of his materials are shown below. The resulting design was 2 × 2, with factors for *grammaticality* (animacy of the matrix subject) and *attraction* (the animacy of the intervener). Experimental sentences also included a reflexive anaphor inside the gerundive clause, assumed to be an indicator of the interpretation assigned to the implicit subject.

(115) a. *Grammatical, No attraction:*
The doctor that the report described meticulously was certified after debunking the urban myth himself in the new scientific journal.

b. *Ungrammatical, No attraction:*
The discovery that the report described meticulously was certified after debunking the urban myth himself in the new scientific journal.

c. *Grammatical, Attraction:*
The doctor that the researcher described meticulously was certified after debunking the urban myth himself in the new scientific journal.

d. *Ungrammatical, Attraction:*
The discovery that the researcher described meticulously was certified after debunking the urban myth himself in the new scientific journal.

As shown in Figure 4.1, the results at the critical gerund – the hypothesized retrieval site – indicated that the presence of an animate distractor had no impact on the processing of matched sentences, where the target subject fully matched the putative retrieval cues. Sentences where both target and distractor were inanimate showed a general slowdown, reflecting that comprehenders correctly identified the erroneous dependency. However, the unacceptable sentences with an inanimate target but an animate intervener showed a facilitation effect, so that the processing at the gerund in these sentences was on par with the acceptable (animate target) sentences. This pattern is analogous to the intrusive licensing of agreement and NPI dependencies, which has formed an integral argument for interference effects in language processing (Wagers,

E. F. Lau, & C. Phillips, 2009; Drenhaus, Saddy, & Frisch, 2005).

4.4 Experiment 4

4.4.1 Design & predictions

Experiment 4 was designed to examine the effect of intervening material on the processing of sentences containing an infinitive complement whose implicit subject, PRO, was controlled by the matrix structural subject. The experimental design manipulated the complexity of an intervening constituent as well as its structural similarity to the target of retrieval. If interpreting PRO requires a retrieval of its antecedent, then we expect that the complexity of intervening DPs will modulate the accuracy of this process (Hofmeister, 2011). Furthermore, if antecedent retrieval is required, and guided by the syntactic properties of the target, then the processing required for PRO's interpretation will be more difficult when intervening DPs are similar to the controller, the matrix subject.

Experimental manipulations varied the properties of a DP intervening between the PRO gap and the controller by including a subject-attached relative clause hosting the intervener. As shown in Table 4.1, experimental design was 2 × 2, with fully crossed factors for intervener *complexity* (simple vs. complex) and intervener *structure* (embedded subject vs. embedded object).

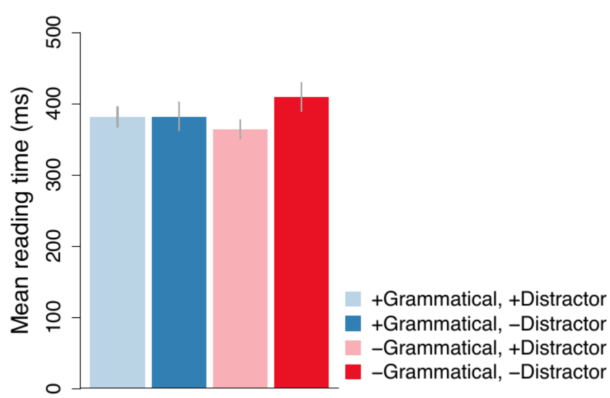


Figure 4.1: Parker (2014) Chapter 3, Experiment 2 results, showing mean reading times (and standard errors) at the gerund verb.

The complexity manipulation compared singular definite DPs such as *the guide* to noun-noun compounds with an adjective modifier, *the quiet forest guide*. The structural similarity manipulation varied the position of the intervener by including it within either an object-extracted relative clause (ORC) with an overt intervening subject to subject-extracted relatives (SRC) with an object intervener.

There were two critical regions, each of which corresponds to a putative retrieval site. The first critical region is the matrix auxiliary verb, where the structural subject is retrieved and integrated with the verb. The second critical region is the infinitive *to*, where the PRO gap is confirmed and its antecedent is identified. Based on previous research (Hofmeister, 2011) and the results of Experiments 1–3, we expect that complex interveners will affect the accuracy of retrieval processes by increasing the distinctiveness of the complex constituent's features. If the intervener is similar to the target, then complexity will make the intervener more intrusive by increasing the associative match between intervener and the cues at the retrieval site, leading to increased difficulty in the complex conditions. On the other hand, if intervener and target are dissimilar, then complexity will make the intervener more distinct, thereby reducing retrieval interference in complex conditions.

Interference due to syntactic similarity between the intervener and the target matrix subject will be reflected in the effect of the structure of the relative clause. If the retrieval at either critical

RC-Type	Sample sentence
Simple intervener conditions:	
SRC	The tourist that ___ helped the guide with travel plans was allowed by the stern judge to leave after the outbreak.
ORC	The tourist that the guide helped ___ with travel plans was allowed by the stern judge to leave after the outbreak.
Complex intervener conditions:	
SRC	The tourist that ___ helped the quiet forest guide with travel plans was allowed by the stern judge to leave after the outbreak.
ORC	The tourist that the quiet forest guide helped ___ with travel plans was allowed by the stern judge to leave after the outbreak.

Table 4.1: Sample materials set for Experiment 4.

region is guided by the syntactic properties of the target, then ORC conditions will be more difficult than SRC conditions, since in ORC conditions both the intervener and target are similar in being structural subjects. To the extent that the distinctiveness effect depends on (syntactic) similarity between intervener and target, then we expect an interaction of structure × complexity, such that complex interveners are more difficult when they are a structural subject like the target (complex ORC sentences).

The final set of predictions concerns potential differences between the retrievals required in the course of parsing, like attaching a verb to its subject, and those required for the interpretation of PRO. If interpreting PRO's antecedent requires a content-addressable access of memory, then the effects at both the PRO regions and the matrix verb regions will be as outlined above. On the other hand, the results of previous experiments suggest that the identification of a PRO gap and the interpretation of its antecedent benefit from lexical information in the control verb. It is possible, then, that identifying PRO's antecedent (i) does not require retrieval, or (ii) that retrieval is required, but involves a structured search of previous material. These alternatives are addressed in Experiment 5, but in the current experiment either will be reflected in contrasting profiles for the critical regions, with PRO showing sharply reduced sensitivity to interference.

4.4.2 *Method*

Participants

95 participants completed the online experiment after responding to advertisements on social media. All participants were native English speakers. To verify native fluency, a pre-experiment form included two questions: "Please list the languages you learned before the age of 10 years" and "Are you a native English speaker?". Participants received gratitude for completing the experiment, but were otherwise not compensated.

Materials

A full sample item set is shown in Table 4.1. The full set of materials for Experiment 4 are included in Appendix A. Experimental materials consisted of 24 item sets instantiating a 2 × 2 factorial design, crossing factors for intervener *complexity* (simple/complex) and *structure* of the relative clause (SRC/ORC). Experimental sentences were arranged in a Latin Square design and combined with 48 filler sentences of a similar length, so that the ratio of relative clauses and/or control verbs to unrelated sentences was 1:1 across the experiment. Excluding practice trials, there were 72 sentences total, which were then combined into four pseudo-randomized lists.

Experimental sentences consisted of a passive matrix clause with an overt *by*-phrase subject followed by an infinitive complement clause. The matrix main verb was always an Object Control verb, such as *persuade*, so that the controller of the infinitive's implicit subject was always the matrix structural subject.

A complex matrix subject contained a subject-attached relative clause that intervened between the matrix subject and verb. This relative clause always contained a gap. In subject-relative (SRC) conditions, the subject was co-referent with a gap in the embedded subject position, while in object-relative (ORC) conditions the subject was co-referent with a gap in the embedded object position. The overt argument of the relative clause was the potentially interfering intervener. In simple conditions, this DP was a singular definite DP, *the guide*. In complex conditions, the intervener was an adjective-modified noun-noun compound, *the quiet forest guide*. In all conditions, the relative clause ended with a prepositional modifier, chosen to favor VP-level attachment. This PP was intended to serve as a buffer between the gaps in the relative clause and the matrix verb regions (the first critical region).

The matrix verb regions consisted of an auxiliary verb followed by the main verb, which was always a passive object control verb. Object control verbs were chosen from a list of 12 verbs, inspired by Boland, Tanenhaus, & Garnsey (1990): *allow, advise, ask, convince, encourage, force, invite, permit, require, send, teach, tell.* Each verb appeared twice in experimental item sets. The control verb was followed by an agentive *by*-phrase PP, which was always of the form *the-Adj-*

Noun. This PP serves as both a buffer between the critical gap regions, and as spillover for the first critical region.

The *by*-phrase was immediately followed by an infinitive complement clause. The infinitive contained the T(ense) head, *to*, followed by an intransitive verb. Note that at this point the input is sufficient to confirm the existence of the gap for PRO. Subsequent to the infinitive verb, a sentence-final PP served as a spillover region.

Procedure

Analysis

Ibex data was imported to the R programming environment for statistical computing (R Core Team, 2016). 33 participants were removed for extremely low comprehension accuracy (<65%). Reading times less than 50 ms and greater than 2000 ms were removed, which excluded 1.1% of the data. Residual reading times were computed from filler data, using a procedure identical to that described in Experiments 1 and 2 (see Ch.2, §4). Z-scores were computed for residual reading times, by word-region and condition. Observations whose z-score exceeded |2.5| were removed. This exclusion affected 2.93% of the data.

The remaining residual RT data was analyzed in regions consisting of a single word. Residual RTs at each region were fit to a series of linear mixed-effects models using the *lme4* package in R (D. Bates et al., 2013). Models contained fixed-effects for experimental factors, coded using deviation coding (-.5, .5), with object-relative and complex intervener levels as the positive coefficients. All models were 'maximal' in the sense of Barr et al. (2013), with by-participant and by-item random intercepts and slopes for experimental design. See Experiment 1 for discussion of non-convergent models. Significance of experimental factors was evaluated by treating the coefficient t-score as a z-score, taking those coefficients whose t/z-score exceeds |2| (Gelman & Hill, 2007; Baayen, Davidson, & D. M. Bates, 2008).

Comprehension question accuracy was modeled using logistic regression (Jaeger, 2008). Contrast coding for these models was identical to that described in the RT analysis. Since comprehension questions did not vary within item sets, these models contained only random intercepts by item.

Unless otherwise noted, I report only significant coefficients (β) and their standard errors (S.E.).

4.4.3 *Results*

Comprehension accuracy

Mean comprehension accuracy for each condition are shown in Table 4.2. There were no significant effects on accuracy for experimental factors. Models fit to response times for comprehension questions revealed a significant interaction of complexity and structure, such that complex interveners led to significantly longer response times in ORC structures (β = 401 ms, s.e. = 179 ms).

	Complex	Simple
ORC	75	77
SRC	78	78

Table 4.2: Experiment 4 comprehension accuracy means (% correct) by condition.

Reading times

Residual reading times for Experiment 4 are summarized in Figure 4.2. Overall, residual reading times indicated a recurring facilitation for sentences containing complex interveners. The processing at the verb showed fleeting effects of structure, with greater difficulty in the object relative conditions (ORC) compared to subject-extracted relatives (SRC).

R17–R21: Matrix verb & *by*-phrase regions.— There were no significant effects at the first critical region (matrix auxiliary and main verb), but there were delayed effects of both structure and size at the PP phrase immediately following the matrix verb regions.

A significant effect of structure arose immediately after the main verb, at the prepositional head, reflecting greater difficulty in ORC conditions compared to SRC conditions (β = 11 ms, s.e. = 5 ms). This difficulty was resolved by the determiner of the *by*-phrase, where there were no

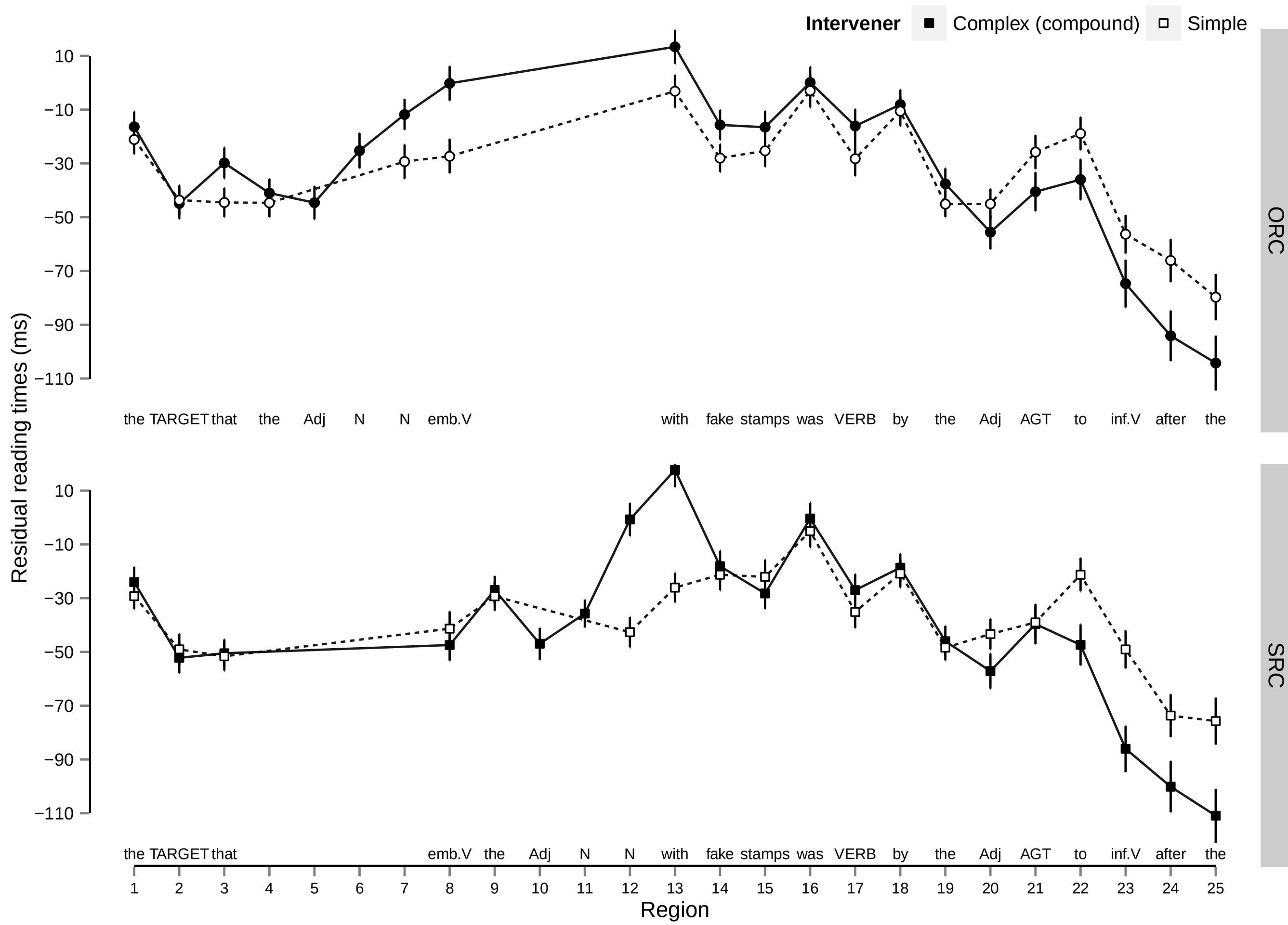


Figure 4.2: Experiment 4 residual reading times by condition, aggregating over participants and items. Solid lines and filled squares indicate complex intervener conditions, dotted lines and open squares indicate simple intervener conditions. Error bars show standard error of the mean.

significant effects for experimental factors. A significant effect of size arose at the adjective of the *by*-phrase, reflecting a facilitation for complex intervener conditions relative to simple intervener conditions (β = -14 ms, s.e. = 6 ms).

R22–R26: PRO & spillover regions.— At the second critical region, the infinitive regions, there were significant effects of size reflecting faster residual reading times for sentences with complex interveners. The facilitation for complex interveners arose at the infinitival *to* (β = -26 ms, s.e. = 8 ms) and the infinitive verb (β = -33 ms, s.e. = 8 ms).

The facilitation for complex interveners relative to simple intervener conditions also persisted through the spillover region. Complex intervener conditions were read faster at the spillover preposition (β = -32 ms, s.e. = 8 ms), the spillover determiner (β = -33 ms, s.e. = 7 ms), and the spillover head noun (β = -30 ms, s.e. = 9 ms).

R1–R3: Target regions.— A significant effect of structure arose at the initial three word regions, consisting of the target subject and complementizer *that*. There were no effects at the determiner or head noun of the target subject. There was a significant effect of structure at the complementizer *that*, reflecting a slowdown for ORC structures (β = 13 ms, s.e. = 5 ms). Note, however, that there were no experimental differences between the conditions at this point, and so I take this effect to be spurious.

R8–R15: Embedded VP regions.— Residual reading times inside the relative clause regions indicated significant effects of size, structure, and their interaction. At the embedded verb, a significant effect of structure reflected a slowdown for ORC conditions relative to SRC conditions (β = 30 ms, s.e. = 7 ms). The interaction of complexity and structure was also significant at the embedded verb, such that the difficulty for complex interveners was more acute in ORC conditions (β = 33 ms, s.e. = 16 ms). Pairwise comparisons revealed that this interaction was driven by a significant effect of complexity only in ORC sentences with a subject intervener (β = 27 ms, s.e. = 10 ms), and not SRC sentences. Pairwise comparisons also showed a significant effect of structure in the complex intervener conditions, such that complex ORC sentences were

	Complex interveners		Simple interveners	
Region	Obj. Gap	Subj. gap	Obj. gap	Subj. gap
1	-16 (6)	-24 (6)	-21 (5)	-29 (5)
2	-45 (6)	-52 (6)	-44 (5)	-49 (6)
3	-30 (6)	-50 (5)	-45 (5)	-52 (5)
4	-41 (5)		-45 (5)	
5	-45 (6)			
6	-25 (7)			
7	-12 (6)		-29 (6)	
8	0 (6)	-47 (6)	-27 (6)	-41 (6)
9		-27 (5)		-29 (5)
10		-47 (6)		
11		-36 (5)		
12		-1 (6)		-43 (6)
13	13 (6)	18 (6)	-3 (6)	-26 (6)
14	-16 (5)	-18 (6)	-28 (5)	-21 (6)
15	-17 (6)	-28 (6)	-25 (6)	-22 (6)
16	0 (6)	0 (6)	-3 (6)	-5 (6)
17	-16 (6)	-27 (6)	-28 (7)	-35 (6)
18	-8 (5)	-19 (5)	-11 (5)	-21 (5)
19	-38 (6)	-46 (6)	-45 (5)	-48 (5)
20	-56 (6)	-57 (6)	-45 (6)	-43 (6)
21	-41 (7)	-40 (7)	-26 (6)	-39 (6)
22	-36 (8)	-47 (8)	-19 (6)	-21 (6)
23	-75 (9)	-86 (9)	-56 (7)	-49 (7)
24	-94 (9)	-100 (10)	-66 (8)	-74 (8)
25	-104 (10)	-111 (10)	-80 (9)	-76 (9)
26	-33 (13)	-29 (13)	-6 (11)	2 (12)

more difficult than complex SRC sentences (β = 46 ms, s.e. = 11 ms).

