

Memory mechanisms supporting syntactic comprehension

David Caplan · Gloria Waters

Published online: 15 January 2013
© Psychonomic Society, Inc. 2013

Abstract Efforts to characterize the memory system that supports sentence comprehension have historically drawn extensively on short-term memory as a source of mechanisms that might apply to sentences. The focus of these efforts has changed significantly in the past decade. As a result of changes in models of short-term working memory (ST-WM) and developments in models of sentence comprehension, the effort to relate entire components of an ST-WM system, such as those in the model developed by Baddeley (Nature Reviews Neuroscience 4: 829–839, 2003) to sentence comprehension has largely been replaced by an effort to relate more specific mechanisms found in modern models of ST-WM to memory processes that support one aspect of sentence comprehension—the assignment of syntactic structure (parsing) and its use in determining sentence meaning (interpretation) during sentence comprehension. In this article, we present the historical background to recent studies of the memory mechanisms that support parsing and interpretation and review recent research into this relation. We argue that the results of this research do not converge on a set of mechanisms derived from ST-WM that apply to parsing and interpretation. We argue that the memory mechanisms supporting parsing and interpretation have features that characterize another memory system that has been postulated to account for skilled performance—long-term working memory. We propose a model of the relation of different aspects of parsing and interpretation to ST-WM and long-term working memory.

Keywords Language/memory interactions · Language comprehension

Language comprehension requires memory over time scales ranging from deciseconds (in the case of sound categorization and word recognition) through several seconds and more (to accomplish sentence and discourse comprehension). This article critically reviews recent work on the memory system that supports one aspect of sentence comprehension—the assignment of syntactic structure (parsing) and its use in determining aspects of sentence meaning (interpretation)—and presents a new view of the memory systems that support these aspects of sentence comprehension.

Background

Even adjacent words must be assigned a syntactic relationship, and syntactic relations often span many words. For instance, in (1), the subject and object of *grabbed*, the subject of *lost*, and the antecedent of *his* must be retrieved at the points at which *grabbed*, *lost*, and *his* are encountered (or later, if parsing and interpretation is deferred):

1. The boy who the girl who fell down the stairs grabbed lost his balance.

Historically, the memory system that has been most often connected to parsing and interpretation is *short-term memory* (STM). The hypothesis that the memory system that supports parsing and interpretation utilizes STM is intuitively appealing because the temporal intervals over which parsing and interpretation usually apply are roughly the same as those over which STM operates. In addition, STM is an appealing construct to apply to sentence memory because it is thought to have capacity and temporal limitations that might account for the difficulty of comprehending certain sentences.

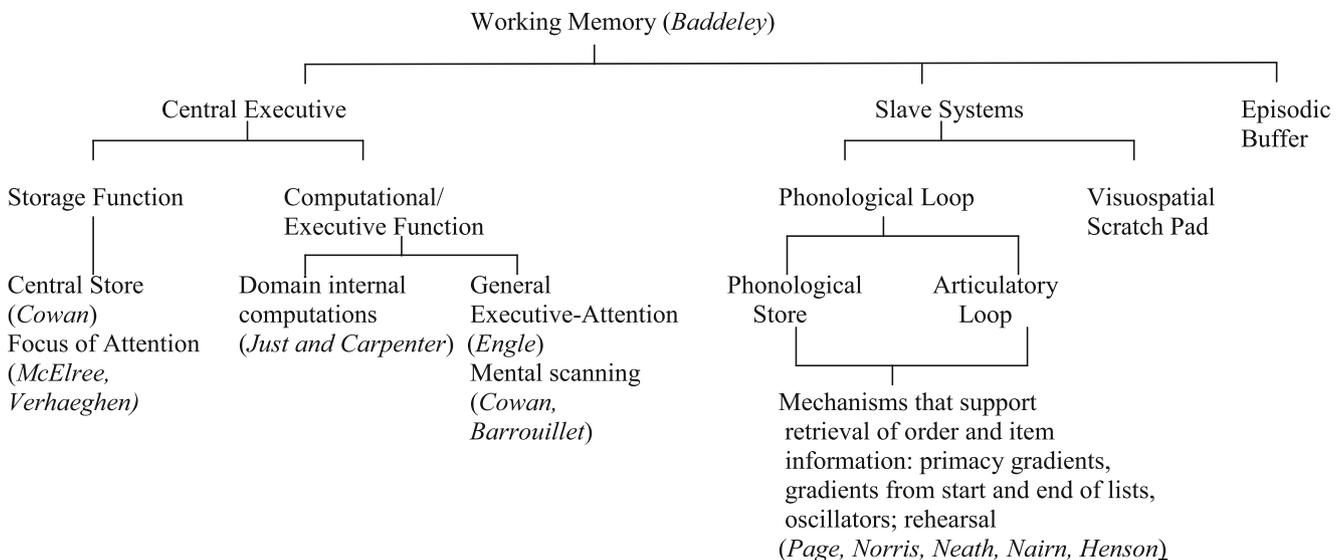
D. Caplan (✉)
Neuropsychology Laboratory, Department of Neurology,
Massachusetts General Hospital, 175 Cambridge Street, Suite 340,
Boston, MA 02114, USA
e-mail: dcaplan@partners.org