The effect of size and its interaction with structure also arose at the preposition signaling the end of the relative clause. The effect of complexity was significant at the embedded preposition, reflecting a slowdown for complex interveners (β = 29 ms, s.e. = 7 ms). The interaction of complexity and structure was also significant at the embedded preposition, such that complex interveners *facilitated* reading times in the ORC structures (β = -27 ms, s.e. = 12 ms). Pairwise comparisons revealed that the interaction was driven by a significant effect of complexity in the ORC conditions (β = 15 ms, s.e. = 8 ms), and a significant effect of structure in the simple intervener conditions (β = 23 ms, s.e. = 8 ms). There were no significant complexity effects in SRC conditions, and no effects of structure in complex intervener conditions.

4.4.4 Discussion

We set out to probe for interference effects in the reading of sentences containing a controlled infinitive complement. Experimental design manipulated the complexity of an intervening argument and its syntactic similarity to PRO's controller, the matrix subject. The rationale for this design was that if interpreting PRO requires retrieval of its antecedent, then the accuracy of this retrieval will be modulated by the relative complexity and/or similarity of a more recent, but grammatically-inappropriate, argument. I assume that any effects on accuracy will be reflected as differences in either reading times or comprehension questions. The reading times results showed that complex interveners facilitated reading times at regions immediately following the two hypothesized gap positions, the matrix verb and the implicit subject of the infinitive. On the other hand, complex interveners inhibited the response times for comprehension questions, but only when the complex intervener and the controller were structurally similar, and both subjects. In contrast, the effects of syntactic similarity between intervener and controller was relatively fleeting, arising at only one word in the agentive *by*-phrase. Consistent with an elaboration account, these results indicate that increasing constituent complexity can modulate the accuracy of retrieval processes by increasing distinctiveness and making it easier to discriminate between candidate encodings. I take these results to suggest that interpreting PRO, and in particular identifying its antecedent, requires retrieval of a co-referent argument from memory.

However, the results of the current experiment are somewhat at odds with previous findings, in that here complex interveners led to a facilitation effect at the gap site, regardless of structural similarity to the target controller, and thus does not appear to make intervening arguments more intrusive. Furthermore, in contrast to Experiments 2–3, the complexity effect was not sensitive to the syntactic position of the intervener, at least in reading times measures. All four experiments are consistent with an interference approach, in that the properties of a grammatically-inappropriate constituent modulate processing at the points where we expect retrieval to occur. The current experiment differs in that here the complexity effect is uniformly facilitatory (cf. Experiment 1), and insensitive to the syntactic similarity between the complex constituent and the target of retrieval (cf. Experiments 2–3). One potential source of the contrasting profiles is the fact that in the current experiment the potential interferer was always the co-argument of a gap. If the WH-gap of the relative clause also reactivates its antecedent, then the intervener will be less intrusive to the extent that this reactivation makes co-arguments more distinctive. An additional difference is that in the current experiment, the identity of PRO's antecedent is predictable from the lexical information in the control verb. Thus, recognition of the control verb may lead the parser to expect a PRO gap and create an expectation about the identity of its antecedent, which might serve to mitigate interference from structurally similar constituents.

Nonetheless, the current experiment indicates that the parser attempts to establish the identity of PRO's antecedent quickly, around the time that the gap is confirmed. Sentences containing a noun-noun compound as the embedded argument were read faster at both the agentive *by*-phrase following the matrix verbs, as well as in the infinitive regions immediately following the PRO gap. While the reading times gave no indication that the complexity effect was modulated by structural similarity, the response times for comprehension questions did show evidence of greater difficulty for complex interveners that were structurally similar to the target of retrieval. The fact that the complexity the intervener matters is consistent with interference from grammatically-inappropriate constituents, while the fact that the complexity effects arose at or immediately after the putative retrieval sites suggests that complexity affects the accuracy of retrieval processes.

Taken together, these results indicate that the complexity of intervening arguments exerts an immediate influence on the processing at a gap site, but the structural properties of that intervener don't. This result is consistent with the hypothesis that PRO's antecedent is reactivated at the gap via retrieval of its antecedent from memory. However, the materials of the current experiment cannot rule out a potential benefit from the predictability of PRO's antecedent given access to the control verb. This possibility is examined in the next experiment.

4.5 Experiment 5: The role of predictability

In Experiment 4, complex interveners facilitated processing at both the matrix verb and the implicit subject of a controlled infinitive clause. While these results are at odds with the interactions of complexity and structural similarity observed in Experiments 1–3, they are broadly consistent with an interference approach. The processing at the retrieval site was sensitive to the properties of structurally illicit constituents, a hallmark of cue-driven interference effects.

Experiment 5 uses the logic of intrusive interference to examine whether interference in the identification of PRO's antecedent is sensitive to the predictability of the upcoming gap. Previous work suggests that plausability and recency are more important cues to the correct antecedent when the gap is unpredictable (Moises Betancort, Carreiras, & Acuña-Fariña, 2006, e.g.). If resolving the controller of a gap is interference-prone, then we expect an animate, but structurally illicit, controller to facilitate processing at the gap. If, furthermore, the predictability of an upcoming gap matters, then interference should be greater in adjunct control sentences than complement control sentences, since only the latter gap is predictable prior to the gap site.

4.5.1 *Design & predictions*

Experiment 5 compared sentences containing two types of control dependencies: complement control and adjunct control. The critical region is the infinitive *to* or gerund, the earliest point at which a gap can be confirmed, and the putative trigger for retrieval of PRO's antecedent:

(116) *Complement control*

The officer$_i$ that the city$_j$ respected ___ was advised ___ delicately PRO$_{i/*j}$ to retire from service.

(117) *Adjunct control*

The officer$_i$ that the city$_j$ respected ___ was disgraced ___ after PRO$_{i/*j}$ retiring from service.

In complement control sentences the presence of a gap is predictable from the control verb, e.g. *advise*. The implicit subject of the gerund headed by *retiring*, though, is not required, and the adjunct is optional. Thus, if the predictibility of the gap matters, then the processing at the gap should be sensitive to the type of control dependency, giving rise to a main effect of control type.

We also vary the animacy of the matrix subject, the grammatical controller, so that detection of the illicit controller – a main effect of controller animacy – indexes the point at which readers have made a commitment as to PRO's antecedent. An animate controller is required for adjunct control, but there is no such restriction for complement control. To create a strong bias against inanimate controllers, the infinitive verb of complement control sentences was selected to require an animate subject, e.g. *the officer… retired* versus *#the precinct… retired*. The preceding verbs were designed to be consistent with either DP, regardless of animacy (e.g., *the precinct was advised*).

(118) *Complement control*

* The precinct$_i$ that the city$_j$ respected ___ was advised ___ delicately PRO$_{i/*j}$ to retire from service.

(119) *Adjunct control*

* The precinct$_i$ that the city$_j$ respected ___ was disgraced ___ after PRO$_{i/*j}$ retiring from service.

In order to probe for interference effects, Experiment 5 used an intrusion design (Sturt, 2003), manipulating the animacy of the matrix subject – the grammatical controller, shown above – and the subject of an intervening object-relative clause:

(120) a. *Complement control*

* The precinct$_i$ that the mayor$_j$ respected ___ was advised ___ delicately PRO$_{i/*j}$ to retire from service.

b. *Adjunct control*

* The precinct$_i$ that the mayor$_j$ respected ___ was disgraced ___ after PRO$_{i/*j}$ retiring from service.

The crucial sentences are those with an inanimate controller but an animate intervener. In this configuration, the target subject is a poor controller, due to the mismatching animacy feature, but the (structurally illicit) intervener is both a subject and animate. If resolving the identity of PRO requires cue-directed retrieval, and this is guided by animacy requirements of the critical verb (Van Dyke, 2007), then the partially-matching intervener should facilitate processing at the critical region due to misretrieval of the intervener, at least on some proportion of trials. Furthermore, the attraction effect should arise only in unacceptable sentences where the target does not fully match the retrieval cues (Wagers, E. F. Lau, & C. Phillips, 2009).

4.5.2 Method

Materials

A sample set of materials is shown in Table 4.3. Experimental materials consisted of 40 item sets implementing a 2 × 2 × 2 experimental design, with factors for *control-type* (complement/adjunct control), *grammaticality* (± animate target subject), and *distractor* (± animate embedded subject). Sentences consisted of an intransitive matrix clause containing a subject-attached relative clause, followed by an infinitive clause, either a complement clause (lexical control verb) or a gerund

Condition	Sample sentence
	Complement control conditions
+gramm., -intrusion	The student that the college praised highly was allowed finally to sing the anthem at the ceremony.
-gramm., -intrusion	The library that the college praised highly was allowed finally to sing the anthem at the ceremony.
+gramm., +intrusion	The student that the president praised highly was allowed finally to sing the anthem at the ceremony.
-gramm., +intrusion	The student that the president praised highly was allowed finally to sing the anthem at the ceremony.
	Adjunct control conditions
+gramm., -intrusion	The student that the college praised publicly was honored after singing the anthem at the ceremony.
-gramm., -intrusion	The library that the college praised publicly was honored after singing the anthem at the ceremony.
+gramm., +intrusion	The student that the president praised publicly was honored after singing the anthem at the ceremony.
-gramm., +intrusion	The library that the president praised publicly was honored after singing the anthem at the ceremony.

Table 4.3: Sample materials set for Experiment 5. Experimental design was 2 × 2 × 2, with factors for control-type, controller animacy, and intervener animacy.

(adjunct control). Sentences were designed such that the matrix and embedded verbs would allow inanimate subjects, so that animacy violations would only arise at the infinitive verb, after the gap.

In complement control conditions, the verb of the matrix clause was the passive form of an object control verb, so that the antecedent of the infinitive's PRO subject was always the matrix structural subject. In adjunct control structures the matrix verb was a simple intransitive verb. The matrix verb was separated from the following infinitive clause by an adverb in complement control conditions, and by a preposition in adjunct control conditions. This preposition was always *while*, *before* or *after*, so that the infinitive was a temporal adjunct.

The 40 target items were arranged in a Latin Square design and pseudo-randomly interspersed with 60 filler sentences, for a total of 100 sentences with a 1:1.5 ratio of experimental items to fillers. 10 fillers contained an animacy-violating dependency, so that 50% of the sentences were unacceptable. Half of the sentences presented (and all of the experimental items) were followed by a comprehension question. This was done to both discourage superficial reading strategies and provide another measure of potential interference effects, at least at later measures.

Participants

47 participants completed the experiment. Participants of the Linger study were members of the University of California, Santa Cruz, community, and received course credit for participation. Participants of the online study were demographically more diverse, but all were native English speakers from North America. Compensation for the online study was $8.00 USD.

Procedure

Experiment 5 used a self-paced reading methodology, with a word-by-word moving window presentation. Each trial sentence began with the words of the sentence masked by dashes. Participants advanced through the sentence by pressing the space-bar key, revealing the next word and re-masking the previous word. Half of the trials – and all of the experimental trials – were followed by a yes/no comprehension question. Participants answered the question by pressing the 'f' key for *yes* and the 'j' key for *no*. Feedback was presented for incorrect answers only. After answering the question, participants were cued to begin the next trial by pressing any key.

The experiment was conducted using both Linger (Doug Rohde, MIT) and Amazon Mechanical Turk. Other than the experimental setting, there were no differences in task or instructions.

Analysis

Reading times and comprehension data from both Linger and Ibex were imported into the R statistical programming environment for analysis (R Core Team, 2016). Reading times were analyzed as follows. Extreme values were excluded by removing .5% from the tails of the distribution of raw reading times (1% of total observations). The remaining RT data was analyzed in regions consisting of a single word. Residual reading times were computed, using a procedure identical to that described in previous experiments (see §3.4). Prior to analysis, seven participants were removed for extremely low comprehension accuracy (less than 65%).

Significance testing was accomplished by fitting the RT and comprehension data to a series of multi-level regression models using the `lme4` package in R (D. Bates et al., 2013). Contrast coding for experimental factors used a deviation coding scheme (.5, -.5), so that coefficients reflect deflections from the grand mean, and can be interpreted in a manner analgous to traditional ANOVA analyses. The positive coefficients were complement control, animate target and animate distractor. Unless otherwise noted, all models contained the 'maximal' random-effects structure, allowing both the slopes and intercepts for experimental effects to vary by both subject and item. In cases of non-convergence, model random effects structure was simplified by inspecting the non-convergent model and (i) removing correlation parameters, by subject and then items, and failing that (ii) dropping the slope associated with the lowest variance. In practice, the first term removed was always an interaction term. Outliers were removed after model fitting, using the procedure described in Baayen & Milin (2010): we first fit a model to the data at the region of interest, then remove observations associated with residuals that exceed $|2.5|$, and re-fit the model.

Comprehension accuracy data was modeled using a series a logistic mixed-effects models, using the same coding scheme for experimental factors as the reading times models. Since comprehension questions did not vary within items, and are thus within-items, comprehension models included only random intercepts for items. Following Gelman & Hill (2007), I assume that significant effects in both RT and accuracy models are those coefficient estimates whose *t/z*-score was greater than |2|.

4.5.3 *Results*

Residual reading times

Residual reading times are summarized in Table 4.3. Overall, residual reading times indicated greater difficulty in the infinitive regions for adjunct control sentences compared to complement control sentences. Inanimate target sentences also led to greater difficulty in these regions,

The effect of structure.— Complement control sentences were read significantly more slowly than adjunct control sentences at the adverb/preposition following the matrix verb (region 10; β = 15 ms, s.e. = 7 ms), and the first word of the infinitive (region 11; β = 12 ms, s.e. = 6 ms). Structural differences reversed at subsequent regions, with adjunct control sentences being read more slowly than complement control sentences at regions 12 (β = 14 ms, s.e. = 4 ms), 14 (β = 11 ms, s.e. = 4 ms), and 16 (β = 72 ms, s.e. = 10 ms).

Target & distractor animacy.— The only significant effects before the matrix verb occurred at region 7, the embedded RC verb, where sentences with an inanimate distractor were read more slowly than sentences with an animate distractor (β = 30 ms, s.e. = 11 ms). The results at this region also showed an interaction, with sentences where both target and distractor were animate were significantly slower than sentences containing either two inanimates or a mixture of animacies (β = 36 ms, s.e. = 16 ms).

A significant effect of target animacy arose at the infinitive verb (region 12). Consistent with a detection of grammaticality at this region, inanimate target sentences were significantly slower than animate target sentences (β = 9 ms, s.e. = 4 ms). Target animacy also exerted significant

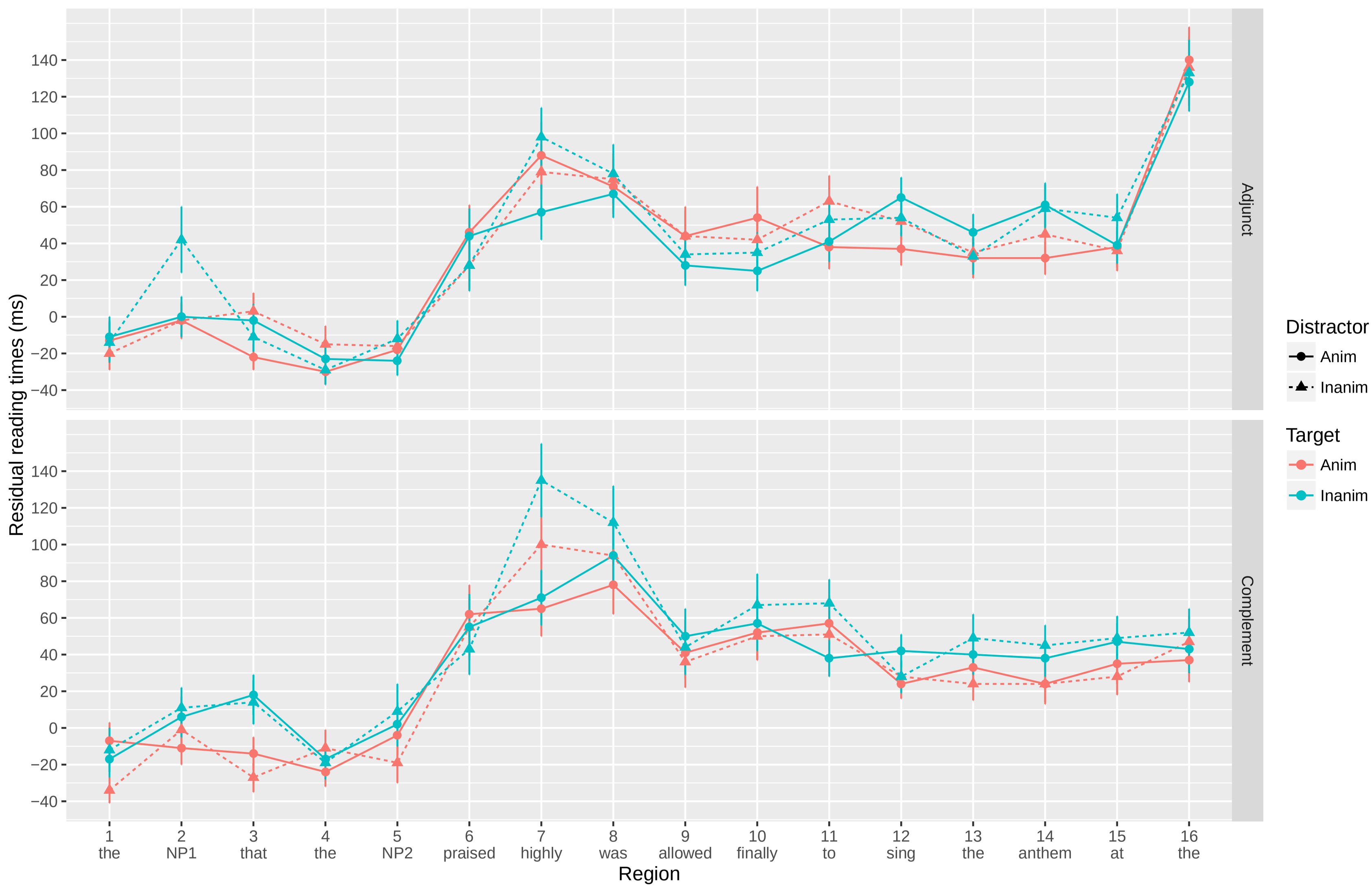


Figure 4.3: Experiment 5 residual reading times and standard error (ms) for adjunct control (top) and complement control (bottom) sentences, aggregated by participants and items. The critical retrieval was region 11, the target was NP1, and the distractor was NP2.

influence downstream of the critical region. A significant effect of target animacy arose at region 14 in both structures, such that inanimate target sentences were read significantly faster than animate target sentences (β = 13 ms, s.e. = 4 ms). Further modeling revealed that complement control sentences with an inanimate target were read significantly more slowly than sentences with an animate target (β = 14 ms, s.e. = 6 ms), as were adjunct control sentences, though the difference was only trending significant (β = 11 ms, s.e. = 6 ms; t = 1.67). The effect of target animacy persisted through region 15 in both structures. Inanimate target sentences were again read significantly more slowly than animate target sentences at region 15 (β = 13 ms, s.e. = 4 ms). At later regions, a slowdown for animate distractor sentences was trending significant at region 17, but only for complement control sentences (β = 15 ms, s.e. = 8 ms).

Intrusion effects.— Visual inspection of the data at region 12 suggested an interaction of structure and target/distractor animacy in the adjunct control conditions. A significant three-way interaction arose at region 13, indicating an intrusion effect. Adjunct control sentences with an animate target and inanimate distractor were significantly slower than adjunct control sentences where both target and distractor were animate (β = 35 ms, s.e. = 17 ms). The interaction reflects the fact that no such effect of distractor animacy was observed in complement control sentences.