G. Waters
Department of Speech and Hearing Sciences, Sargent College,
Boston University, Boston, MA, USA

However, connecting STM to the memory system that supports parsing and interpretation has proven to be difficult. Kane, Conway, Hambrick, and Engle (2007) introduced a chapter on variability in working memory with the comment that “failed attempts to link STM to complex cognitive functions, such as reading comprehension, loomed large in Crowder’s (1982) obituary for the concept” (p. 21). Kane et al. went on to say that Baddeley and Hitch (1974) “tried to validate immediate memory’s functions” by introducing the concept of working memory (which we will call *short-term working memory* [ST-WM]). Evidence for a functional role for ST-WM came from interference from concurrent six-, but not three-, item memory loads in reasoning, comprehension, and learning tasks, which

suggested that “small memory loads are handled by a phonemic buffer . . . whereas larger loads require the additional resource of a central executive. Thus working memory was proposed to be a dynamic system that enabled maintenance of task-relevant information in support of the simultaneous execution of complex cognitive tasks” (Kane et al., 2007, p. 21).

Baddeley’s model of ST-WM represents one influential model among many models of short-term and working memory. We begin this article with an overview of models of these memory systems, beginning with Baddeley’s, which was directly related to sentence comprehension. A guide to terminology and the evolution of models is found in Box 1.



Box 1 Components and processes in models of short-term working memory. This chart summarizes features of models of short term memory, illustrating how components of Baddeley’s “working memory” model are related to other constructs. The outline does not present all models and all mechanisms; it shows aspects of models that are mentioned in the text.

Baddeley’s initial model of ST-WM (e.g., Baddeley, 1986) contained two major components. A *central executive* (CE) maintained multidimensional representations and was also considered to have some computational functions (see below). Visuospatial and verbal *slave systems* maintained domain-specific representations. The verbal slave system, the *phonological loop* (PL), consisted of two components: a *phonological store* (PS) that maintained information in phonological form subject to rapid decay and an articulatory mechanism that rehearsed items in the PS and transcoded written verbal stimuli into phonological form. Baddeley (2000) introduced a third type of store: the *episodic buffer* (EB), which retained integrated units of visual, spatial, and

verbal information marked for temporal occurrence. The slave systems and the EB had no computational functions themselves.

From approximately 1980 to 2000, a number of researchers sought to relate the memory system that supports aspects of sentence comprehension to components of the model of ST-WM developed by Baddeley and his colleagues. The main question that was considered was what component(s) of the ST-WM system, if any, supported aspects of sentence memory; as has been noted, our focus is on the sentence memory that supports syntactic comprehension. Either the CE or the PL could play this role (the role of the EB, which had not been introduced at the time much of this work was

done, has not been investigated). The CE maintains abstract representations that could include syntactic and semantic information, making it suitable for this purpose. The PL maintains phonological representations that are linked to lexical items that contain the needed semantic and syntactic information, and it might be easier for the memory system to maintain such representations in the PL as pointers to the needed information than to maintain what are arguably more complex representations in the CE.¹

Data regarding these possibilities come from several sources: interference effects of concurrent tasks that require the CE or the PL on parsing and interpretation (King & Just, 1991; Waters, Caplan, & Hildebrandt, 1987); tongue twister effects (Acheson & Macdonald, 2011; Ayres, 1984; Keller, Carpenter, & Just, 2003; Kennison, 2004; McCutcheon, Bell, France, & Perfetti, 1991; McCutcheon & Perfetti, 1982; Perfetti & McCutchen, 1982; Zhang & Perfetti, 1993); effects of homophones on comprehension (Coltheart, Patterson, & Leahy, 1994; Waters, Caplan, & Leonard, 1992); correlations between measures of the capacity of the CE and performance in parsing and interpretation (Just & Carpenter, 1992; King & Just, 1991; MacDonald, Just, & Carpenter, 1992; Miyake, Carpenter, & Just, 1994; Waters & Caplan, 1996); associations and dissociations of effects of brain damage on the CE or the PL and parsing and interpretation (Caplan & Waters, 1996; Emery, 1985; Grossman et al., 1991; Grossman, Carvell, Stern, Gollomp, & Hurtig, 1992; Kontiola, Laaksonen, Sulkava, & Erkinjuntti, 1990; Lalami et al., 1996; Lieberman, Friedman, & Feldman, 1990; R. C. Martin, 1990; Natsopoulos et al., 1991; Rochon & Saffran, 1995; Rochon, Waters, & Caplan, 1994; Stevens, Kempler, Andersen, & MacDonald, 1996; Tomoeda, Bayles, Boone, Kaszniak, & Slauson, 1990; Waters, Caplan, & Hildebrandt, 1991; Waters, Caplan, & Rochon, 1995); and neural evidence in the form of differences in neural activation in high- and low-span participants during parsing and interpretation and overlap or nonoverlap of brain areas activated in tasks that involve the CE and ones that involve parsing and interpretation (Fiebach, Schlesewsky, & Friederici, 2001; Fiebach, Vos, & Friederici, 2004). In Caplan and Waters (1990), we argued that these sources of data suggested that

the PL did not support the memory requirements of the initial assignment of the preferred structure and interpretation of a sentence (what we called *first-pass* parsing and interpretation) but may play a role in supporting review and reanalysis of previously encountered information and in using the products of the comprehension process to accomplish a task (e.g., in maintaining a sentence in memory to later answer a question about its meaning). We think that it is fair to say that there is a general consensus that this view is correct; for example, in an article relating the memory system that supports syntactic comprehension to ST-WM, Just and Carpenter (1992) explicitly said that they were not advocating a role for the PL, but only the CE, in this process. The role of the CE in parsing and interpretation has been more controversial. In Caplan and Waters (1999), we argued that evidence of the sort described above indicated that, like the PL, the CE does not support first-pass parsing and interpretation, but other researchers have disagreed (Just & Carpenter, 1992).

We will not revisit this literature because developments in models of ST-WM have radically changed constructs in Baddeley's model of ST-WM. Phonological effects in immediate serial recall and other STM tasks have been discredited as evidence for storage of purely phonological representations (Jones, Hughes, & Macken, 2006, 2007; Jones, Macken, & Nicholls, 2004; see Caplan, Waters, & Howard, 2012, for a review). Neuropsychological data have provided evidence that lexical representations, including semantic features, not purely phonological representations, are stored in ST-WM (N. Martin & Ayala, 2004). These changes have effectively eliminated the PS. Duration-based word length effects in immediate serial recall have been found to be due to lexical properties of words—in particular, neighborhood size (Jalbert, Neath, & Surprenant, 2011)—undermining a critical piece of evidence for a role for articulatory-based rehearsal in sustaining decaying phonological representations. Some contemporary models of ST-WM include a rehearsal mechanism that applies to multidimensional representations (Burgess & Hitch, 1992, 1999; Page & Henson, 2001), a significant departure from the PL model. Many modern models postulate an important role for a mental refreshing process in maintaining representations in ST-WM (Barrouillet, Bernardin, & Camos, 2004). This mechanism is not part of Baddeley's PL.

The notion of the CE has also undergone changes. One line of research has investigated its computational function. Baddeley and Hitch (1974) limited the computational role of the CE to control processes, attributing domain-specific computations to domain-specific processors. Most recent work has accepted this view and has explored the role of

¹ A store that is not part of Baddeley's formulation of ST-WM but that has been argued to play a role in sentence comprehension is a "short term semantic store" postulated by R. Martin and her colleagues (R. C. Martin, Shelton, & Yaffee, 1994). Martin and He (2004) have argued that the "short term semantic store" is used when items in a sentence are not assigned propositional meaning, as in pre-nominal sequences of adjectives (*the old, rusty, twisted . . . N*). This would be at most a minor role in sentence comprehension, given the incremental nature of comprehension.

domain-independent executive functions of the CE in supporting performance on memory tasks (Barrouillet et al., 2004; Engle, Kane, & Tuholski, 1999; Engle, Tuholski, Laughlin, & Conway, 1999; Oberauer & Lewandowsky, 2011). The *mental refreshing* process mentioned above can be seen as one way the CE supports ST-WM; the CE has also been said to support ST-WM by strategically transferring items between *primary* and *secondary* memory (Engle, Kane, & Tuholski, 1999; Engle, Tuholski, et al., 1999). Engle and his colleagues have argued that the functional role of the domain-independent executive functions of the CE extends beyond supporting performance on memory tasks to the point of being a major determinant of “fluid intelligence.” They have suggested that the combination of storage and the executive functions of task goal maintenance and inhibition of competing responses constitute an integrated “executive attention” capacity that “is specifically responsible for the covariation between measures of WM and higher order cognition” (Kane & Engle, 2003, p. 47).

Other investigators have conceived of the computational functions of the CE as supporting domain-specific cognitive functions as well as domain-independent executive attention (Craig & Lewandowsky, *in press*). In the area of sentence comprehension, Just and Carpenter (1992) modeled behavioral results in sentence comprehension in terms of the load exerted on a single “resource” by both memory and computational requirements of parsing and interpretation and argued that that resource was the CE. Just and Carpenter developed a model in which resources could be flexibly allocated to computations (parsing operations) or storage to minimize the effects of overload. However, there have been no other models of this type that specifically relate both computational and storage functions of the CE to parsing and interpretation.

The aspect of the CE that has been examined in recent studies and that has been related to parsing is its memory component. As was noted, Baddeley’s CE includes a limited-capacity memory store of multidimensional representations, which can be accessed by the parser. This store, which has been named the *central store* (CS; Cowan, 2000; Ricker, AuBuchon, & Cowan, 2010), has been studied separately from the attentional/executive/computational functions of the CE. Entry of items into the CS depend upon attention (Cowan, 1995), but this aspect of the functioning of the CS has not been related to memory used in sentence comprehension (all recent studies have involved attended sentences). The capacity of the CS and the nature of retrieval of items from the CS have been related to the memory processes used in sentence comprehension—in particular, the memory system that supports parsing and interpretation. We now therefore turn to studies of the CS.