Given the importance of this region and the interaction for experimental predictions, data at this region was also analyzed in models fit to each structure independently. Models fit separately to adjunct and complement control sentences revealed that the interaction of target and distractor animacy at the object DP of the infinitive object (region 13) was significant for adjunct control sentences, but not complement control sentences. Sentences in which both target and distractor were inanimate were significantly slower than other adjunct control conditions (β = 26 ms, s.e. = 12 ms). Importantly, no such interaction was found in models fit to complement control sentences at regions 13 or 14 (the direct object of the infinitive in these conditions). At subsequent regions, these models also showed a trending effect of distractor animacy in complement control sentences at region 17, such that sentences with an animate distractor were read more slowly than sentences with an inanimate distractor (β = 15 ms, s.e. = 8 ms).

Comprehension question accuracy

Comprehension question accuracy for each condition is shown in Table 4.4. Mean comprehension accuracy across all conditions was 57%. Accuracy on adjunct control sentences (58%) was comparable to accuracy on complement control sentences (57%), but accuracy for animate target sentences (61%) was better than accuracy for inanimate target sentences (54%).

Target	Distractor	Adjuncts	Complements
Anim	Anim	59% (3%)	58% (3%)
Anim	Inanim	64% (3%)	63% (3%)
Inanim	Anim	51% (3%)	52% (3%)
Inanim	Inanim	57% (3%)	55% (3%)

Table 4.4: Experiment 5 comprehension question accuracy, mean percentage correct and standard error for each condition.

Animate targets led to significantly greater accuracy in both complement control (β = 0.5, S.E. = 0.16) and adjunct control (β = 0.47, S.E. = 0.17). The effect of distractor animacy was trending significant in the adjunct control sentences (β = -0.28, S.E. = 0.16), but non-significant in complement control sentences (β = -0.26, S.E. = 0.16).

4.5.4 Discussion

The general slowdown at the infinitive regions for adjunct control sentences relative to complement control sentences suggests that something about licensing a controlled infinitive is more difficult for (temporal) adjuncts than complements. Aside from the syntactic position of the infinitive, adjunct control sentences differed from complement control sentences in the matrix verb (and so in the predictability of the infinitive), the category of the word following this verb, and the morphology of the infinitive verb. Any differences due to the form of the infinitive verb would be unable to account for why the structural differences arose before the form of the verb could be known. Furthermore, while adjunct infinitives were preceded by a preposition such as *after*, and complement infintives were preceded by an adverb, in both structures this word served to signal a clause boundary. In adjunct control sentences, any predictive benefit of the preposition is likely quite mild, since the preposition is often consistent with non-clausal continuations, e.g. *after lunch*. I thus take the matrix verb to be the source of the differential processing in the infinitive, in that the predictability of the upcoming infinitive in complement control eases the processing burden associated with licensing the implicit subject.

This conclusion is also supported by the effects of target animacy. In both structures, inanimate targets were poor controllers of the infinitive subject. The implicit subject of a controlled temporal adjunct is required to be animate by a general constraint (see Parker, 2014, for evidence from acceptability judgements), while in the complement control sentences used in the current experiment this constraint was enforced by using infinitive verbs that strongly preferred sentient agents, e.g. *sing*. Despite these differences, both structures showed evidence that poor controllers were detected quite rapidly, one word after the null subject was confirmed by the input. This is also the putative retrieval site for identification of the null subject's antecedent. Thus, poor controllers were detected quickly in both structures, immediately after the gap, but controlled adjuncts were nonetheless more difficult than controlled complements. These patterns are consistent with the view that the degree of difficulty in processing a controlled infinitive depends, at least in part, on the predictability of its null subject given (here) a lexical control verb that requires an infinitive complement.

The question at the center of this is experiment was whether this predictability would interact with interference in retrieving the antecedent of the implicit subject. One indicator of interference would be an effect of distractor animacy on inanimate target sentences, an intrusion effect. On the other hand, an effect of control structure on this interference would implicate a role of predictability. Consistent with this prediction, adjunct control sentences with two inanimates were significantly slower than the other three adjunct control conditions. No such effect was present in the complement control sentences. Recall also that both structures showed a significant slowdown for animate targets at the preceding region (12). Thus, another way of framing this interaction is that in adjunct control sentences only, the processing of ungrammatical sentences (inanimate target) with an animate target were on par with grammatical sentences.

That adjunct control sentences are prone to intrusive interference is consistent with the findings of Parker (2014), while the relatively stronger effect of intrusion in that study can be explained by the larger number of mismatching features on the target. Here we extend these findings – or lack thereof – to complement control sentences, and find that the processing in regions at or around a null subject gap is sensitive to (i) similarity between the controller and other, structurally illicit constituents, and (ii) the predictability of the gap given prior sentential material.

The finding of differential processing for adjunct control and complement control is reminiscent of Moises Betancort, Carreiras, & Acuña-Fariña's (2006) study on the processing of controlled PRO in Spanish sentences. Thus, in conjunction with other evidence (reviewed above) the results of the current experiment add supporting evidence to the hypothesis that one way the sentence processing architecture overcomes the challenge of interference is by engaging predictive mechanisms alongside a content-addressable retrieval mechanism.

4.6 Implicit subjects, complexity, & predictability

In this chapter, we extended our examination of subject retrieval to include implicit or phonologically null subjects (PRO). I reviewed some evidence that control dependencies involve reactivating the antecedent of a gap, potentially via content-addressable retrieval. Previous research also suggests that control dependencies are processed actively, in that the parser attempts to resolve open dependencies as quickly as possible, often in advance of any direct evidence in the input. The question then became how these predictive processes interact with retrieval, if at all. The answer to this question, based on the current experiments and previous work, is that predictability of a gap matters, at least insofar as unpredictable dependencies are more interference prone.

A second question was how the complexity of interfering elements modulates difficulty in resolving the antecedent of PRO. Recall that in Experiment 3, the interference engendered by complex interveners in object control sentences was modest, at best. The results of Experiment 4 contrast with those of Experiment 3 in that the effect of intervener complexity was a general facilitation at the retrieval site. We return to the issue of complexity in the next chapter, but for now note that one difference between the materials of these experiments is that, while both experiments separated the control gap from its antecedent by an intervening relative clause, only in Experiment 5 did this clause itself contain a gap. This gap, in turn, was coreferent with the matrix subject, the controller of the implicit subject of the infinitive. This could potentially facilitate the retrieval of the antecedent by repeatedly reactivating the controller before it is retrieved at the infinitive gap. Relatedly, whatever the source of complexity effects is, in Experiment 4 its outcome is an increased ability to discriminate the correct controller from other retrieval outcomes. Broadly speaking, I think that these effects can be unified by the intuition that items in memory are easier to remember when they are familiar (reactivated) and/or distinct (complex).

Simulating subject retrieval in ACT-R

5

5.1 Introduction

Up to this point we have been concerned with how subject encodings are encoded and retrieved in the course of sentence processing, and how these processes are affected by consituent complexity. In Chapter 3, we saw that the retrieval cues for subjects must be able to distinguish between the nominative subject of a finite clause and the accusative subject of a non-finite clause. In chapter 4, we looked for interference effects in the retrieval required to identify the antecedent of an implicit subject. In this chapter, we revisit these questions in a series of computational simulations, using a retrieval model implemented in the ACT-R theory of cognition (Lewis & Vasishth, 2005; J. R. Anderson & Lebiere, 1998).

The goal of these simulations is to resolve some residual issues with the results of the previous chapters. One, we seek to sharpen our theory of retrieval cues to account for the interactions complexity and syntactic similarity observed in the experiments of chapter 3. Two, we seek to resolve a tension between the generally inhibitory effect of complexity observed in Chapter 3, and the facilitatory effect of complexity in Chapter 4.

Recall from Chapter 3 that a complex embedded subject consistently led to difficulty in subject retrieval when it matched the target subject in both case and position (S-complement sentences). In sentences containing an embedded ECM-verb complement, where the intervener matched the target in position, but not case, this effect was attenuated (Experiment 2) or reversed (Experiment 3). These patterns contrasted with the lack of an effect of complexity in object control sentences.

I argued that these effects can be understood as evidence of interference in subject retrieval, modulated by a distinctiveness effect arising from the complexity manipulations. In brief, this additional processing has the effect of increasing the availability of the complex constituents and their features. The impact of complexity on retrieval depends on the level of background similarity between the complex constituent and the retrieval target: items that are similar to the target will become more intrusive, while those that are already dissimilar are unaffected because the complex constituent is not a candidate for retrieval. The interesting case is the effect of complexity on items that are only partially similar to the target, like the ECM sentences of Experiments 2–3. Here, the effect of complexity is to make the mismatching features more distinctive, thereby reducing the difficulty of discriminating between candidate encodings. The retrieval model here incorporates a notion of representational strength, *activation*, and thus allows us to probe both the effect of complexity and its interactions with the putative retrieval structure.

These simulation also allow us to test an issue related to questions of the retrieval structure, namely the *target* of retrieval. Up to this point, I have assumed that encountering a verb initiates a retrieval of its subject from memory, without being particularly sanguine about this assumption. In the sections that follow, I discuss some reasons that we might instead think that the target of retrieval is clause that contains the subject. In brief, a subject DP in the input leads the parser to expect an upcoming clause. When the verb is encountered, this expectation must be reactivated from memory, so that the verb can be bound as its (lexical) head. It may be, then, that the observed interference profiles reflect interference in clausal retrieval. Testing this hypothesis is difficult in English, since the properties of a clause are inextricably linked to properties of their subjects, like case and position. Computational simulations have no such issue, though, since we can specify precisely what the retrieval targets are. Thus, one goal of this chapter is to examine which models give a better fit to the observed data, given different retrieval structures, including different retrieval targets.

Returning now to the second goal, another challenge for interpreting the effect of complexity comes from its effects on the retrieval of overt subjects in Chapter 3 and implicit subjects in Chapter 4. At the heart of this contrast is that identifying the antecedent of a controlled subject was uniformly easier in sentences containing a complex distractor. This effect seemed insensitive to similarity, in that the strength or direction of this effect was unaffected by the syntactic position of the intervener.

Since the retrieval site in all experiments was a verb in the root clause, the retrieval cues seem unlikely to account for the contrast. One potential explanation here is non-linguistic factors, such as recency of the target/distractor to the retrieval site. On the other hand, the materials of Experiment 4 differed from those in Experiments 1–3 in that the embedded relative clause that hosted the intervener always contained a gap that was referentially linked to the retrieval target. Processing the gap, then, might increase the availability of the target, if the ease of remembering is a function of the utility of that memory in the past. Note that the issue of clausal retrieval reappears here, in that properties of the clause that intervenes between a verb and its retrieval target – gapped or ungapped – may modulate the efficiency of memory access at a later point. Thus, the second goal of this chapter is to provide an explanation as to why the complexity of an intervener interacts with similarity in some configurations, but in others seems to engender a robust facilitation effect.

5.1.1 *Outline of the chapter*

We begin by discussing some general constraints on the processing model and factors that determine the efficiency of memory access (§5.2). In particular, I focus on the effects of recency of the to-be-remembered item to the retrieval point, the complexity of the retrieval target and distractor encodings, and the effects of similarity. While this is mostly a summary of topics from previous chapters, it is important to understand what the constraints are and how the model captures them.

Section 5.3 articulates the model of subject retrieval used in the simulations. The model is implemented within the Active Control of Thought-Rational (ACT-R) model of cognition (J. R. Anderson & Lebiere, 1998; Lewis & Vasishth, 2005). One advantage of this model is that the architecture is *computationally complete*, in that it specifies the memories, processes, and control structures required in any computational architecture. I sketch these elements and show how they are realized within the sentence processing model.

The remainder of the chapter addresses using the resulting model to simulate the experiments of Chapters 3 & 4. Section 5.4 discusses how the predictions of the model were derived and how the fits were obtained. In brief, the predictions reduce to questions about how constituents are encoded and when they are retrieved. I discuss the representational assumptions about the encoded constituents. Constructing the retrieval schedule requires some specific assumptions about the parsing algorithm that determines the trajectory of structure-building operations. I sketch the workings of a left-corner parser, which is used to derive both the creation times for constituent representations and when they are retrieved.

We then turn to the simulations themselves. Section 5.5 simulates the retrieval of a subject at a verb, using the materials of the experiments in Chapter 3. Section 5.6 simulates the retrieval of controlled PRO, using the materials of the experiments in Chapter 4. Section 5.7 discusses the results of the simmulations. Section 5.8 concludes.

5.2 The efficiency of memory access

The theory we will adopt takes processing difficulty to reflect working memory constraints (Lewis & Vasishth, 2005). In this section, I review the evidence from previous chapters for three of the primary determinants of the efficiency of memory access.

5.2.1 *Recency of the target to retrieval site*

Perhaps the most important determinant is recency—recent memories are available for immediate processing, while everything else needs to be retrieved. Several studies of retrieval dynamics have found that recall/recognition of the most recent item (with nothing intervening between study and test) is retrieved about 50% faster than non-recent items (Wickelgren, Corbett, & Dosher, 1980). Similar results are reported in McElree & Dosher (1989) using word lists. The most recent word showed significantly faster latency (44% faster SAT rate) and higher accuracy, especially for larger memory-set sizes. Dosher (1981) reports the same pattern in a word-word paired association recognition task. Summarizing these studies, McElree (2006) observes that the probability of retrieval (the SAT asymptote) decreases as a function of recency, so that accuracy is lower for less-recent items. On the other hand, the accessibility of a memory (the SAT rate and intercept) shows only a two-way split: the most recent item is retrieved significantly faster than non-recent items, but the retrieval speed for non-recent items is unaffected by the degree of recency. In other words, the only thing that matters for retrieval speed is whether the item is the most recent or not.

The same pattern of facilitated processing for the most recent item has also been observed in language processing. In cleft constructions like those in (121), retrieval dynamics (SAT) for sentences where the verb and its subject are adjacent (121a) show significantly faster retrieval speeds than sentences where subject and verb are separated by one or two relative clauses (121b–c). The recency advantage extends beyond cleft constructions, too; The same pattern holds in filler-gap questions and subject-verb dependencies (McElree, Foraker, & Dyer, 2003, see also §2.??).

(121) a. This was the book that the editor admired.

b. This was the book that the editor who the receptionist married admired.

c. This was the book that the editor who the receptionist who quit married admired.

The facilitated processing of the most-recent item/memory is taken to reflect the item being maintained in a active, attended state – 'primary memory' – where it is immediately available for processing (Wickelgren, Corbett, & Dosher, 1980). Thus, recency matters because the most-recent item is often in this active state – the focus of attention – while other, non-attended memories are in a passive storage state, and must be retrieved for processing (McElree, 2006). In other words, recency is an important determinant of the efficiency of memory access because recency is a factor in whether retrieval is required. Recent memories are more likely to be in the focus of attention, obviating the need for their retrieval, while non-attended memories must be retrieved from memory, which takes time.

Additional evidence for this view comes from studies indicating that the processing advantage is tied to an actively attended item, and not just the most recent item. For one, the advantage can be extended to more than a single item when the experimental task encourages 'chunking', like categorized lists (McElree, 1998). Two, studies of list-memory indicate that the advantage extends to other, non-recent items when the task encourages participants to maintain non-recent items in the focus of attention via covert rehearsal (McElree, 2006).

Returning to language processing, it is still an open question as to when a consituent is displaced from the focus of attention. We have already seen that in a sentence like *the book ripped*, the subject is still in focal attention at the verb (McElree, Foraker, & Dyer, 2003). Other evidence indicates that the advantage afforded to constituents in focal attention can be extended to non-recent arguments when their verb is itself in focal attention.

Reactivation of the antecedent for a reflexive anaphor is generally robust to interference effects, and highly accurate (Sturt, 2003). However, Joseph King, Andrews, & Wagers (2012) observe that the reflexive in these studies is always verb-adjacent, so that retrieval of the antecedent can potentially benefit from the verb (and its argument structure) being in focal attention. In order to probe for interference in constructions where the reflexive is separated from the verb, Joseph King, Andrews, & Wagers used sentences containing verbs of transfer and VPs with benefactive arguments. The key conditions are shown in (122). I set aside conditions where the target mismatched the features of the reflexive. Their results indicated evidence of interference from the distractor when the reflexive was not verb-adjacent.

(122) a. Adjacent reflexive:

The bricklayer who employed *Helen* shipped herself sacks of mortar.

b. Non-adjacent reflexive:

The bricklayer who employed Helen shipped sacks of mortar to herself.

These results suggest that a verb is in focal attention at its direct object, but has been displaced by the time the indirect object is encountered. This is important for sentence processing because if structure-building operations require retrieval – e.g. binding a subject or direct object to a verb – then parsing will be facilitated in those cases where the otherwise-to-be-retrieved constituent is in focal attention.

5.2.2 Similarity-based interference

The second determinant of difficulty in memory access, somewhat familiar at this point, is similarity-based interference. Intuitively, items are more difficult to remember when they are similar to other items in memory (M. C. Anderson & Neely, 1996). Here I briefly review some of the core properties of interference, since the model of retrieval discussed below is, at bottom, a formal and computational definition of these properties.

Similarity is a property of the item in context. The primary determinant of retrieval success is the ability of the retrieval cues to correctly discriminate between the target and other items in memory, the 'diagnosticity' of the cue(s) (Nairne, 2002). Holding the probe-to-target similarity constant, processing is affected as a function of similarity between the target and other items in memory. The focus on cue-diagnosticity is an outgrowth of the now-classical conception of interference as driven by competition between items sharing a retrieval cue (??McGeoch, 1942). More recently, this formulation lies at the heart of computational models of retrieval. As we will see, the broad view is that retrieval success for some memory (like a syntactic constituent) is a function of the similarity between that memory and the retrieval cues, proportional to the degree of similarity between the cues and other items in memory.

5.2.3 Complexity

Another determinant of retrieval efficiency that we have been concerned with in the present work is how the complexity of (non-)target memories impacts retrieval. The evidence from sentence processing indicates that increasing the complexity of constituents can modulate the accuracy of retrieval processes. Increasing the syntactic/semantic content of a to-be-remembered constituent leads to longer reading times at encoding regions, but faster RTs at retrieval regions (Hofmeister, 2011; Hofmeister & Vasishth, 2014). For example, Hofmeister (2011) found that sentences like those in (123) were read more slowly at the clefted noun, *communist*, as its complexity increased, but additional modifiers facilitated processing at the verb, *banned*.

(123) a. It was *a communist* who the members of the club banned from ever entering the premises.

b. It was *an alleged communist* who the members of the club banned from ever entering the premises.

c. It was *an alleged Venezualan communist* who the members of the club banned from ever entering the premises.

These effects are typically explained by appeals to the complex constituent being somehow more *distinctive*. This facilitates retrieval because increasing the distinctiveness of an item eases difficulty in discriminating between encodings. Thus, for instance, Hofmeister (2011) accounts for the comparatively easier retrieval of complex targets by an appeal to a notion of representational strength and similarity-based interference. The account is based on the idea that increasing the complexity of some constituent increases the likelihood that the constituent will bear a unique feature. The complex constituent, then, will be easier to retrieve to the extent that this unique feature can distinguish it from other competitors in memory and/or increases the match between the complex constituent and the retrieval cues. A nice result of this view is that it blends two strands of memory research: complex consituents require additional processing, *elaboration*, which can increase the diagnostic value of associated features, a distinctiveness effect.