The Central Store

The CS is a relatively new concept in theories of STM. As has been noted, it does not correspond to any construct in Baddeley’s working memory model, differing from the CE in not having an executive or computational function and from the EB in not binding different items into an episodic unit. It has some similarity to the constructs of *primary memory* (PM) and *secondary memory* (SM) in the pre-Baddeley literature (James, 1890) but differs from both. It is similar to PM in that it is capacity limited. It differs from conceptualizations of PM that maintain that PM contains phonological representations (Colle, 1980; Colle & Welsh, 1976; Conrad, 1963, 1964; Shallice, 1975; Shallice & Vallar, 1990) and is more similar to conceptualizations of PM that maintain that PM contains multidimensional representations (Baddeley & Hitch, 1974; Craik, 1970; Craik & Levy, 1970). The CS is similar to SM because it contains multidimensional representations and constitutes the activated portion of long-term memory (LTM; see below), but it differs significantly from SM because of its limited capacity. In this section, we briefly outline recent views regarding the capacity of the CS and retrieval of items in STM tasks.

The capacity of the CS has been estimated on the basis of accuracy and speed of recall data. Using tasks that prevent the persistence of sensory stores and the use of rehearsal, the capacity of the CS based on accuracy has been estimated at three to five items (see Cowan, 2000, for a review). For instance, Saults and Cowan (2007) exposed participants to four spoken digits presented in four different voices by loudspeakers in four spatial positions and four colored spots presented at four different locations in a recognition task. Participants recognized three to four items regardless of whether only the visually presented colors or both the visually presented colors and the auditorily presented digits had to be remembered.

Temporal data result in a similar, although more subtle, picture. Verhaeghen and Basak (2005) found that RTs in the *n*-back task showed an increase from *n* = 1 to *n* = 2 and no changes from *n* = 2 to *n* = 5. On the assumption that shifts to items outside the focus of attention incur a cost seen in reaction times (RTs) when accuracy is near ceiling, they argued that only one item was maintained in the focus of attention in this task. However, Verhaeghen, Cerella, and Basak (2004) showed that the point at which RTs began to increase was increasingly displaced toward *n* = 5 with increasing practice with the task. Verhaeghen et al. (2007) argued that the difference in RTs between *n* = 4 and *n* = 5 even in highly practiced tasks suggests an upper bound of four on the number of items that can be maintained in the focus of attention. This number is consistent with Cowan’s (2000) conclusion regarding the capacity of the CS, although the fact that an expanded focus of attention requires

practice on a task differs from Cowan's (2000) view of the capacity of the CS.

The capacity of ST-WM has also been studied through the analysis of speed–accuracy trade-offs (SATs). In the SAT approach, participants are presented a stimulus, followed by a probe at a variable lag, and are required to make a two-alternative forced new–old choice within a very short time (usually less than 200 ms; responses less than 100 ms are discarded as anticipations). Asymptotic d' is considered to be a measure of the availability of an item, and the temporal point at which d' rises above 0 (the intercept) and the slope of the rise of d' from the point at which it rises above 0 to its asymptotic level reflect how an item is accessed. McElree and Doshier (1989) found that set size (three and five) affected asymptotic accuracy but not temporal dynamics, except for the set-final item, even when the probe was presented in a different font. McElree (1998) found that there were different retrieval speeds for probes drawn from the last and all other categories when triads of words in different semantic categories were presented. McElree (1996) found that recognizing rhyming words and semantically related words resulted in uniformly slower dynamics than did recognizing previously presented words as wholes. McElree (2006) argued that ST-WM consisted of a focus of attention with one item, which was accessed by a *matching* procedure based on abstract (nonsensory) features of the retrieval cue, and items outside the focus of attention, in LTM, that were accessed by a content-addressable retrieval mechanism, with *chunking* of the memory set in accordance with semantic features.

There are important differences among these models regarding the capacity-limited portion of ST-WM. Cowan (2000) argued for a tripartate division of memory, in which items could be in focal attention, in an activated part of LTM outside of focal attention (the CS), or in LTM. McElree (2006) and Verhaeghen et al. (2007) argued for a bipartate division of memory into a state of focal attention and LTM. These models do not include a CS. McElree's (2006) estimate of the capacity of the focus of attention as one chunk is consistent with results for unpracticed trials of the n -back task reported by Verhaeghen and Basak (2005) but is inconsistent with the expanded capacity of this "store" after practice and with the existence of a CS.

The capacity limit of ST-WM, whether the CS or the focus of attention, is important because, as Lewis (1996) has emphasized, syntactic processing requires at least two items. Therefore, both the CS and an expanded focus of attention provide a capacity-limited memory store that could both satisfy the demands of parsing for access to more than one representation and also impose limits on parsing related

to the availability of information, but McElree's (2006) capacity-limited part of ST-WM (the focus of attention), which is limited to one item, cannot support parsing.² McElree's (2006) model therefore entails that parsing involves retrieval from LTM and is not capacity limited, a position that McElree endorses (McElree, personal communication, University of Massachusetts Workshop on Parsing and Memory).

These studies also provide information about the nature of retrieval of items in the different stores available in STM tasks. In general, items could be retrieved by matching features of all items in memory against those of a retrieval cue and selecting the best match (*content-addressable retrieval*) or by searching for one or more items with particular features (*search*). Search could proceed in many ways. We consider items inside and outside the limited-capacity portion of ST-WM separately.

McElree (2006) suggested that items within the focus of attention were accessed by direct matching. Verhaeghen et al. (2007) considered items in their expanded focus of attention by analyzing the time to report items in the practiced trials of the n -back task reported in Verhaeghen et al. (2004). There was a 30-ms/item increase over the $n = 1$ to $n = 4$ range, and an ex-Gaussian decomposition of the RT data indicated that the effect was due to changes in the skew of the distribution of RTs (τ). These features suggest a limited-capacity parallel search process, not a content-addressable process. It has also been argued that items within the focus of attention that are not grouped into a higher level chunk are not content addressable but require search to be retrieved (Basak, 2005; Hockley, 1984, Experiment 1; Trick & Pylyshyn, 1994; Verhaeghen et al., 2007).

The flat RT curves for items outside the focus of attention in the work by Verhaeghen and his colleagues cited above are consistent with a content-addressable retrieval model for those items, as are the uniform SAT dynamics for these items reported by McElree and his colleagues. Verhaeghen et al. (2007) pointed out, however, that substantial slopes of RTs-over- n for n above the step function have been described for older participants (Basak, 2005) and in unpracticed trials in younger participants when the item to be matched could be any one between 1 and n back. These

² In a review of an earlier version of this article, Rick Lewis pointed out that McElree (2006) must claim that at least two items are available in a privileged fashion to account for the list-final effect in probe tasks—the list-final item (the item in the focus of attention) and the probe (the cue). A problem arises in parsing if only the cue and the item in the focus of attention are available, however, because the item in the focus of attention is not always what the cue is looking for. An item may be in linguistic focus but not be the item specified by a cue (*It was the girl whose husband the police questioned*). So, while Lewis is right that McElree's model requires that two items be available in a privileged fashion, this point does not change the fact that retrieval must be of items in LTM if McElree's model is applied to retrieval-based parsing.

slopes are much steeper than the ones found for items within the focus of attention and are best modeled by changes in the skew, mean, and standard deviation in ex-Gaussian models, pointing to a search process. McElree (2001) also studied SAT functions for recognition of items either 2 or 3 back in an n -back test and concluded that the best model of the temporal dynamics in this situation was a mixed model, in which some items were accessed by a content-addressable memory and some by a serial search process, and that the extent to which these two mechanisms applied differed in different individuals. Verhaeghen et al. (2007) concluded that items outside the focus of attention are content addressable only under what they called “ideal” circumstances—precisely, predictable switching out of the focus of attention on the part of high-functioning individuals—and require search under other circumstances. Items outside the focus of attention are also not content addressable but require search when order, as well as item information, must be retrieved (Hockley, 1984, Experiment 3; McElree & Doshier, 1993).

Despite these differences in models and results, two features of modern models have been considered to be well enough established to warrant exploration as memory mechanisms in sentence comprehension: (1) the notion that a small number of items—at most, four or five—are maintained in a highly accessible form and that this number of available items sets a capacity limit in parsing and interpretation and (2) the view that retrieval of information in parsing and interpretation is content addressable. We shall review studies that suggest that these features of memory are found in parsing and interpretation (see the Mechanisms in ST-WM and Parsing and Interpretation: Empirical Results and Interpretation section).

Retrieval-based parsing

A model of parsing and interpretation that specifies when and how memory is utilized is needed to relate ST-WM to parsing and interpretation. Models that specify such mechanisms in parsing and interpretation have recently been developed in work on *retrieval-based parsing*. We shall briefly outline the model developed by Lewis and his colleagues, which is one of the most widely cited models (see Lewis, Vasishth, & Van Dyke, 2006, for a summary).

Lewis and his colleagues’ model utilizes Anderson’s ACT-R framework (Anderson, 1976). Information about syntactic structure is maintained in long-term declarative memory as symbols consisting of feature-value pairs. In Lewis and Vasishth (2005), each symbol represents a maximum syntactic projection (the categories are drawn from Chomsky, 1995), and the features include syntactic information about the category (e.g., whether it takes a specifier, head, complement), as well as other syntactic features. Lexical items, containing syntactic information, are also listed in declarative memory. Parsing begins with lexical access, which places a new word and its features in a lexical buffer. Production rules then combine information in the lexical buffer with information in the accruing syntactic representation, based on information about structures of the language that are maintained in declarative memory. The production rules follow a left-corner parsing process: Given a lexical input and a goal category, the input is replaced with the left side of a rewrite rule that then applies, and the process is repeated until no further rules apply (Fig. 1).