If increasing the complexity of a consituent can strengthen its match to the retrieval cues, then increasing the complexity of a non-target (a distractor) should increase interference by making the complex consituent more intrusive. We have already seen evidence that the complexity of structurally-illicit encodings matters, too, supporting an interference-based explanation. Recall from the experiments presented in Chapter 3 that an embedded finite clause such as (124) engendered greater interference when its subject was complex, either an event nominalization or noun-noun compound. At the same time, the form of this interference was sensitive to the structural similarity between the embedded subject and a higher, target subject. Structural subjects that mismatched a nominative target subject in their case feature, such as the ECM subject of (125), facilitated processing at the retrieval site, while the accusative object of clauses like (126) showed no effect of complexity.

(124) …who believed that the chef delayed the guest — S-complement

(125) …who believed the chef to be delaying the guest — ECM

(126) …who persuaded the chef to delay the guest — Obj. control

I take these patterns to be consistent with an interference approach, in that grammatically inappropriate constituents can exert an influence on processing at the putative retrieval site, but only when similar to the target encoding. At the same time, the inhibitory interference for fully-matching (S-complement) subjects and the facilatory interference for partially-matching (ECM) subjects, when complex, suggests that complexity can ease processing by increasing the diagnosticity of (say) a case cue against a background of otherwise similar encodings. In other words, complexity may increase the distinctiveness of a cue given a background of similarity. When this background similarity is lacking, as in the object control sentences, complexity cannot highlight differences between encodings, and has no effect. Indeed, in summarizing memory research on distinctiveness effects, Hunt (2006) concludes that the proper definition of distictiveness is the 'the processing of dissimilarity in the context of similarity', a point to which we return below.

A potential challenge to this view comes from an experiment conducted by Hofmeister & Vasishth (2014). who examined the effects of elaboration and contextual isolation – alternatively, complexity and distinctiveness – on the processing of English object-relative sentences, such as those shown below in (127). In a 2 × 2, self-paced reading experiment, Hofmeister & Vasishth manipulated the complexity of retrieval candidates by adding modifiers to the matrix object (NP1, the target) and/or the preceding matrix subject (NP2, the distractor). The critical retrieval was the verb of the relative clause (*advised*), where the matrix object is retrieved[1].

(127) a. The congressman$_{\text{NP1}}$ interrogated the (victorious four-star) general$_{\text{NP2}}$ who a lawyer for the White House advised to not comment on the prisoners.

[1]This retrieval is also potentially motivated by the presence of the control verb, *advise*, but it is not clear if all the the materials contain similar verbs. The discussion of the materials suggests that not this was not a systematic confound, and that at least some materials contained simple transitive verbs like *banned*.

b. The (conservative U.S.) congressman$_{NP1}$ interrogated the general$_{NP2}$ who a lawyer for the White House advised to not comment on the prisoners.

If elaborating the target constituent facilitates retrieval, then complex targets should speed processing at the critical retrieval site, as in Hofmeister (2011). The additional manipulation of distractor complexity probed whether elaborative processing of non-targets eases processing by facilitating the discrimination of candidate encodings, and controlled for length effects.

Consistent with previous studies (above), Hofmeister & Vasishth observed slower reading times at the complex constituents, presumably due to additional processing at encoding, and faster RTs at the retrieval region when the target (NP2) was complex. Crucially, though, complex distractors also facilitated processing at the retrieval region, but only when NP2 was simple. The effects of complexity also diverged in that the effect of target complexity persisted to the word following the critical verb, but the complexity of the distractor did not.

These results are seemingly at odds with the results of the experiments here, where complex distractors *did* modulate retrieval difficulty (though the form of this effect varied). Nonetheless, I think that the results of Hofmeister & Vasishth (2014) do not consitute a counter-argument to an interference-based account of complexity, and that overall the studies discussed above are broadly consistent. For one, the lack of an effect for distractor complexity when the target was also complex might be explained as the facilitation for complex targets mitigating the intrusion of complex distractors. In essence, the complexity effects cancel each other out.

More importantly, the complexity manipulation of Hofmeister & Vasishth (2014) involved distractors that were not structurally similar to the target. In their materials, the distractor was a matrix object, while the target was a matrix subject. If the effect of complexity is to make constituents more distinct by highlighting differences between otherwise similar consituents, then the lack of structural similarity between target and distractor will not provide a sufficient level of background similarity. Nonetheless, it's worth reiterating that *any* effect of the distractor is compatible with a content-addressable retrieval mechanism, since a structured search would be much more accurate.

Hunt (2006) stress the importance of this contextual similarity, and give the following example. What are the differences between *cat* and *dog*? What are the differences between *gasoline* and *tree*? In the first case, the dissimilarity is likely based on the dimensions of shared similarity: both are animals, usually pets, but cats meow and dogs bark. In the second pair, however, the differences are so manifold that it seems hard to know where to start. This example is meant to illustrate that processing dissimilarity is easier along some dimension of similarity. The processing of these differences, in turn, is what is claimed to give rise to the distinctiveness effect. I think that complexity can be understood as a form of elaborative processing givng rise to a distinctiveness effect. To see how this works, it is necessary to briefly discuss distinctiveness effects.

Distictiveness in memory

The classic pattern of distictiveness is an isolation effect, where memory performance is improved for items that differ from their surrounding context, eponymously called the *von Restorff effect* (Von Restorff, 1933). Von Restorff's original experiments illustrate the core phenomenon of distictiveness, and also provide evidence against some potential explanations for the effect (like salience), and so I describe them briefly here. (This paper has not been translated to English, and so my description is based on the discussions in Nairne (2006) and Hunt (2006).)

Von Restorff measured the participants' accuracy on the recall of three lists of ten items. The first list contained 10 unrelated items, the second list contained nine numbers and one nonsense syllable, and the third list contained nine nonsense syllables and one number. List 1 served as the control condition, while in lists two and three instatiate the isolation conditions. Von Restorff showed that memory for the isolates was significantly better than items in the control condition, even when the unrelated items occupied the same ordinal position in the control list as the isolate.

Importantly, the isolation effect obtained even when the isolate occupied the first position in the list. This provided the first piece of evidence that the distinctiveness effect is not driven by salience of the isolate, since the initial item could not draw more attention due to its differences from the list context. Von Restorff's experiments also illustrate that distinctiveness is not a property of the item itself, but rather depends on the relative differences between the isolate and background items.

Nonetheless, the differences between the isolate and background items are a necessary, but not sufficient, condition for distictiveness effects. For instance, Epstein, W. D. Phillips, & S. J. Johnson (1975) studied the recall of related word pairs like *dog-cat* and unrelated pairs like *dog-beer*, while asking participants to perform an additional 'orienting task'. The orienting task required listing either the similarities between items or their dissimilarities. Recall of the related pairs was significantly better for those participants who performed the similarity judgement task, while the recall of unrelated pairs was better for those participants who performed the similarity orienting task (see also Begg, 1978). However, there was no benefit for the dissimilarity judgement task on unrelated pairs, which had no contextual similarity. Relatedly, related words showed benefits for the dissimilarity judgement, but not the similarity judgement. Taken together, these results illustrate that neither processing similarity nor processing dissimilarity, on their own, give rise to the isolation effect. Rather, the distinctiveness effect is the result of processing the dissimilarity of otherwise similar items (Hunt, 2006).

5.2.4 *Two models of distinctiveness*

In this section, I discuss two models of retrieval that have been proposed to account for distinctiveness effects in memory and sentence processing. The first of these models is the 'choice rule' model of Nairne (2006); The second approach is the activation-based retrieval model of ACT-R J. R. Anderson & Lebiere (e.g. 1998). In Nairne's (2006) model, distinctive encodings are easier to retrieve because they bear a feature not shared with other encodings. In activation-based approaches, the distinctiveness effect arises in those cases where the distinctive encoding is associated with additional processing, resulting in a more durable encoding and facilitating its retrieval—an elaboration effect.

Looking ahead, I will adopt the activation-based approach, largely for pragmatic reasons. Both models have been proposed to explain the effects of complexity on sentence processing (Hofmeister, 2011; Hofmeister & Vasishth, 2014). Hofmeister & Vasishth (2014) argue in favor of Nairne's model, for reasons I discuss (and reject) below. Nonetheless, the model is sufficiently general that it is not necessarily inconsistent with the activation-based approach, and provides some insight into conceptualizing the effect of complexity on retrieval.

The choice-rule model of retrieval

One model of distinctiveness is the simple choice/ratio model of Nairne (1990) and Nairne (2006). The model conceptualized the problem of retrieval as one of response selection (Nairne, 2006). Memories consist of feature vectors, records of encoding processes such as [C C 2 3 1]. Retrieving an item from memory involves a comparison of the encoding vector to the retrieval cues – 'lingering records of the immediate past' – derived from primary memory or context. Nairne assumes that these records are degraded copies of the original encoding, e.g. [C ? 2 3 1].

The probability of successfully retrieving an item, E_1, is a function of the similarity of a cue, X_1, to E_1, normalized to the similarity of the cue to other encodings ($E_2, E_3, \ldots, E_N$).

$$P_r(E_1|X_1) = \frac{s(X_1, E_1)}{\sum s(X_1, E_i)} \tag{5.1}$$

In Nairne's model, similarity is realized in terms of a distance measure (following Shepard et al., 1987), computed by comparing the features at each vector position. The similarity between an encoding and the retrieval cues is the ratio of the sum of mismatching features to the number of compared features. Items with fewer mismatching features are 'closer' (more similar). Similarity decreases exponentially as the number of matching features decreases.

$$s(X_1, E_1) = e^{-d(X_1, E_1)} \tag{5.2}$$

Nairne is clear that this is not a full model of retrieval; It does not specify how event traces are represented, and does not link the retrieval probabilities to output behavior. As a conceptual framework, though, it has some nice properties. In particular, it makes clear that the probability of successful sampling depends on the degree of match between the retrieval cues and a given memory (the numerator), and is inversely proportional to the degree of cue-overload (the denominator).

Nairne (2006) applies this model to distinctiveness effects. His hypothetical example is shown below in Table 5.1. Three conditions are shown: the control condition shows the similarity and retrieval probabilities for an item embedded in a list of unrelated items, though some small degree of similarity is assumed (the features "C C"). The second list, dubbed 'Isolate', instantiates the standard distinctiveness effect. The encoding to be retrieved remains the same (the first item on each list), and the retrieval structure is always a fully intact version of this encoding. Nairne suggests that in reality the retrieval structure is in fact a degraded or 'blurry' reinstantiation of the original encoding.

To see how the model captures distinctiveness, compare the control list to the isolate list. In both lists, the similarity between the retrieval cues and the target is a perfect match. Nonetheless, the chances of correctly sampling the target in the control list are less than 50%. This is due to the similarity between the target and other memories, or in other words competition between items sharing a retrieval cue.

In the isolate condition, a non-target memory now bears a distinctive feature, X, not borne by other encodings. In terms of the model, this distinctive feature leaves the probe-to-target similarity unaffected, but lowers the denominator in Equation 5.1 and thereby increases the sampling probability of the target (by a modest amount). Abstracting away from the model formalism, the result for the target is a reduction in cue-overload.

	Cue	Traces	Similarity	Sampling Probability
Control	[C C 2 3 1]	[C C 1 2 3]	.55	.26
		[C C 2 3 1]	1.0	.48
		[C C 3 1 2]	.55	.26
Isolate	[C C 2 3 1]	[C C 1 2 3]	.45	.24
		[C X 2 3 1]	1.0	.53
		[C C 3 1 2]	.45	.24
Iso/Sim	[C C 2 3 1]	[B B B B 3]	.37	.21
		[C C 2 3 1]	1.0	.58
		[B B B B 2]	.37	.21

Table 5.1: Nairne (2006) retrieval model of distinctiveness, sampling probabilities & similarity values

Hofmeister & Vasishth (2014) argue that Nairne's (2006) model provides a better account of the effects of consituent complexity on encoding and retrieval processes observed in their experiments (discussed in the previous section). In brief (they do not give language-specific examples), they argue that complexity facilitates retrieval because the retrieval cues are derived, in part, from reinstantiating a trace of the original memory encoding. Increasing constituent complexity increases the number of 'meaning-related' features associated with that encoding, and to the extent that these features are reinstantiated by the retrieval structure, the amount of cue-overload will decrease, and the likelihood of successfully retrieving the item will increase (as in Nairne's example, above). This view essentially casts the complexity effect as a form of facilatory encoding-interference.

In arguing for this view, Hofmeister & Vasishth explicitly reject activation-based approaches, and in particular the model of Lewis & Vasishth (2005). I discuss this model in detail in subsequent sections. For now, activation is a measure of representational strength, which fluctuates as a function of (i) how often the item has been retrieved (its utility), (ii) the resonance between the retrieval cues and the target item, and inversely (iii) the degree of match between the cues and other items. Retrieval cues are derived from a combination of grammatical knowledge and the retrieval context. The probability of successful retrieval is determined directly by activation: the encoding with the highest is retrieved. As Hofmeister & Vasishth (2014) note, this model has a ready explanation for complexity effects as an increase in activation driven by repeated reactivation of the encoding. Thus, for instance, Lewis & Vasishth (2005) suggest that modifying a constituent involves repeatedly reactivating its head, either as a strengthening of expectations (pre-nominal modifiers) or re-accessing the head to attach (post-nominal) modifiers.

The heart of Hofmeister & Vasishth's (2014) argument against activation-based approaches is that in these models, and unlike Nairne's model, the retrieval cues are wholly derived from the current word and its syntactic context. Insofar as the relevant dimensions of similarity that modulate processing difficulty are features "not directly invoked by the local sentence context", then the model of Lewis & Vasishth is at loss for explanation (Hofmeister & Vasishth, 2014, pp.9–10). For example, Gordon, Hendrick, & Marcus Johnson's (2001) findings that the processing asymmetry between object and subject extracted relatives is attenuated in sentences containing a mixture of definite descriptions and pronouns. This effect is inexplicable, they argue, if the retrieval cues are derived from the retrieval context, since verbs do not select for these properties. As another example, Hofmeister & Vasishth point to the German case attraction effects reported in Logačev & Vasishth (2012). (See Chapter 3 for a discussion of these and other, related findings.) More generally, Hofmeister & Vasishth reject activation-based approaches on the grounds that they cannot account for encoding-interference.

As far as I can tell, an explanation of encoding-interference is indeed lacking in activation-based approaches. This is not to say that these models deny the existence of such effects, though, or that such models are incompatible with such effects. Indeed, Peter Gordon and colleagues, in discussing their studies, repeatedly claim that interference exerts its effects on both storage (encoding) and retrieval (Gordon, Hendrick, & Marcus Johnson, 2001). Given this, I am inclined to see encoding-interference effects as grounds for fruitful future research in activation-based approaches, and not a fatal flaw. In what follows, I adopt such a model largely because it has the advantage of being a precise, computationally complete model that has been applied to an admirably diverse array of cognitive tasks with some success (see, e.g. J. R. Anderson & Lebiere, 1998, for comprehensive review). For all its failings, the model is capable of generating precise predictions given some inputs. In the following sections, I discuss how this is accomplished.

5.3 A model of subject retrieval

Computational modeling is a natural outgrowth of the assumption, standard in cognitive science generally, that language and cognition are computation. This view extends to both the functional architecture of the mind, and its physical instantiation in the brain. The view of cognition as computation has been with cognitive science since the 1950's (R. Sternberg & K. Sternberg, 2016). One of the earliest applications of computational architectures to cognitive science, unsurprisingly, comes from George A. Miller, in the form of the Test-Operate-Test-Exit (TOTE) cycle of cognitive processes (George A Miller, Galanter, & Pribram, 1960), but Allen Newell advocated a central role for computational architectures in theories of cognitive architecture (Newell, 1980; Newell, 1990; Newell, Simon, et al., 1972; See also Zenon Walter Pylyshyn, 1984).

The core of this assumption rests on the view that cognition is a form of information processing, and information processing is computation. More precisely, some subset of human behavior is determined by mental representations, so that cognitive processes can be understood in terms of formal operations that manipulate these symbolic structures (Zenon W Pylyshyn, 1980). The minimal architecture required to describe any computational system is composed of three parts: a memory to store information, the primitive processes that alter the contents of memory and are composed into more complex operations, and the control structure that determines how information flows through the system and the trajectory of processes in time (Lewis, 2000).

In the following section, I describe a model of sentence processing in terms of these components, based on the work of Richard Lewis and colleagues (Lewis & Vasishth, 2005; Lewis, Vasishth, & Van Dyke, 2006; Van Dyke & Lewis, 2003). The model is implemented within the Active Control of Though-Rational (ACT-R) theory of cognition, from which many of the qualitative predictions are derived (J. R. Anderson & Lebiere, 1998; J. R. Anderson, Bothell, et al., 2004; J. R. Anderson, 2005).

5.3.1 The computational architecture of ACT-R

The ACT-R theory of cognition is computationally complete, in that it specifies the computational primitives embodied in any computational system. This property has the potential advantage of unifying theories of sentence processing with more domain general cognitive architectures. It is also the property that allows the model to derive precise quantitative predictions from simulations run on a computer. The functional architecture that supports computation is decomposed into three parts: memories, processes, and the control structure. Here I briefly describe these components in general and in ACT-R, and show how they inform the sentence processing model. Unless otherwise noted, the discussion below is based on Lewis's (2000) discussion of computational architectures, J. R. Anderson & Lebiere's (1998) text on ACT-R, and the ACT-R model of sentence processing in Lewis & Vasishth (2005). I encourage the reader to consult these sources for more in-depth discussion.

Memories in ACT-R

Memories store and represent information used in the course of computation. Characterizing memory systems requires us to be explicit about three questions. One, the kinds of memory states that encodings can occupy, and how elements are represented in these memory states. Two, the acquisition and retrieval processes that determine how memories are created and reaccessed. Three, the limits on memory systems' ability to perform their required duties, the capacity of memory.

Memories in ACT-R can occupy one of three different states. ACT-R makes a distinction between *declarative* and *procedural* memory, an assumption borrowed from the HAM theory of cognition (J. R. Anderson & Bower, 1973; see also J. R. Anderson, 1983). The interface to these memory states is a limited set of *buffers*, each of which has the capacity of a single encoding.

Declarative memory is the knowledge representation of 'things we know', often consciously accessable (J. R. Anderson & Lebiere, 1998). Memories are entered into declarative memory in two ways: as encodings of perceptual objects in the environment, or as memorized solutions to previous problems in the control structure ('goals', see below). For language processing, this

	Memory
1.	Potential memory states and their coding schemes.
2.	Acquistion and retrieval processes
3.	Capacity limitations
	Processes
1.	How memory contents are altered
2.	Chronometric properties: process duration and the effect of other variables
	Control structures
1.	How processes are initiated
2.	How processes communicate information to other processes
3.	Which parts are fixed, and which depend on variable content
4.	Identify multiple streams of control
5.	How control of comprehension processes is related to control of central cognition

Table 5.2: Functional invariants and their properties required by any physical computational system, from Lewis (2000).

means that constituent encodings are created as the result of perceptual input, or as constituents created in the course of structure-building operations (XPs). In other words, memories of linguistic objects are created either bottom-up, from the input, or top-down, from expectations generated in the course of parsing. In this way, ACT-R captures the classical distinction in philosophy about the way that knowledge is created: from the senses, or through the mind (J. R. Anderson & Lebiere, 1998).

Declarative memories are coded as *chunks*, bundles of FEATURE:VALUE pairs. The representational notion of a chunk has its origins in the work of George A. Miller (1957), who showed that people could remember roughly 7±2 random symbols. However, the size of a chunk is inherently functional: Miller found that recall of digits is roughly on par with recall of words, even though words contain more information (bits) than digits. Van Dyke & Lewis (2003) define chunks functionally, as the 'minimal representational unit that can enter into novel relations with other chunks'. The focus on *novel* relations is important, since it means that inherently relational information, like (say) c-command relations, cannot be encoded on a single chunk since it specifies a relation between two encodings.

Nonetheless, it is hard to imagine encoding linguistic representations without some relational information. I follow Lewis & Vasishth (2005) in assuming that chunks correspond to maximal projections (XPs) that encode the core X-bar relations of SPECIFIER, HEAD, and COMPLEMENT. I further assume that linguistic chunks bear features corresponding to those grammatical dimensions implicated in interference effects (see Chapter 3). Some example chunks are shown in the right-hand side of Figure 5.1. The figure also implies an important ramification of the assumption that chunks do not encode relational information: there is no global syntactic representation. Rather, the broad syntactic structure (left side of Figure 5.1) is an emergent property of chunks occurring as the values of argument-structure features on other chunks.

In addition to the declarative memory of chunks, ACT-R assumes a procedural memory that, roughly, represents the way that knowledge is brought to bear in solving problems (J. R. Anderson & Lebiere, 1998). The contents of procedural memory take the form of production rules, which are sets of CONDITION-ACTION pairs that encode compiled knowledge. We return to

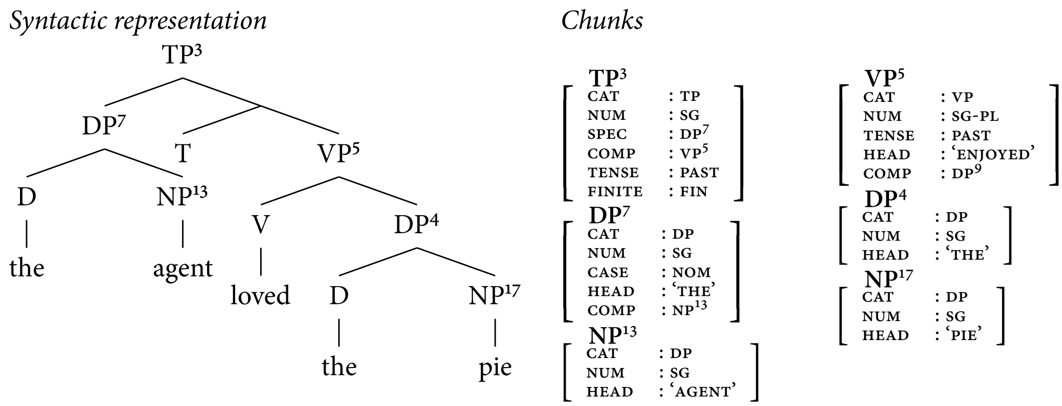


Figure 5.1: Example syntactic structure (left) and the declarative memory chunks of its constituents (right), based on Lewis & Vasishth (2005).

production rules presently, but the important aspect now is that these production rules comprise grammatical knowledge and parsing skill.

Processes in ACT-R

Processing primitives are composed into larger computational operations. These processes are characterized in the way that they alter the contents of memory, and their chronometric properties, both their duration and how it is affected by other variables.

Processes in ACT-R take the form of *production rules*, sets of CONDITION-ACTION pairs stored in procedural memory. Both condition and action of a production rule correspond to memory chunks. The condition side of a production rule corresponds to a test on the contents of memory buffers that determines whether the production is applicable. A matching condition triggers the action side of the production rule, e.g. a declarative memory retrieval or syntactic attachment operation. The primary action of production rules is to alter the contents of control buffers. Production rules in ACT-R determine the grain size of cognition: each production is a step in cognition, as well as the basic building-block for the acquisition of procedural knowledge, and thus by hypothesis corresponds to a mental state (J. R. Anderson & Lebiere, 1998; Zenon W Pylyshyn, 1980).

Control structures in ACT-R

Control structures determine how information flows through the system and the trajectory of computational operations in time. Control in ACT-R is *open* and *dynamic*. It is open because all knowledge is brought to bear on each stage of processing (Lewis, 1993). Control is dynamic because the processing at each step is determined by the set of available productions and the contents of the memory buffers, in particular the goal buffer, which encodes the task at hand. The goal buffer organizes the behavior of the system by ensuring that declarative and procedural knowledge is brought to bear in the service of this goal.

Productions are initiated when the contents of memory buffers match the condition specified by a given production. The contents of these buffers are variable, but we assume specialized linguistic representations. The action side of productions typically involves altering the goal structure – altering the current goal, creating a new one, or popping a completed goal – and/or initiating a retrieval, so that processes can communicate with each other by updating the contents of the goal or retrieval buffers. Changing the goal buffer also prevents a given production from firing in an iterative loop, since once the goal buffer has been altered, the conditions for the production rule will no longer be met when the next production is selected.

While the goal structure allows for multiple (sub)goals, so that multiple goals can be pursued simultaneously, the control structure also serves as a bottleneck to processing, in that only one production can fire at a time. Thus, at the computational level, cognition in ACT-R is serial, characterized by a series of changes to the control structure.

5.3.2 *Retrieval in ACT-R*

Retrieval in ACT-R depends on the *activation* of the chunks in memory—the chunk with the highest activation is retrieved. Activation is a formalization of representational strength that reflects the accessability of the encoding in memory. Activation in ACT-R is defined below in Equation 5.3. The activation of a chunk, A_i, is a function of its usage history (B_i) and the degree of associative match between the retrieval cues and the chunk ($W_j S_{ji}$).

$$A_i = B_i + \sum_j W_j S_{ji} \tag{5.3}$$

The *base-level activation*, B_i, reflects both the time since the chunk was created as well as its utility, as determined by the number of times it was retrieved in the past. The 'associative component', $W_j S_{ji}$, is determined by the utility of the chunk in the current context, i.e. the associative match of the chunk to the retrieval structure. In ACT-R, the activation equation captures the effects of recency, complexity, and associative cue-match, and indirectly interference. I discuss the two components of Equation 5.3 in turn.

The base-level activation of chunk *i*, or B_i, is defined in Equation 5.4. The equation is based on the 'rational analysis' of J. R. Anderson & Schooler (1991), and broadly speaking tracks the log-odds that an encoding will need to be retrieved given its prior usage history. In particular, the usage history is a function of a chunk's retrieval history, summing over *n* retrievals. It is the time since the *j*th retrieval, t_j, modulated by the decay parameter, d.

$$B_i = \ln\left(\sum_{j=1}^{n} t_j^{-d}\right) \tag{5.4}$$

The base-level activation of a chunk also represents one form of learning in ACT-R, allowing the system to statistically adapt to its environment. Chunks that are retrieved often also have a relatively higher activation, reflecting a more durable representation for those (declarative) memories that are frequently retrieved. For language processing, it is the base-level activation that gives rise to frequency effects (Lewis & Vasishth, 2005).

While useful chunks are easier to retrieve, successful retrieval also depends on the degree of associative match between the target memory and the retrieval structure. In the activation formula given in Equation 5.3, this is the associative component $W_j S_{ji}$, where S_{ji} is the strength of association between the chunk, i, and the retrieval cues, j. The parameter W_j is the attentional weighting of cues in the retrieval structure. Following Lewis & Vasishth (2005, a.o.), I will assume that the weight associated with each cue is set by G/j, where j is the number of cues and G is the total amount of activation, assumed to be 1. Chunks receive activation for each matching cue in the retrieval structure, and assuming equal weighting (as I do), the total activation boost is simply the sum of these activations.

Retrieval interference arises when multiple chunks match a given cue. This reduces the diagnostic value of the cue, an effect sometimes called 'cue-overload' (Nairne, 2002). In ACT-R, a cue, j, that matches multiple encodings will have its associative strength, S, reduced as a function of the number of items matching that cue (the 'fan' of j):

$$S_{ji} = S - \log_e(\text{fan}_j) \tag{5.5}$$

Thus, the effect of interference on retrieval in ACT-R is to reduce the activation boost afforded to a chunk via the associative component.

In summary, the retrieval of a chunk in ACT-R is determined by the relative activation of the chunks in memory—the chunk with the highest activation is retrieved. More recent chunks, or those that have been retrieved frequently in the past, are easier to retrieve by virtue of a higher base-level activation.

$$T_i = Fe^{-A_i} \tag{5.6}$$

While the activation of a chunk determines its retrieval probability, activation also determines the latency of retrieval, so that items with higher activation are retrieved faster. In ACT-R, the latency of retrieval, T, for chunk i is related to activation (scaled by F) via Equation 5.6.

5.4 Simulating subject retrieval in ACT-R

In the experiments that follow, I describe the results of retrieval simulations based on the materials of the experiments in previous chapters. Each experiment examines the effect of different retrieval structures on two retrieval targets: (i) the head of the subject constituents, and (ii) the clausal node.

5.4.1 *How predictions are derived in the model*

The inputs to the simulations are the representations of the constituent chunks, including their creation times, and a schedule of the retrievals. In the following simulations, I derive these inputs from the experiments of previous chapters. In an effort to reduce the degrees of freedom in the model, I make the following assumptions about the inputs.

The contents of chunks, the FEATURE:VALUE pairings, are based on assumptions from syntactic theory and empirical evidence about the dimensions of similarity that engender interference discussed in Chapter 3. I assume that the features of chunks are those shown below (see §3.9). Features for nominals are based on evidence reviewed in Chapter 3. Features for CATEGORY, NUMBER, POSITION, CASE. I leave out the animacy feature because all targets were animate, while the animacy of other consituents was balanced—not experimentally manipulated. For clausal constituents: separate categories for gaps and non-gap constituents; features for CATEGORY, NUMBER and TENSE (finite or non-finite); I also include features for X-Bar relations, following Lewis & Vasishth (2005). In particular, a HEAD feature takes the value 'open', and distinguishes those constituents that have been sucessfully integrated/parsed from predicted categories.

The creation times for the input chunks were derived from the reading times of the relevant experiments. For Experiment 1, the creation times for DPs were estimated from the cumulative self-paced reading times at the head noun. For the materials of Experiments 2 and 3, I estimated the creation times from the cumulative regression path times of Experiment 3, at the region of interest (B. W. Dillon, 2011).

The retrieval points for each simulation were estimated the same way as the creation moments, using the cumulative reading times for the embedded and matrix verb regions. Each simulation includes a table of the cumulative raw reading times used to derive the inputs.

I will assume that the retrieval probabilities of the model provide an estimate of the predicted availability of the item in memory. Relative differences in retrieval probability, then, can be interpreted as predictions of interference.

Assumptions about clausal retrieval

I assume that TP nodes are built predictively, after a subject constituent is built. This assumption is motivated for two reasons. There is a substantial body of evidence that real-time language comprehension is predictive, with information derived from preceding material generating expectations about upcoming input. In the visual-world paradigm, comprehenders anticipate upcoming words based on the semantics of a verb, its agent, the case marking of preverbal arguments and/or their syntactic/semantic constraints (Altmann & Kamide, 1999; Kamide, Altmann, & Haywood, 2003; Kamide, Scheepers, & Altmann, 2003). ERP evidence indicates that semantically atypical arguments of a verb engender a negativity (N400) compared to expected, typical nominals (Federmeier & Kutas, 1999; Federmeier, 2007), and that comprehenders show rapid sensitivity to syntactic constraints imposed by the preceding context, like the availability of ellipsis (E. Lau et al., 2006; for a review of the ERP evidence, see Federmeier, 2007). These predictive processes are highly sensitive to grammatical knowledge, as evidenced from the accuracy of predictive gap-filling processes (Stowe, 1986; C. Phillips, 2006; Wagers & C. Phillips, 2009). Clausal nodes are also highly predictable, in general, especially for root clauses. For argumental/modifier clauses, the existence of an upcoming clause is predictable from the existence of a relative pronoun, or the selectional restrictions of an embedding verb. It thus seems reasonable to suppose that clausal constituents are hypothesized before encountering a verb in the input, at least.

The second motivation comes from Lewis & Vasishth's (2005) assumption of a left-corner parsing algorithm (Aho & Ullman, 1972). In a left-corner parser, input of the first constituent of a phrase-structure rule (the 'left-corner') leads the parser to hypothesize the existence of a higher projection on the left-hand side of the PS rule (Steven P Abney & Mark Johnson, 1991). Incoming constituents are attached when material in the input matches the hypothesized structure (Resnik, 1992).

This is important to present concerns, because a (hypothesized) subject leads to an expectation of an upcoming clause (TP). To illustrate, assuming the phrase-structure rules in (128), a determiner in the input leads the parser to hypothesize a DP. This DP, in turn, satisfies the left-corner of the TP rule, so the TP node is hypothesized as well. In this simple example, no further rule applications are possible, so the parser returns to processing the input.

(128) *Sample Phrase-structure rules for subjects*
  a. TP → DP T'
  b. T' → VP
  c. DP → D NP
  d. NP → N (PP)

The situation in the ACT-R model is similar. Thus, Lewis & Vasishth (2005) give the two production rules in (129–130).

(129) *Predict clausal TP node production*
  IF goal category is TP…
    and lexical entry is category DET.
  THEN set retrieval cues to TP expectation.

(130) *Attach subject to predicted TP production*
  IF Lexical entry is category DET…
    and retrieved constituent is a predicted TP.
  THEN set goal category to be NP…
    and create new DP with DET as head…
    and attach new DP as subject of predicted TP.

The first production builds a subject consituent given a determiner, and predicts an upcoming clause by setting the expectations in the goal buffer to a TP node. This production corresponds to a left-corner implementation of the phrase-structure rule in (128c) In the model of Lewis & Vasishth (2005), clausal TP nodes are primitive constituents, and always require a retrieval. The action side of the production thus sets the retrieval cues to retrieve this predicted TP from memory. The second production is also triggered by a determiner in the input, but the condition is further restricted by the contents of the retrieval buffer. Given a determiner in the input and a TP expectation, the second production will attach the hypothesized DP as the subject of the expected clause.

The ACT-R minimum duration for a process is 50 ms. To allow for both processing the input and triggering the first production, I assume that TP nodes are created 150 ms after the subject is created, and retrieved after an additional 150 ms (300 ms after the subject is created).

5.5 Thematic and structural subjects

5.5.1 *Model inputs*

I make the following assumptions about the parsing of the sentences used in Experiment 1. I discuss the assumptions about the retrieval of the subject head first, then discuss assumptions about the retrieval of the clause.

DP retrievals

The crucial events of the simulation are summarized below in 131. Subscripts correspond to points at which constituents are created or retrieved. I show only conditions with a complex intervener and possessor. No possessor conditions dropped the possessor, but were otherwise identical, while simple intervener conditions swapped the positions of the nominalization and the embedded object, *the server*.

(131) The hostess$_a$ who$_b$ thought that the$_c$ chef's careful preparation of the blowfish$_d$ delayed$_e$ the server was$_f$ yelling.

The target subject is created when the determiner is encountered, and completed when the head noun is read (a). The target subject is retrieved almost immediately, at the relative pronoun, in order to attach the relative clause (b). The representation of the intervening embedded subject is created at the determiner that initiates the embedded clause (c). In complex intervener conditions, the representation of the intervener is retrieved at the end of the nominalization to attach the complement (e.g. *of the blowfish*). In simple intervener conditions, this additional retrieval occurs downstream, at the embedded object region (d). The intervening subject is retrieved (again) at the embedded verb, and integrated with the verb (e). The critical retrieval occurs at the matrix auxiliary verb, where the target subject is integrated (f).

Retrieval of the clause

(132) The hostess$_a$ who$_b$ thought$_c$ that the$_d$ chef's careful preparation of the blowfish delayed$_e$ the server was$_f$ yelling.

The matrix clause (TP) node is created after the matrix subject is built (a), and retrieved. I assume that this occurs 300 ms after the head-noun of the matrix subject is read: 150 ms to build the subject and TP node, and another 150 ms to set retrieval cues for a TP expectation and retrieve it. The relative clause TP node is created after reading the relative pronoun, which signals the presence of an upcoming embedded clause. Binding the subject gap requires retrieval of this embedded gapped TP node (b). Following Lewis & Vasishth (2005), I assume a distinct syntactic category for constituents containing a gap. I assume further that the relative clause TP remains in focal attention when the RC verb is read, so the verb can be integrated without an additional retrieval (c).

The embedded TP is created after the distractor subject is built (d). As in the matrix clause, I assume that setting retrieval cues and retrieving the predicted TP node takes time (100 ms). For simple subject conditions, the predicted TP node remains in the retrieval buffer when the embedded verb is encountered, and so does not require retrieval. For complex subject conditions, however, processing the additional structure of the nominalization displaces the embedded TP from the retrieval buffer, and so integrating the embedded verb requires an additional retrieval (e). Finally, at the matrix verb, the predicted matrix TP node must be retrieved (f). This is the critical retrieval, potentially prone to interference from the intervening clauses.

Capturing the subject complexity effect

The two simulations discussed above differ in their explanations of why complex embedded subjects should lead to difficulty. If the constituent retrieved at the verb is the subject head-noun, then additional interference for complex interveners is due to the extra retrievals incurred in processing the nominalization, in particular the retrieval required for modifying the head noun of the nominalization. In contrast, if the matrix verb triggers retrieval of a TP node predicted at the matrix subject, then interference comes from the embedded clausal node. The amount of interference from this constituent is modulated by the additional retrieval required in the complex subject conditions, where the additional material displaces the predicted TP from the focus of attention and the TP must be retrieved again at the embedded verb.

5.5.2 *Results & discussion*

Nominal retrieval

The simulated activation profiles of the critical constituents are shown in Figure 5.2. The DP retrieval model correctly predicts a contrast between complex and simple intervener conditions. The direction of this effect, though, goes in the wrong direction—complex intervener conditions are predicted to be easier than simple intervener conditions.

In complex intervener conditions, the target subject is successfully retrieved at the critical main verb. This retrieval suffers only minor interference from the intervening subject. Interference from the intervener is somewhat more substantial at the embedded verb, which might be interpreted as a prediction of greater processing difficulty or lower comprehension accuracy inside the relative clause.

The results for the simple intervener conditions show substantially less interference from the embedded subject. This is because the activation of the intervener has decayed below the activation threshold required for retrieval. The target itself has also decayed substantially by the critical retrieval, so much so that despite its exact match to the retrieval cues, the winning candidate is in fact the prepositional object of the nominalization (not shown). This DP mismatches the retrieval cues in both case and position; It wins out solely because it is the most recent DP, and thus has the highest activation.