The phrase marker created by a production rule is stored in a problem state buffer, and the categories that are needed

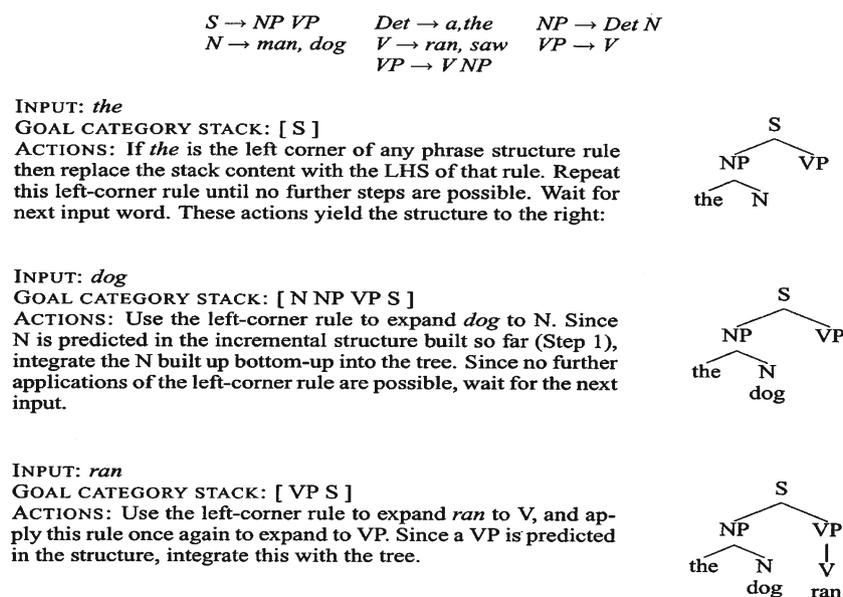


Fig. 1 Rewrite rules from Lewis and Vasishth (2005, Fig. 3)

to complete phrase markers in this buffer are stored in a control buffer. Items in the problem buffer decay over time. The syntactic features of the current lexical entry and the structure in the control buffer combine to form retrieval cues, which may retrieve previously constructed constituents or symbols in LTM. On the basis of the retrieved item and the current lexical content, a production creates a new syntactic structure and attaches it to the retrieved constituent. This increases the activation of previously constructed constituents and thereby partially or, at times, fully counteracts the effect of temporal decay. The control buffer is updated, and a production rule then guides attention to the next word.

The retrieval process for lexical items and syntactic symbols in declarative memory is content addressable. It consists of matching features of a retrieval cue to features of items in the memory set. Retrieval success is determined by the ratio of the *resonance* of the retrieval cue with the retrieved item divided by its resonance with other items in the memory set (Nairne, 1990a; Ratcliff, 1978), subject to noise. This model has two features: (1) Items in the memory set that share features with a retrieval cue interfere with retrieval of a target, and (2) shared features of items in the memory set do not affect retrieval unless these features are part of the retrieval cue.

These developments in models of ST-WM and parsing specify mechanisms in ST-WM that can apply to parsing and interpretation. Work in the past 15 years has made connections between these mechanisms and parsing and interpretation. In the next section, we review research into three phenomena—interference effects, retrieval dynamics, and capacity limitations—that have been interpreted as providing a mutually consistent set of results that establish such links. In the A Critical Look at the Evidence Regarding Features of Retrieval-Based Parsing section, we will argue that, despite its important contributions to describing and modeling new phenomena, the effort to relate these phenomena to mechanisms in ST-WM encounters significant problems. In the An Alternative Framework for Viewing the Memory System for Parsing and Interpretation section, we outline a different perspective on the memory system that supports parsing and interpretation.

Mechanisms in ST-WM and parsing and interpretation: Empirical results and interpretation

In this section, we selectively review representative results that bear on the nature of memory processes in parsing and interpretation. The range of results is summarized in Box 2, along with issues we shall raise in later sections.

Box 2 Phenomena and issues in studies of the memory processes that support parsing and interpretation (with sections of the article where the topic is discussed)

Empirical phenomenon	Issues raised
I. Semantic interference effects (Semantic Similarity Effects)	
a. of items in lists (Similarity of External Load Load to Sentence-Internal Items)	<ul style="list-style-type: none"> - statistically insignificant results (Empirical Issues in Studies Demonstrating Semantic Interference Effects) - unexpected results (Empirical Issues in Studies Demonstrating Semantic Interference Effects) - consistent with source memory confusion (The Number of Memory Systems Needed to Account for Effects of External Load)
b. of items in sentences (Similarity of Sentence-Internal Items to One Another)	<ul style="list-style-type: none"> - requires chaining model of similarity effects if effect arises in recall (Semantic Similarity Effects and Retrieval) - consistent with retrieval of item and order information, unlike SAT results (Semantic Similarity Effects and Retrieval) - effect may arise in encoding, which is unlike encoding in ST-WM (Semantic Similarity Effects and Retrieval)
II. Syntactic Interference Effects (Syntactic Interference Effects)	<ul style="list-style-type: none"> - not found at times where expected (Empirical Issues in Studies Demonstrating Syntactic Interference Effects) - selectivity of effects inconsistent with solely content-addressable retrieval (Syntactic Interference Effects and Retrieval Cues)
III. SAT results (Retrieval Dynamics)	<ul style="list-style-type: none"> - potential task effects, enhanced by strategic effects (Empirical Issues in Studies Using the SAT Technique) - requires retrieval of only item information, unlike I.b (Semantic Similarity Effects and Retrieval) - reflects access of items in LTM, not the CS (The Role of Capacity and Temporal Limits in Parsing and Interpretation and Their Relation to ST-WM)
IV. Capacity Limitations (Capacity Limits and Temporal Decay)	<ul style="list-style-type: none"> - capacity limits in parsing models not those found in models of ST-WM (The Role of Capacity and Temporal Limits in Parsing and Interpretation and Their Relation to ST-WM)

Empirical phenomenon	Issues raised
V. Temporal Decay (Capacity Limits and Temporal Decay)	<ul style="list-style-type: none"> - chunks in capacity-limited parsing models are transparent (The Role of Capacity and Temporal Limits in Parsing and Interpretation and Their Relation to ST-WM) - existence of temporal decay in ST-WM questionable (The Role of Capacity and Temporal Limits in Parsing and Interpretation and Their Relation to ST-WM)
VI. Other issues	<ul style="list-style-type: none"> - some mechanisms described in ST-WM not found in parsing/interpretation (Overlap and Nonoverlap of Retrieval Mechanisms in ST-WM and in Retrieval-Based Parsing) - retrieval in parsing/interpretation differs from retrieval in recognition and recall (Overlap and Nonoverlap of Retrieval Mechanisms in ST-WM and in Retrieval-Based Parsing)

Interference effects in parsing and interpretation

The first set of studies we shall review are ones that document what have been interpreted as interference effects in retrieval during parsing and interpretation. Two types of interference effects have been described in parsing and interpretation and related to retrieval mechanisms described in ST-WM: those due to semantic features of the items in memory and those due to syntactic features of those items.

Semantic similarity effects

Effects of the semantic similarity of noun phrases on online sentence processing measures have been described for both sentence-internal and sentence-external noun phrases.

Similarity of external load to sentence-internal items Two studies have been said to show that semantic properties of noun phrases in a sentence and in a to-be-recalled list affect processing times for verbs of relative clauses more in object- than in subject-relative clauses. On the assumption that the verb of an object-relative clause creates a retrieval cue for the noun phrases that serve as its subject and object, whereas the verb of a corresponding subject-relative clause cues only retrieval of its subject, a superadditive interaction between semantic properties (similar, dissimilar) and sentence type (subject or object relative) has been interpreted as being due

to interference of semantically similar nouns on retrieval in parsing and interpretation.

Gordon et al. (2002) presented word lists consisting of either three proper nouns (Joel, Andy, Greg) or three common nouns (banker, lawyer, accountant) for recall while participants self-paced their way through cleft subject and cleft object sentences (2), which also contained either proper nouns or common nouns.

- 2a. Subject cleft: It was Sam (the manager) that liked Tony (the clerk) before the argument began.
- b. Object cleft: It was Sam (the manager) that Tony (the clerk) liked before the argument began.

The difference in reading times for semantically matched and unmatched list/sentence pairs was numerically greater for object than for subject relatives.

Fedorenko, Gibson, and Rohde (2006) had participants recall either one or three common or proper nouns and self-space themselves through sentences with subject- or object-relative clauses (3) containing common nouns in a verification task.

- 3a. Subject relative: The physician who consulted the cardiologist checked the files.
- b. Object relative: The physician who the cardiologist consulted checked the files.

The interaction of list type and syntactic structure on reading times for the relative clause was significant: The effect of match (longer reading times when participants recalled matching, as compared with nonmatching, lists) was found only in the object-relative clauses in the three-item list condition.

A third study of external load has been taken to show that semantic properties of NPs in to-be-recalled lists interfere with retrieval of NPs in object-relative clauses when they have semantic features specified in the retrieval cue. Van Dyke and McElree (2006) reported self-paced reading times for object clefts with or without concurrent recall of sets of three words that could or could not be integrated into the sentence, as in (4):

- 4: Word list: table–sink–truck

Sentence: It was the boat that the guy who lived by the sea sailed/fixe*d* in two sunny days.

At the verb of the embedded clause, there were longer reading times in the integrated (*fixe*d**) than in the unintegrated (*sailed*) load conditions.

Similarity of sentence-internal items to one another The second set of semantic similarity effects involves sentence-internal noun phrases. Gordon et al. (2001) reported self-paced reading times in subject- and object-extracted relative clauses (5a, 5b) and clefts (6a, b) with common nouns (definite descriptions), proper nouns, or pronouns:

- 5a. The banker that praised the barber/Sue/you climbed the mountain...
- b. The banker that the barber/Sue/you praised climbed the mountain...
- 6a. It was the banker/Sue that praised the barber/Dee...
- b. It was the banker/Sue that the barber/Dee praised...