Clausal retrieval

In comparison to direct retrieval of the subject head, simulations of clausal retrieval provide a better fit to the experimental data. The simulation correctly models the effect of intervener

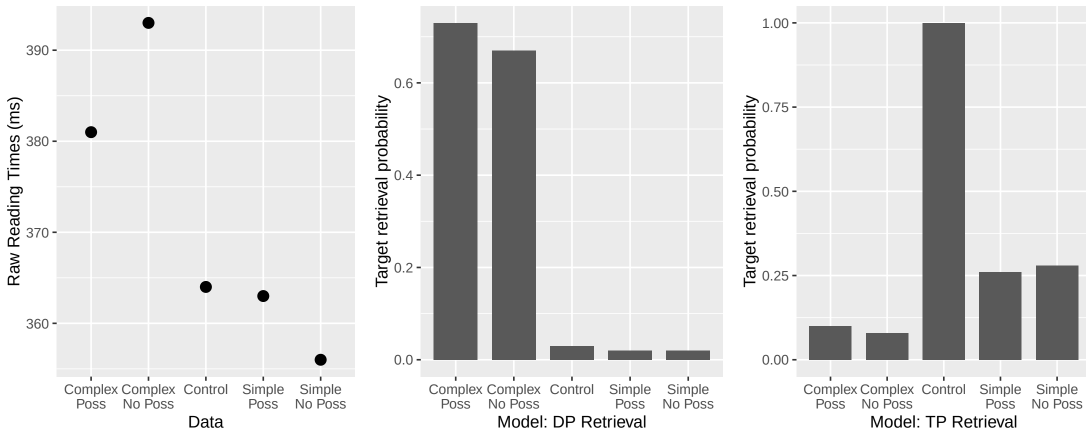


Figure 5.2: Experimental results (left) compared to model results for DP retrieval targets (middle) and TP retrieval targets (right). Retrieval probabilities reflect the predicted probability of sampling the target, matrix subject.

complexity. Complex subject conditions are more difficult, due to interference from the embedded clause. As shown in Figure 5.3, the embedded TP is retrieved more often than the matrix clause in complex intervener conditions. There is minimal interference from the relative clause constituent, due to the category difference. In simple intervener conditions, the retrieval of the matrix TP fares better, though there is some mild interference in the possessor condition.

The source of the differences between simple and complex intervener conditions is the time between the creation of the TP node and when it is retrieved to be integrated with the verb. In complex intervener conditions, processing the nominalization displaces the TP node from the focus of attention, so that its activation decays until the TP is retrieved. This retrieval, in turn, provides an activation boost to the embedded TP, with the result that it is more intrusive in the retrieval of the root TP at the matrix verb (the critical retrieval point). On the other hand, in simple intervener conditions the complex nominalization in object position serves to make the embedded TP less intrusive. In these conditions, the embedded TP is retrieved immediately after it is created, then integrated with the verb, but the delay between the embedded TP retrieval and the retrieval at the matrix clause is increased due to the presence of the complex object.

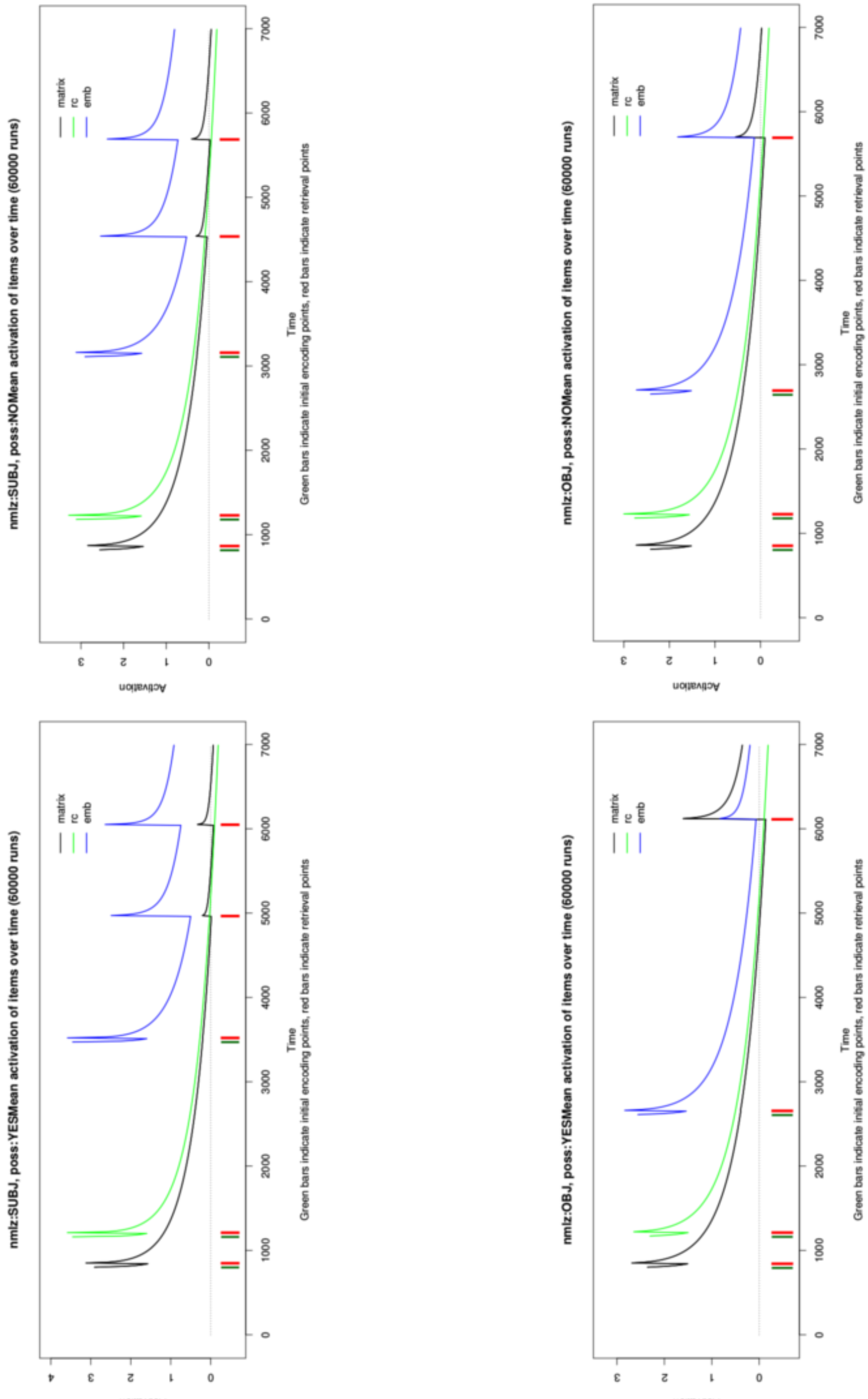


Figure 5.3: Simulation results, complex interveners with possessor.

S-complement

	Target	Intervener	Emb.Obj
cat	N	N	N
num	sg	sg	sg
pos	spec-T	spec-T	comp-V
case	nom	nom	acc

ECM

	Target	Intervener	Emb.Obj
cat	N	N	N
num	sg	sg	sg
pos	spec-T	spec-T	comp-V
case	nom	acc	acc

Object Control

	Target	Intervener	Emb.Obj
cat	N	N	N
num	sg	sg	sg
pos	spec-T	comp-V	comp-V
case	nom	acc	acc

Table 5.3: Constituent representations for S-complement, ECM, and object control sentences.

5.6 Case & position

5.6.1 Model inputs

Constituent representations for the simulation inputs are shown in Table 5.3. The key features are those for CASE and POSITION, with the intervener of S-complement sentences fully matching the target, ECM sentences matching in position, but not case, and object control sentences mismatching the target in both case and position.

The significant retrievals of the simulation, summarized in (133), were as follows. At the relative pronoun, the matrix subject is retrieved and the relative clause is attached [1]. The second retrieval occurs at the embedded verb, where the intervener is retrieved and attached to the verb. The critical retrieval occurred at the matrix auxiliary[3], where the matrix subject was retrieved and integrated. In complex conditions, an additional retrieval of the intervener was used

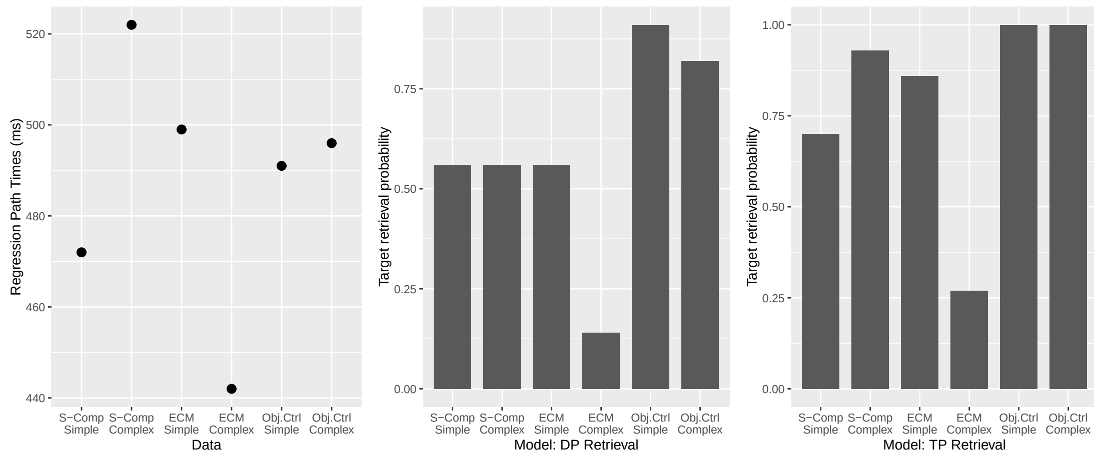


Figure 5.4: Simulation results, showing regression path times from Experiment 3 (left) alongside the model results for DP retrieval targets (middle) and TP retrieval targets (right). Retrieval probabilities reflect the predicted retrieval probabilities for the target, matrix subject.

to model the reactivation involved in elaborating the complex constituent [C].

(133) The explorer who_1 believed the (ancient alien) monster_C to_2 be prowling the ruins was_3 insane after the journey.

For the simulation of clausal retrieval, there were four significant retrievals. First, the matrix TP is predicted at the subject, and following Lewis & Vasishth (2005) is retrieved from memory. The relative clause TP is predicted at the relative pronoun, and again as a primitive consituent is retrieved from memory. I assume that the relative clause TP remains in the focus of attention when the RC verb is read, and can be integrated without retrieval. The third retrieval occurs at the embedded verb, where the TP predicted by the intervener is retrieved from memory. Finally, the critical retrieval occurred at the matrix auxiliary, where the initial expectation of an upcoming TP for the root clause is retrieved.

5.6.2 *Results & discussion*

Retrieval of the subject head

The modeling results for both DP and TP retrievals are shown in Figure 5.4. In simple intervener conditions, the model predicts that object control sentences are easiest at the critical retrieval (the matrix auxiliary), with a retrieval probability of .91. ECM sentences are predicted to be substantially less accurate, with a retrieval probability of .56 for the matrix subject. S-complement sentences are predicted to be only slightly less accurate than object control sentences, with a retrieval probability of .56.

These relative differences were repeated in complex intervener conditions. Object control sentences are again predicted to be the most accurate (.82 retrieval probability), and ECM sentences are predicted as the least accurate, with a target retrieval probability of .14. The predicted accuracy of target retrieval in S-complement sentences was .56.

The simulation successfully models the impact of higher activation for complex interveners, in that an additional retrieval for the complex consituent leads to reduced accuracy at the critical retrieval. Furthermore, the effect of complexity on the object control sentences is comparatively mild, a .9 decrement in accuracy, compared to the more substantial effects on S-complement (-.31) and ECM (-.42) sentences. On the other hand, the model over-predicts differences between sentences in simple intervener sentences, especially for the ECM sentences. Within the complex intervener conditions, the relative differences between sentence types are consistent with the observed reading time data, but only if we interpret the poor accuracy in ECM sentences as reflecting misanalysis (another DP is retrieved). However, this would also lead us to expect the comprehension question accuracy data to show a similar interaction of structure and complexity, contrary to the observed patterns.

Retrieval of the clause

In simple intervener conditions, the model predicts that object control sentences are easiest, with a retrieval probability of 1.0. In other words, the model predicts no interference in these conditions. S-complement sentences are predicted to be hardest, with a target retrieval probability of .70. ECM sentences fall between these conditions, with a predicted retrieval probability of .86.

Complex intervener conditions show the intervener complexity effect, which varied as a function of intervener to target similarity. Object control sentences were again predicted to be the easiest, with a target retrieval probability of 1.0. The model thus predicts no effect of complexity in object control sentences, as in Experiment 3. Complex interveners substantially reduce the accuracy in ECM sentences, which showed a target retrieval probability of .27. The form of the complexity effect was reversed for S-complement sentences, where the matrix TP had a retrieval probability of .93.

As in the simulation of Experiment 1, the interference patterns for clausal retrieval more closely approximate the observed reading times data, though not without issues. The model successfully predicts an absence of complexity in the object control conditions, due to the featural distinctions between the matrix and embedded clauses. It also successfully predicts a complexity effect in the S-complement and ECM sentences, which also varies in form, as desired, but in the wrong direction. In S-complement sentences, complex interveners reduce interference in clausal retrieval, while in ECM sentences the effect of complexity is to reduce the availability of the matrix clause, lowering accuracy. This, of course, is the opposite of the observed pattern. Thus, while the model fails to approximate the experimental data, it nonetheless indicates that activation-based approaches are capable of predicting an effect of non-target complexity, while at the same time allowing the form of this effect to vary as a function of other factors, such as similarity to the retrieval cues.

5.7 Complexity effects in control dependencies

5.7.1 Model inputs

Constituent chunks for simulated sentences are shown in Table 5.4. The contents of memory for these sentences are four chunks corresponding to the matrix subject, the intervener argument of the embedded relative clause, the propositional object of the embedded modifier PP, and the

	Target	NP2 (src/orc)	NP3	NP4
category	N	N	N	N
number	sg	sg	pl	sg
position	spec-T	comp-V/spec-T	comp-P	adjunct
case	nom	acc/nom	acc	acc
animacy	anim	anim	inanim	anim

Table 5.4: Simulation 3, assumed constituent representations and their creation times for each of four conditions.

demoted subject of the passivized object control verb in the matrix clause. The target of the critical retrieval inside the infinitive was the matrix subject, NP1, while the intervener, NP2, was interfering (or distractor) constituent. I assume that chunks bear features for POSITION, CASE, NUMBER, and ANIMACY. Creation times for these chunks were estimated from the results of Experiment 4 (§4.5). Consistent with a left-corner parsing algorithm, I assume that nominal chunks are created when their determiner is read, but assume an additional 150 ms to allow for processing the determiner, triggering the NP production, and setting retrieval cues for the next constituent.

Each condition contained four significant DP retrievals, summarized in (134). I assume that these sentences involved other retrievals, but since these were always retrievals of categories other than nominals (e.g. a VP to attach a modifier) I did not include them in the simulation.

(134) The tourist that$_1$ ___ $_2$ helped the (quiet forest) guide$_C$ with travel plans was$_3$ allowed by the stern judge to$_4$ leave.

(135) The tourist that the (quiet forest) guide$_C$ helped ___ with travel plans was allowed by the stern judge to leave.

At the relative pronoun, the matrix subject is retrieved and the RC is attached [1]. The matrix subject is retrieved again inside the relative clause, and co-indexed with the gap [2]. At the matrix auxiliary, the subject is retrieved a third time [3]. The critical retrieval occurred inside the infinitive clause, where *to* triggers retrieval of the implicit subject's controller, the matrix subject. As in previous simulations, the effect of complexity was modeled as an elaboration effect, accomplished via an additional retrieval of the complex constituent [C].

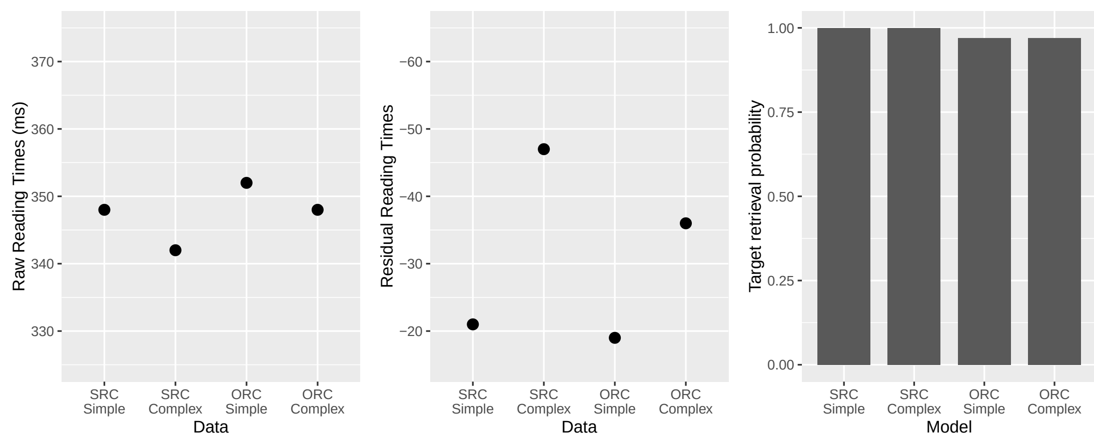


Figure 5.5: Simulation 3 results, showing empirical results from Experiment 4 in both raw RTs (left) and residual RTs (middle) compared to predicted retrieval probabilities for the matrix subject (right).

5.7.2 Results

The predicted probabilities of retrieving the target subject are shown in Figure 5.5. Overall, the simulations correctly predict high accuracy in the complex constituent conditions, but nonetheless fail to model the effect of complexity. The small effect of complexity on the retrieval probabilities for the target antecedent more closely approximate the non-significant differences in the empirical raw reading times, where differences between conditions were very small (<10 ms). Raw reading times do not account for differences in individual reading rates, however, and for this reason the results of Experiment 4 were reported in terms of residual RTs (see §4.4). The model performs less well in approximating the residual reading times results.

Both SRC and ORC structures fail to show an effect of complexity, and there were only minimal differences between structures. SRC sentences showed perfect accuracy (1.0) in both simple and complex conditions. Similarly, ORC structures showed high accuracy in both simple and complex intervener conditions, with a retrieval probability of .97 in both conditions.

The lack of a complexity effect in this simulation is due to the fact that the retrieval of the complex consituent erroneously retrieves the matrix subject, not the complex intervener, even when structural cues heavily favor the intervener in SRC conditions. The matrix subject outcompetes the intervener at this retrieval due to the activation boost it receives when it is retrieved during attachment of the relative clause subject.

5.8 Discussion

Overall, the results of these simulations are more consistent with the view that subject interference is driven by difficulty in retrieving a predicted clausal node when its verbal head is encountered than the view that the subject head is retrieved directly. Such a view is also more consistent with the predictions of a left-corner parsing algorithm. This result is most clear in the simulations of Experiment 1, where the results indicated that it was possible to retrieve the target, matrix subject, but the model incorrectly predicted that this retrieval would be more difficult when the complex consituent was located in the embedded object position, immediately before the retrieval site. This suggests that the model has a tendency to over-estimate the effect of recency on retrieval, in that recency is more important than similarity. In contrast, the simulated clausal retrieval of Experiment 1 correctly predicted that retrieval is more difficult when the complex intervener was a subject compared to when it was an object. Thus, the results of the first simulation correctly predict that complexity modulates the accuracy of retrieval processes, and that this effect is sensitive to the degree of similarity between the complex constituent and the retrieval cues.

In contrast, the results of the simulation for Experiment 2 failed to correctly model the structural differences, though they are interesting in predicting a complexity effect modulated by similarity. In these simulations, the critical retrieval for both DP and TP targets was very accurate in object control conditions, consistent with the experimental evidence. The model incorrectly predicted ECM sentences to be easier than S-complement sentences, though, the inverse of the observed pattern. This was true for both simple and complex intervener conditions, and in the retrievals of both the subject head and the clausal node, though in the reading times results of Experiment 2 structural differences were restricted to complex intervener conditions only.

The simulation of Experiment 2 correctly predicted an effect of intervener complexity, in line with the first simulation. I take this as suggestive evidence that an activation-based approach can model the complexity effect as elaborative processing, using minimal additional assumptions. Moreover, the second simulation correctly predicted that the effect of complexity can take the form of either an increase or decrease in accuracy. This is an interesting result, and perhaps consitutes the strongest success of the simulations presented here. However, the direction of the complexity effect on retrieval accuracy went in the wrong direction, with complex interveners reducing accuracy in ECM sentences, but increasing accuracy in S-complement sentences. Here again the effect seems to be driven by the model over-estimating recency effects, in that complexity facilitates retrieval insofar as it decreases the delay between the critical retrieval and an intermediate retrieval.

Finally, the third simulation, which used the materials of Experiment 4, failed to predict any effect of complexity at all. Here the lack of an effect was driven by the fact that the elaborative retrieval erroneously retrieved the matrix subject, which had high activation due to attachment of the relative clause. I do not consider this an issue with the retrieval architecture. Rather, it arises because the only means of reactivating an item from memory is retrieval, which is potentially error-prone. While it is likely that processing a phrase such as *the ancient alien monster* involves a momentary mis-analysis, such as taking the first noun as the head, it seems unlikely that such phrases would be prone to misanalyses that attach the modifier elements to a structurally distant head noun, like the subject of a higher clause.

In closing, I would like to suggest that many of the issues with the current simulations could be overcome by altering the constituents' features or the schedule of retrieval. I have deliberately avoided doing so in an effort to restrict the degrees of freedom open to the modeler. It is tempting to adjust these simulations, but in so doing we open ourselves to the criticism that we are simply tuning the model to fit the data. I am more inclined to make only minor adjustments, with the minimum of representational assumptions, and view the failures of the model as more instructive than its successes.

Conclusions 6

The empirical focus of this dissertation has been how linguistic representations are encoded and retrieved from memory in the course of sentence processing. I reported the results of a series of self-paced reading and eye-tracking experiments on the retrieval of subjects in two situations: subject-verb attachment and controlled PRO. The experiments manipulated the similarity between a target matrix subject and a potentially-interfering, intervening subject. We also examined the effects of constituent complexity on encoding and retrieval processes by crossing similarity manipulations with manipulations that varied the syntactic/semantic complexity of the intervening subject. The goal of these experiments was to use interference effects to diagnose how the retrieval cues characterize subject constituents. More generally, the objective of this dissertation was to fortify existing proposals that sentence comprehension is supported by a content-addressable memory architecture with a theory of retrieval cues.

6.1 Complex subject attachment

6.1.1 Structural and thematic subjects (Experiment 1)

Increasing the complexity of a subject leads to processing difficulty at its verb – the putative retrieval site – when the complex subject contains constituents that are syntactically and semantically similar to the target subject (Van Dyke & Lewis, 2003; Van Dyke, 2007). This experiment extended these findings to stimuli that control for depth of embedding, holding constant both the number of syntactic subjects and the number of clauses.

(136) The hostess thought that the (chef's) careful preparation of the blowfish delayed the guest.

(137) The hostess thought that the guest delayed the (chef's) careful preparation of the blowfish.

We found that the processing at the verb was more difficult when the potentially interfering subject was an event nominalization (136) than when it was a simple *det-noun* constituent (137).

Crucially, though, there was no additional difficulty when the nominalization also contained a possessor, a thematic subject.

We concluded that subjects are retrieved using cues that target their syntactic properties, such as case and/or structural position, rather than their thematic properties. This conclusion was also supported by the offline comprehension data. Comprehension question accuracy for sentences containing a potentially-interfering subject (whether simple or complex) was significantly lower than control sentences that contained only null subjects.

6.1.2 *Case & syntactic position (Experiments 2–3)*

The observed interference from embedded syntactic, but not thematic, subjects is consistent with retrieval cues for syntactic position and/or case. These cues were examined in Experiments 2–3, which probed for interference from embedded subjects such as those shown in (138–140). In these sentences, the critical subject, *Freyja*, occupies the structural subject position (SPEC-T) and bears nominative case. Potential inteference was manipulated by varying the case and/or syntactic position of the embedded (italicized) subject:

(138) Freyja believed that *the captain* was sleeping.

(139) Freyja believed *the captain* to be sleeping.

(140) Freyja persuaded *the captain* to sleep.

As in Experiment 1, we found that complex interveners engendered inhibitory retrieval interference in sentences with an embedded finite clause (138), where both the intervener and target were structural subjects bearing nominative case. However, when the intervener matched the target in structural position, but differed in its case feature, as in sentences with an embedded ECM complement clause (139), complex interveners facilitated processing at the putative retrieval site. The sentences where both intervener and target were structural subjects patterned differently than object control sentences (140), where the intervener was a matrix, accusative object and thus maximally different than the target. In these sentences, increasing the complexity of the intervener had no effect on the processing at the critical retrieval.

The interference from structural subjects, and not objects, was taken as evidence that subject retrieval is guided by a cue for syntactic position. Crucially, the contrasting form of this interference effect – inhibitory interference in S-complement sentences, but facilitatory interference for ECM sentences – was taken to implicate a retrieval cue for case properties. The positional similarity between the intervener and the target leads to interference, but the mismatching case feature in the ECM sentences facilitates retrieval by distinguishing otherwise similar subjects. Thus, the general conclusion of Experiments 2–3 is that the retrieval cues characterize subjects in terms of *both* their case and syntactic position. Importantly, these cues are informed by abstract grammatical knowledge, since none of the stimuli used interveners that overtly encoded case morphology (there were no pronouns). More generally, Experiments 1–3 indicate that interference depends on both the number of subjects and their similarity to the target of retrieval.

6.2 Implicit subjects

Chapter 4 presented the results of two self-paced reading experiments that probed for interference effects in subject retrieval, when that retrieval was triggered by detection of a gap corresponding to the implicit subject (PRO) of a controlled infinitive. A major motivation for these experiments was to probe for subject interference under a different retrieval context, under different grammatical constraints, and in the absence of morphological cues provided by a verb. These experiments also allowed us to separate the retrieval of a subject from other processes triggered by a verb.

Experiment 4 used sentences such as those in (141). The critical retrieval occured at the infinitive, *to*. Interference in the retrieval of the target antecedent was manipulated by varying the structural position of the intervener, which was either a subject or object. As in the previous experiments, the similarity manipulation was crossed with intervener complexity:

(141) The tourist$_i$ …[a,b]… was allowed ___ by the judge [PRO$_i$] to leave.

a. …that the (old forest) guide helped ___ with travel plans…

b. …that ___ helped the (old forest) guide with travel plans…

We saw that the processing at the retrieval site was sensitive to the complexity of structurally illicit interveners. This finding extended the results of Experiments 1–3 to control dependencies. However, in contrast to those experiments, retrieval interfernece in sentences with a complex intervener was facilitatory, and insensitive to the syntactic position (or Case) of the intervener.

I argued that the facilitation effect was due to the presence of a gap in the embedded relative clause that hosted the intervener. This gap was always coreferent with the matrix, target subject, so that processing these sentences required repeatedly reactivating the target prior to the critical retrieval. This reactivation served to make the target subject more salient, facilitating its retrieval. As in the ECM sentences of Experiments 2–3, I argued that complex interveners engendered a distinctiveness effect. In essence, the additional processing required for complex interveners facilitated retrieval by highlighting the differences between the complex constituent and the target of retrieval (see Hofmeister, 2011, for a similar proposal). Thus, the presence of a gap in the relative clause made the target more salient, while the presence of a complex distractor constituent made it more distinct.

6.2.1 Predictability & the challenge of interference

Experiment 5 explored an alternative account of the facilitatory interference observed in Experiment 4. Namely, that lexical information in the matrix control predicate generated an expectation for a PRO subject, with the result that some information was predictively carried forward from the verb, and helped to guide retrieval of the correct antecedent at the gap site.

To examine this possibility, Experiment 5 compared the retrieval of PRO's antecedent in sentences containing a complement control verb (142), where the PRO subject is predictable, to sentences containing a controlled PRO inside an optional – and so unpredictable, by hypothesis – temporal adjunct clause (143).

(142) Complement control

The officer was *allowed* eventually ___ to retire after the scandal.

(143) Adjunct control

The officer was *honored* while ___ retiring after the scandal.

The crucial effect of Experiment 5 was a three-way interaction of structure, target animacy, and distractor animacy. In adjunct control, but not complement control, the effect of target animacy was reduced in sentences sentences when the distractor was animate. In other words, a slowdown for inanimate targets – which make poor controllers – was attenuated when the structurally illicit distractor was animate, so that these sentences were read on par with animate target sentences. I took the effect of structure in this interaction to indicate that the intrusion effect is sensitive to the predictability of the PRO subject, in that the predictability of PRO in complement control sentences helps to mitigate interference from the distractor.

6.3 Modeling results

Chapter 5 presented a model of cue-based retrieval implemented within the ACT-R cognitive architecture (J. R. Anderson & Lebiere, 1998; Lewis & Vasishth, 2005). The currency of this model is *activation*, a measure of representational strength reflecting both the utility of an encoding and the degree to which it matches the retrieval cues. In this model, both the probability of successfully retrieving a constituent from memory and the latency of that retrieval is a function of activation. The item with the highest activation is retrieved, and the higher the activation, the easier this retrieval is.

Activation also provides a straightforward way to capture the effects of complexity. Increasing the complexity of a constituent involves repeatedly reactivating the modified constituent from memory, which increases its activation and helps to distinguish the item in memory (Lewis & Vasishth, 2005; Hofmeister, 2011). In this way, the model also captures the view that distinctiveness and elaboration are the result of additional processing (see Hunt & Worthen, 2006, for more discussion).

The simulation results that while it was possible to retrieve the target subject, due to an activation boost in early processing, but retrieving the embedded subject was highly error-prone. The simulations also over-predicted an effect of recency, in that the most recent DP was often the winning candidate. In contrast, when the target of retrieval was a clausal node, retrieving an expectation for the matrix clause at the verb suffered interference from embedded clauses. I argued that this evidence is consistent with a left-corner parsing algorithm, and that previous findings of difficulty from intervening subjects stems from similarity between the matrix and embedded clauses, and not their subjects per se.

6.4 General conclusions

This thesis is part of a broader trend in sentence processing that hypothesizes that language comprehension, like memory more generally, is supported by a content-addressable memory architecture. In this vein, the results of the experiments presented here are supportive of such a view. Across all of the experiments presented here, we consistently saw that grammatically inappropriate constituents exerted an influence on the processing at the retrieval site. These effects would be unexpected if memory for language utilized a structured search mechanism, which should be much more myopic in its consideration of retrieval candidates.

Making predictions in such a memory architecture requires specific proposals about when retrieval is required and the structure of the retrieval cues. At time of writing, empirically diagnosing when retrieval occurs seems still over the horizon. I think that here we must rely the rich body of work on parsing algorithms to provide independent support. In terms of retrieval cues, though, the possibility of immediate results is much more promising. Linguistic theory has furnished us with a rich body of knowledge about the information relevant to our linguistic capacities. In this dissertation I have attempted to leverage syntactic theory, in particular, to define the space of possible retrieval cues. My hope is that psycholinguistic research can reciprocally inform linguistic theory by showing what information is used by the parser in the course of sentence processing. While it is likely, if not certain, that real-time comprehension processes will utilize information outside the scope of the grammatical system, the simplest model would seem to be one that does not take linguistic competence to be completely distinct from linguistic performance.

One of the more surprising and exciting results of the experiments here to show that the sentence processor makes use of abstract grammatical knowledge, such as abstract case (Experiment 3) or expectations generated from lexical information (Experiment 5). The use of such information provides a means of overcoming the challenge of interference from highly similar linguistic representations, and highlights how domain specific knowledge (linguistic representations and grammar) can operate within more domain-general constraints like content-addressability and limited memory capacity.

In closing, the use of this abstract grammatical knowledge is both surprising and yet expected. As a general conclusion, though, I find it reminiscent of a point made by McCloskey (1997), which I will attempt to paraphrase here. Questions of how and why such sophisticated grammatical knowledge is used are interesting, but are also exactly what syntactic theory would lead us to expect. After all, syntax is precisely a system of dependency formation. A more interesting question, perhaps, and one with more surprising answers, is why the linguistic system should be organized in such an odd manner.

Bibliography

Abney, Steven P & Mark Johnson (1991). "Memory requirements and local ambiguities of parsing strategies." In: *Journal of Psycholinguistic Research* 20.3, pp. 233–250.

Abney, Steven Paul (1987). "The English noun phrase in its sentential aspect." Doctoral dissertation. Massachusetts Institute of Technology.

Aho, Alfred V & Jeffrey D Ullman (1972). *The theory of parsing, translation, and compiling*. Prentice-Hall, Inc.

Altmann, Gerry & Yuki Kamide (1999). "Incremental interpretation at verbs: Restricting the domain of subsequent reference." In: *Cognition* 73.3, pp. 247–264.

Anderson, John Robert (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

— (2005). "Human symbol manipulation within an integrated cognitive architecture." In: *Cognitive science* 29.3, pp. 313–341.

Anderson, John Robert, Daniel Bothell, et al. (2004). "An integrated theory of the mind." In: *Psychological review* 111.4, p. 1036.

Anderson, John Robert & Gordon H Bower (1973). *Human associative memory.* Washington, DC: Winston & Sons.

Anderson, John Robert & Christian J Lebiere (1998). *The atomic components of thought*. Psychology Press.

Anderson, John Robert & Lael J Schooler (1991). "Reflections of the environment in memory." In: *Psychological science* 2.6, pp. 396–408.

Anderson, Michael C & James H Neely (1996). "Interference and inhibition in memory retrieval." In: *Memory* 22, p. 586.

Antón-Méndez, Inés, Janet L Nicol, & Merrill F Garrett (2002). "The relation between gender and number agreement processing." In: *Syntax* 5.1, pp. 1–25.

Baayen, R Harald, Douglas J Davidson, & Douglas M Bates (2008). "Mixed-effects modeling with crossed random effects for subjects and items." In: *Journal of Memory and Language* 59.4, pp. 390–412.

Baayen, R Harald & Petar Milin (2010). "Analyzing reaction times." In: *International Journal of Psychological Research* 3.2, pp. 12–28.

Babyonyshev, Maria & Edward Gibson (1999). "The complexity of nested structures in Japanese." In: *Language*, pp. 423–450.

Badecker, William & Frantisek Kuminiak (2007). "Morphology, agreement and working memory retrieval in sentence production: Evidence from gender and case in Slovak." In: *Journal of Memory and Language* 56.1, pp. 65–85.

Badecker, William & R. L. Lewis (2007). "A new theory and computational model of working memory in sentence production: Agreement errors as failures of cue-based retrieval." In: *20th annual CUNY sentence processing conference. San Diego, La Jolla, CA: University of California*.

Bader, Markus (1994). "Sprachverstehen: Syntax und Prosodie beim Lesen [Language processing: Syntax and prosody in reading]." Unpublished doctoral dissertation, University of Stuttgart.

— (1997). "Syntactic and morphological contributions to processing subject-object ambiguities." In: *Manuscript submitted for publication*.

Bader, Markus & Michael Meng (1999). "Case attraction phenomena in German." In: *Unpublished Manuscript. University of Jena, Jena*.

Bader, Markus, Michael Meng, & Josef Bayer (2000). "Case and reanalysis." In: *Journal of Psycholinguistic Research* 29.1, pp. 37–52.

Barr, Dale J et al. (2013). "Random effects structure for confirmatory hypothesis testing: Keep it maximal." In: *Journal of Memory and Language* 68.3, pp. 255–278.

Bates, Douglas et al. (2013). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.0-5. URL: http://CRAN.R-project.org/package=lme4.

Bayer, Josef, Markus Bader, & Michael Meng (2001). "Morphological underspecification meets oblique case: Syntactic and processing effects in German." In: *Lingua* 111.4, pp. 465–514.

Begg, Ian (1978). "Similarity and contrast in memory for relations." In: *Memory & Cognition* 6.5, pp. 509–517.

Betancort, Moises, Manuel Carreiras, & Carlos Acuña-Fariña (2006). "Processing controlled PROs in Spanish." In: *Cognition* 100.2, pp. 217–282.

Betancort, Moisés, Enrique Meseguer, & Manuel Carreiras (2004). "The empty category PRO: Processing what cannot be seen." In: *The online study of sentence comprehension: Eyetracking, ERPs and beyond*, pp. 95–118.

Bever, Thomas G (1974). "The ascent of the specious, or there'sa lot we don't know about mirrors." In: *Explaining linguistic phenomena*, pp. 173–200.

Bever, Thomas G & Brian McElree (1988). "Empty categories access their antecedents during comprehension." In: *Linguistic Inquiry* 19, pp. 35–43.

Blaubergs, M.S. & M Braine (1974). "Short-term memory limitations on decoding self-embedded sentences." In: *Journal of Experimental Psychology* 102, pp. 745–748.

Blumenthal, Arthur L (1966). "Observations with self-embedded sentences." In: *Psychonomic Science*.

Bock, Kathryn & J Cooper Cutting (1992). "Regulating mental energy: Performance units in language production." In: *Journal of memory and language* 31.1, pp. 99–127.

Bock, Kathryn & Carol A Miller (1991). "Broken agreement." In: *Cognitive psychology* 23.1, pp. 45–93.

Boland, Julie E, Michael K Tanenhaus, & Susan M Garnsey (1990). "Evidence for the immediate use of verb control information in sentence processing." In: *Journal of Memory and Language* 29.4, pp. 413–432.

Boland, Julie E, Michael K Tanenhaus, Susan M Garnsey, & Greg N Carlson (1995). "Verb argument structure in parsing and interpretation: Evidence from wh-questions." In: *Journal of Memory and Language* 34.6, pp. 774–806.

Bresnan, Joan (1982). *The mental representation of grammatical relations*. Vol. 1. The MIT Press.

Chomsky, Noam (1957). *Syntactic structures*. Mouton de Gruyter.

— (1970). "Remarks on Nominalization." In: *Readings in English Transformational Grammar*. Ed. by L. Jacobs & P. Rosenbaum. Waltham, MA: Ginn & Comany, pp. 184–221.

— (1981). *Lectures on government and binding*. Dordrecht: Foris.

— (1986). *Knowledge of language: Its nature, origin, and use*. Greenwood Publishing Group.

— (1995). *The Minimalist program*. Vol. 28. The MIT Press.

— (2008). "On phases." In: *Current Studies in Linguistics Series* 45, p. 133.

Chomsky, Noam & Howard Lasnik (1993). "The theory of principles and parameters." In: *Syntax: An international handbook of contemporary research* 1, pp. 506–569.

Clifton, Charles, Lyn Frazier, & Patricia Deevy (1999). "Feature manipulation in sentence comprehension." In: *Rivista di lingüística* 11.1, pp. 11–40.

Clifton, Charles, Adrian Staub, & Keith Rayner (2007). "Eye movements in reading words and sentences." In: *Eye movements: A window on mind and brain*, pp. 341–372.

Cowan, Nelson (2005). *Working memory capacity*. Psychology Press.

Crain, Stephen & Janet D Fodor (1985). "How can grammars help parsers." In: *Natural language parsing: Psychological, computational, and theoretical perspectives*, pp. 94–128.

Daneman, Meredyth & Patricia A Carpenter (1980). "Individual differences in working memory and reading." In: *Journal of verbal learning and verbal behavior* 19.4, pp. 450–466.

Demestre, Josep & José E García-Albea (2007). "ERP evidence for the rapid assignment of an (appropriate) antecedent to PRO." In: *Cognitive science* 31.2, pp. 343–354.

Demestre, Josep, Sheila Meltzer, et al. (1999). "Identifying the null subject: Evidence from event-related brain potentials." In: *Journal of Psycholinguistic Research* 28.3, pp. 293–312.

Dillon, Brian W (2011). "Structured Access in Sentence Comprehension." Doctoral dissertation. PhD dissertation, University of Maryland, College Park.

Dillon, Brian et al. (2013). "Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence." In: *Journal of Memory and Language* 69.2, pp. 85–103.

Dosher, Barbara A. (1981). "The effects of delay and interference: A speed-accuracy study." In: *Cognitive Psychology* 13.4, pp. 551–582.

Drenhaus, Heiner, Douglas Saddy, & Stefan Frisch (2005). "Processing negative polarity items: When negation comes through the backdoor." In: *Linguistic evidence: Empirical, theoretical, and computational perspectives*, pp. 145–165.

Eberhard, Kathleen M, J Cooper Cutting, & Kathryn Bock (2005). "Making syntax of sense: number agreement in sentence production." In: *Psychological Review* 112.3, p. 531.

Epstein, Michael L, W Daniel Phillips, & Shirley J Johnson (1975). "Recall of related and unrelated word pairs as a function of processing level." In: *Journal of Experimental Psychology: Human Learning and Memory* 1.2, p. 149.

Fanselow, Gisbert et al. (1999). "Optimal parsing: Syntactic parsing preferences and optimality theory." In: *Rutgers Optimality Archive* 367.

Featherston, Samuel et al. (2000). "Brain potentials in the processing of complex sentences: An ERP study of control and raising constructions." In: *Journal of Psycholinguistic Research* 29.2, pp. 141–154.

Federmeier, Kara D (2007). "Thinking ahead: The role and roots of prediction in language comprehension." In: *Psychophysiology* 44.4, pp. 491–505.

Federmeier, Kara D & Marta Kutas (1999). "A rose by any other name: Long-term memory structure and sentence processing." In: *Journal of memory and Language* 41.4, pp. 469–495.

Fedorenko, Evelina, Maria Babyonyshev, & Edward Gibson (2004). "The nature of case interference in on-line sentence processing in Russian." In: *Proceedings of New England Linguistics Society*. Vol. 34, pp. 215–226.

Ferreira, Fernanda & Charles Clifton (1986). "The independence of syntactic processing." In: *Journal of Memory and Language* 25.3, pp. 348–368.

Ford, Marily (1983). "A method for obtaining measures of local parsing complexity throughout sentences." In: *Journal of verbal learning and verbal behavior* 22.2, pp. 203–218.

Franck, Julie, Gabriella Vigliocco, & Janet Nicol (2002). "Subject-verb agreement errors in French and English: The role of syntactic hierarchy." In: *Language and Cognitive Processes* 17.4, pp. 371–404.

Frazier, Lyn, Charles Clifton, & Janet Randall (1983). "Filling gaps: Decision principles and structure in sentence comprehension." In: *Cognition* 13.2, pp. 187–222.

Frazier, Lyn & Keith Rayner (1982a). “Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences.” In: *Cognitive psychology* 14.2, pp. 178–210.

— (1982b). “Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences.” In: *Cognitive psychology* 14.2, pp. 178–210.

Gelman, Andrew & Jennifer Hill (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gordon, Peter C, Randall Hendrick, & Marcus Johnson (2001). “Memory interference during language processing.” In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27.6, p. 1411.

Gordon, Peter C, Randall Hendrick, & William H Levine (2002). “Memory-load interference in syntactic processing.” In: *Psychological Science* 13.5, pp. 425–430.

Grimshaw, Jane (1990). *Argument structure*. the MIT Press.

Grodner, Daniel & Edward Gibson (2005). “Consequences of the serial nature of linguistic input for sentenial complexity.” In: *Cognitive Science* 29.2, pp. 261–290.

Hakes, David T & Helen S Cairns (1970). “Sentence comprehension and relative pronouns.” In: *Perception & Psychophysics* 8.1, pp. 5–8.

Hakes, David T, Judith S Evans, & Linda L Brannon (1976). “Understanding sentences with relative clauses.” In: *Memory & Cognition* 4.3, pp. 283–290.

Hartsuiker, Robert J et al. (2003). “Morphophonological influences on the construction of subject-verb agreement.” In: *Memory & Cognition* 31.8, pp. 1316–1326.

Hemforth, Barbara & Lars Konieczny (2000). *German sentence processing*. Vol. 24. Springer Science & Business Media.

Hofmeister, P. (2011). “Representational complexity and memory retrieval in language comprehension.” In: *Language and cognitive processes* 26.3, pp. 376–405.

Hofmeister, P. & S. Vasishth (2014). “Distinctiveness and encoding effects in online sentence comprehension.” In: *Frontiers in Psychology* 1237.5, pp. 22–54. DOI: 10.3389/fpsyg.2014.01237.

Holmes, Virginia M & J Kevin O’Regan (1981). “Eye fixation patterns during the reading of relative-clause sentences.” In: *Journal of Verbal Learning and Verbal Behavior* 20.4, pp. 417–430.

Hommel, Bernhard (1998). “Event files: Evidence for automatic integration of stimulus-response episodes.” In: *Visual Cognition* 5.1-2, pp. 183–216.

Hudgins, Jo Carol & Walter L Cullinan (1978). “Effects of sentence structure on sentence elicited imitation responses.” In: *Journal of Speech, Language, and Hearing Research* 21.4, pp. 809–819.

Hunt, R Reed (2006). "The concept of distinctiveness in memory research." In: *Distinctiveness and memory*, pp. 3–25.
Hunt, R Reed & James B Worthen (2006). *Distinctiveness and memory*. Oxford University Press.
Inhoff, Albrecht Werner (1984). "Two stages of word processing during eye fixations in the reading of prose." In: *Journal of Verbal Learning and Verbal Behavior* 23.5, pp. 612–624.
Jaeger, T Florian (2008). "Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models." In: *Journal of Memory and Language* 59.4, pp. 434–446.
Jaeger, T Florian, Evelina Fedorenko, & Edward Gibson (2010). "Anti-locality in English: Consequences for theories of sentence comprehension." In: *Manuscript submitted for publication*.
Just, Marcel A, Patricia A Carpenter, & Jacqueline D Woolley (1982). "Paradigms and processes in reading comprehension." In: *Journal of Experimental Psychology: General* 111.2, p. 228.
Kamide, Yuki, Gerry Altmann, & Sarah Haywood (2003). "The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements." In: *Journal of Memory and Language* 49.1, pp. 133–156.
Kamide, Yuki, Christoph Scheepers, & Gerry Altmann (2003). "Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English." In: *Journal of psycholinguistic research* 32.1, pp. 37–55.
Kawasaki, Noriko (1993). "Control and arbitrary interpretation in English." Doctoral dissertation. University of Massachussetts.
King, Jonathan & Marcel Adam Just (1991). "Individual differences in syntactic processing: The role of working memory." In: *Journal of memory and language* 30.5, pp. 580–602.
King, Joseph, Caroline Andrews, & M. Wagers (2012). "Do reflexives always find a grammatical antecedent for themselves." In: *The 25th annual CUNY conference on human sentence processing, CUNY, NYC*.
Kuno, S. (1973). *The structure of the Japanese language*. Cambridge, MA: MIT Press.
Kwon, Nayoung & Patrick Sturt (2014). "The use of control information in dependency formation: An eye-tracking study." In: *Journal of Memory and Language* 73, pp. 59–80.
Ladusaw, William & David Dowty (1988). "Toward a nongrammatical account of thematic roles." In: *Syntax and semantics* 21, pp. 61–73.
Larkin, Willard & David Burns (1977). "Sentence comprehension and memory for embedded structure." In: *Memory & Cognition* 5.1, pp. 17–22.
Lau, Ellen et al. (2006). "The role of structural prediction in rapid syntactic analysis." In: *Brain and Language* 98.1, pp. 74–88.
Levy, Roger, Evelina Fedorenko, & Edward Gibson (2013). "The syntactic complexity of Russian relative clauses." In: *Journal of Memory and Language* 69.4, pp. 461–495.

Lewis, R. L. (1993). “An architecturally-based theory of human sentence comprehension.” Doctoral dissertation. Carnegie Mellon University.
— (2000). “Specifying architectures for language processing: Process, control, and memory in parsing and interpretation.” In: *Architectures and Mechanisms for Language Processing*. Ed. by Matthew W. Crocker, Martin Pickering, & Charles Clifton Jr. Cambridge: Cambridge University Press.
Lewis, R. L. & Mineharu Nakayama (2002). “Syntactic and positional similarity effects in the processing of Japanese embeddings.” In: *Sentence processing in East Asian languages*, pp. 85–110.
Lewis, R. L. & S. Vasishth (2005). “An activation-based model of sentence processing as skilled memory retrieval.” In: *Cognitive Science* 29.3, pp. 375–419.
Lewis, R. L., S. Vasishth, & Julie A Van Dyke (2006). “Computational principles of working memory in sentence comprehension.” In: *Trends in Cognitive Sciences* 10.10, pp. 447–454.
Logačev, Pavel & S. Vasishth (2012). “Case Matching and Conflicting Bindings Interference.” In: *Case, Word Order and Prominence*. Springer, pp. 187–216.
Manzini, Maria Rita (1983). “On control and control theory.” In: *Linguistic inquiry*, pp. 421–446.
Marks, Lawrence E (1968). “Scaling of grammaticalness of self-embedded English sentences.” In: *Journal of Verbal Learning and Verbal Behavior* 7.5, pp. 965–967.
Mauner, Gail, Michael K Tanenhaus, & Greg N Carlson (1995). “Implicit arguments in sentence processing.” In: *Journal of Memory and Language* 34.3, pp. 357–382.
McCloskey, James (1997). “Subjecthood and Subject Positions.” In: *Elements of Grammar*. Ed. by Liliane Haegeman. Kluwer Academic Publishers, 197–235.
McElree, Brian (1998). “Attended and non-attended states in working memory: Accessing categorized structures.” In: *Journal of Memory and Language* 38.2, pp. 225–252.
— (2000). “Sentence comprehension is mediated by content-addressable memory structures.” In: *Journal of Psycholinguistic Research* 29.2, pp. 111–123.
— (2006). “Accessing recent events.” In: *Psychology of learning and motivation* 46, pp. 155–200.
McElree, Brian & Thomas G Bever (1989). “The psychological reality of linguistically defined gaps.” In: *Journal of Psycholinguistic Research* 18.1, pp. 21–35.
McElree, Brian & Barbara A. Dosher (1989). “Serial position and set size in short-term memory: The time course of recognition.” In: *Journal of Experimental Psychology: General* 118.4, p. 346.
McElree, Brian, Stephani Foraker, & Lisbeth Dyer (2003). “Memory structures that subserve sentence comprehension.” In: *Journal of Memory and Language* 48.1, pp. 67–91.
Meng, Michael & Markus Bader (2000). “Ungrammaticality detection and garden path strength: Evidence for serial parsing.” In: *Language and Cognitive Processes* 15.6, pp. 615–666.

Meseguer, Enrique, Manuel Carreiras, & Charles Clifton (2002). "Overt reanalysis strategies and eye movements during the reading of mild garden path sentences." In: *Memory & Cognition* 30.4, pp. 551–561.

Miller, George A. (1957). "The magic number seven, plus or minus two." In: *The Psychological Review* 63, pp. 81–97.

Miller, George A, Eugene Galanter, & Karl H Pribram (1960). "Plans and the structure of behavior." In:

Miller, George A & Stephen Isard (1964). "Free recall of self-embedded English sentences." In: *Information and Control* 7.3, pp. 292–303.

Nairne, James S (1990). "A feature model of immediate memory." In: *Memory & Cognition* 18.3, pp. 251–269.

— (2002). "The myth of the encoding-retrieval match." In: *Memory* 10.5-6, pp. 389–395.

— (2006). "Modeling distinctiveness: Implications for general memory theory." In: *Distinctiveness and memory*. Ed. by R.R. Hunt & J. Worthen. New York: Oxford University Press, pp. 27–46.

Newell, Allen (1980). "Physical symbol systems." In: *Cognitive science* 4.2, pp. 135–183.

— (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Newell, Allen, Herbert Alexander Simon, et al. (1972). *Human problem solving*. Vol. 104. 9. Prentice-Hall Englewood Cliffs, NJ.

Nicol, Janet & Lee Osterhout (1988). "Reactivating antecedents of empty categories during parsing." In: *Unpublished manuscript*.

Nicol, Janet & David Swinney (1989). "The role of structure in coreference assignment during sentence comprehension." In: *Journal of psycholinguistic research* 18.1, pp. 5–19.

Osterhout, L & J Nicol (1988). "The time-course of antecedent activation following empty subjects." In: *Unpublished manuscript*.

Parker, Daniel J (2014). "The cognitive basis for encoding and navigating linguistic structure." Doctoral dissertation. University of Maryland, College Park. URL: http://hdl.handle.net/1903/15768.

Pearlmutter, Neal J, Susan M Garnsey, & Kathryn Bock (1999). "Agreement processes in sentence comprehension." In: *Journal of Memory and language* 41.3, pp. 427–456.

Phillips, Colin (2006). "The real-time status of island phenomena." In: *Language*, pp. 795–823.

Pickering, Martin & Guy Barry (1991). "Sentence processing without empty categories." In: *Language and Cognitive Processes* 6.3, pp. 229–259.

Pollard, Carl & Ivan Sag (1994). *Head-driven phrase structure grammar*. University of Chicago Press.

Pylyshyn, Zenon W (1980). "Computation and cognition: Issues in the foundations of cognitive science." In: *Behavioral and Brain Sciences* 3.01, pp. 111–132.

Pylyshyn, Zenon Walter (1984). *Computation and cognition*. Cambridge Univ Press.
R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.
Rayner, Keith (1977). "Visual attention in reading: Eye movements reflect cognitive processes." In: *Memory & Cognition* 5.4, pp. 443–448.
Rayner, Keith, Sara C Sereno, & Gary E Raney (1996). "Eye movement control in reading: a comparison of two types of models." In: *Journal of Experimental Psychology: Human Perception and Performance* 22.5, p. 1188.
Reed, Adam V (1973). "Speed-accuracy trade-off in recognition memory." In: *Science* 181.4099, pp. 574–576.
— (1976). "List length and the time course of recognition in immediate memory." In: *Memory & Cognition* 4.1, pp. 16–30.
Resnik, Philip (1992). "Left-corner parsing and psychological plausibility." In: *Proceedings of the 14th conference on Computational linguistics-Volume 1*. Association for Computational Linguistics, pp. 191–197.
Roberts, Rose & Edward Gibson (2002). "Individual differences in sentence memory." In: *Journal of Psycholinguistic Research* 31.6, pp. 573–598.
Rosenbaum, Peter Steven (1965). "The grammar of English predicate complement constructions." Doctoral dissertation. Massachusetts Institute of Technology.
Schlesewsky, Matthias (1996). "Kasusphänomene in der Sprachverarbeitung." Doctoral dissertation. Doctoral dissertation, University of Potsdam.
Shepard, Roger N. et al. (1987). "Toward a universal law of generalization for psychological science." In: *Science* 237.4820, pp. 1317–1323.
Staub, Adrian (2011). "Word recognition and syntactic attachment in reading: Evidence for a staged architecture." In: *Journal of Experimental Psychology: General* 140.3, pp. 407–433.
Sternberg, Robert & Karin Sternberg (2016). *Cognitive psychology*. Wadsworth.
Sternberg, Saul (1966). "High-speed scanning in human memory." In: *Science* 153.3736, pp. 652–654.
— (1969a). "Memory-scanning: Mental processes revealed by reaction-time experiments." In: *American Scientist*, pp. 421–457.
— (1969b). "The discovery of processing stages: Extensions of Donders' method." In: *Acta psychologica* 30, pp. 276–315.
— (1975). "Memory scanning: New findings and current controversies." In: *The Quarterly journal of experimental psychology* 27.1, pp. 1–32.
Stowe, Laurie A (1986). "Parsing WH-constructions: Evidence for on-line gap location." In: *Language and cognitive processes* 1.3, pp. 227–245.

Sturt, Patrick (2003). "The time-course of the application of binding constraints in reference resolution." In: *Journal of Memory and Language* 48.3, pp. 542–562.

Sturt, Patrick & Nayoung Kwon (2015). "The processing of raising and nominal control: an eye-tracking study." In: *Frontiers in psychology* 6.

Sturt, Patrick, Martin J Pickering, & Matthew W Crocker (1999). "Structural change and reanalysis difficulty in language comprehension." In: *Journal of Memory and Language* 40.1, pp. 136–150.

Tabor, Whitney, Bruno Galantucci, & Daniel Richardson (2004). "Effects of merely local syntactic coherence on sentence processing." In: *Journal of Memory and Language* 50.4, pp. 355–370.

Tulving, Endel & Zena Pearlstone (1966). "Availability versus accessibility of information in memory for words." In: *Journal of Verbal Learning and Verbal Behavior* 5.4, pp. 381–391.

Unsworth, Nash et al. (2005). "An automated version of the operation span task." In: *Behavior research methods* 37.3, pp. 498–505.

Van Dyke, Julie A (2007). "Interference effects from grammatically unavailable constituents during sentence processing." In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33.2, p. 407.

Van Dyke, Julie A & R. L. Lewis (2003). "Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities." In: *Journal of Memory and Language* 49.3, pp. 285–316.

Van Dyke, Julie A & Brian McElree (2006). "Retrieval interference in sentence comprehension." In: *Journal of Memory and Language* 55.2, pp. 157–166.

— (2011). "Cue-dependent interference in comprehension." In: *Journal of Memory and Language* 65.3, pp. 247–263.

Von Restorff, Hedwig (1933). "Über die wirkung von bereichsbildungen im spurenfeld." In: *Psychologische Forschung* 18.1, pp. 299–342.

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

— (2014). "Memory mechanisms for wh-dependency formation and their implications for islandhood." In: *Experimental Syntax and Island Effects*. Ed. by J. Sprouse & N. Hornstein. Cambridge.

Wagers, M., Ellen F Lau, & Colin Phillips (2009). "Agreement attraction in comprehension: Representations and processes." In: *Journal of Memory and Language* 61.2, pp. 206–237.

Wagers, M. & Colin Phillips (2009). "Multiple dependencies and the role of the grammar in real-time comprehension." In: *Journal of Linguistics* 45.02, pp. 395–433.

Wang, Marilyn D (1970). "The role of syntactic complexity as a determiner of comprehensibility." In: *Journal of Verbal Learning and Verbal Behavior* 9.4, pp. 398–404.

Wickelgren, Wayne A & Albert T Corbett (1977). "Associative interference and retrieval dynamics in yes-no recall and recognition." In: *Journal of Experimental Psychology: Human Learning and Memory* 3.2, p. 189.

Wickelgren, Wayne A, Albert T Corbett, & Barbara A. Dosher (1980). "Priming and retrieval from short-term memory: A speed accuracy trade-off analysis." In: *Journal of Verbal Learning and Verbal Behavior* 19.4, pp. 387–404.