In the relative clauses (5a, 5b), the sentence type effect was reported as being greater in the definite description (*the barber*) condition than in either the pronoun (*you*) or the proper name (*Sue*) condition. In clefts (6a, 6b), there was a greater sentence type effect in sentences in which the NPs were matched for noun type (both definite descriptions or both proper nouns) than in sentences in which they differed. In a follow-up study, Gordon, Hendrick, and Johnson (2004) reported that there was a reduction in the object extraction cost in sentences with two definite descriptions when the relative-clause-internal NP was a quantified pronoun (*everyone*), but not when it was modified by an indefinite article (*a barber*) or was a generic noun (*barbers*) or a superordinate term (*the person*). Gordon and his colleagues argued that these results could be explained if the verb of an object-relative clause is a retrieval cue for two noun phrases and their order and the verb of a subject-relative clause is a retrieval cue for only one noun phrase. Gordon, Hendrick, and Johnson (2001) argued that “memory for order information is impaired when the items to be remembered are similar because the similarity of the items causes interference in retrieving the order information (Lewandowsky & Murdock, 1989; Murdock & Vom Saal, 1967; Nairne, 1990b) (p. 1420)”.

Syntactic interference effects

Lewis and his colleagues (Lewis, 1996, 2000; Lewis & Vasishth, 2005; Lewis et al., 2006; Van Dyke & Lewis, 2003; Vasishth & Lewis, 2006) have shown that syntactic features of lexical items that intervene between a retrieval cue and the item to be retrieved affect online processing. As in the work by Gordon and his colleagues that we have discussed, these results have been presented as evidence that principles that govern interference in ST-WM models also apply in sentence processing. Lewis and his colleagues have modeled a variety of phenomena in both comprehension and online performance in terms of retrieval-based syntactic interference, including effects of repetition of the same structure, some ambiguity effects, some embedding effects, some “locality” effects, and some “antilocality” effects (Vasishth & Lewis, 2006, present several results). We will describe the phenomenon by presenting one study that is often cited as an example of these effects (Van Dyke & Lewis, 2003).

Van Dyke and Lewis (2003) presented sentences such as (7)–(9). Words in parentheses were omitted in half the presentations of the sentences to produce ambiguities; we

have annotated some words with subscripts for ease of reference.

7. Short distance:
The secretary forgot (that) the student was₂ standing in the hallway.
8. Long distance, low interference:
The secretary forgot (that) the student who was₁ waiting for the exam was₂ standing in the hallway.
9. Long distance, high interference:
The secretary forgot (that) the student who knew (that) the exam was₁ important was₂ standing in the hallway.

Van Dyke and Lewis’s focus was on retrieval of *the student* as the subject of *was standing* when *was standing* was encountered.

Sentences (7) versus (8) and (9) vary the distance between *the student* and *was standing*. The critical sentences for syntactic interference effects are (8) versus (9), which vary the syntactic nature of the items between *was standing* and *the student*, holding distance constant. We focus here on the unambiguous versions, in which *the student* is assigned the role of subject when it is first encountered (the ambiguous versions will be discussed below). If *was standing* is a retrieval cue for a subject, there will be more syntactic interference with retrieval of *the student* in (9) than in (8) because *the exam* is a subject in (9) and an object in (8). Van Dyke and Lewis (2003) presented four studies of accuracy and self-paced reading times in an acceptability judgment task with these sentences. In the three of the four studies in which unambiguous versions of these sentences were presented, there was greater acceptance of the well-formedness of (8) than of (9) and longer self-paced reading times for the critical segment *was₂* in (9) than in (8). The difference between the unambiguous versions of sentences (8) and (9) is consistent with the claim that retrieval cues include syntactic features of items and with the model of interference described in the Retrieval-Based Parsing section. Similar effects of whether an NP intervening between a verb and its subject was in subject or object position were described by Van Dyke and McElree (2011) for reading times and for asymptotes of SAT functions (see below) in sentences where matrix verbs were the retrieval cues.

Retrieval dynamics

McElree and his colleagues have reported a series of studies using the SAT technique that led them to the conclusion that the dynamics of search through memory are similar for items in sentences and in lists. In this application of the SAT paradigm, participants are required to make judgments about the acceptability of

sentences. To study retrieval, the last word in the experimental sentence serves as a retrieval cue for an earlier word, and the position of the retrieved word and the number and type of potentially interfering items are varied across sentence types. Asymptotic d' is taken as a measure of the availability of the earlier word and d' dynamics as the measure of its accessibility.

McElree and his colleagues have used the SAT technique to study many types of sentences. A. E. Martin and McElree's (2008, 2009) work on verb phrase ellipsis such as (10), in which the antecedent of the elided verb phrase cannot be anticipated, produced characteristic results. A. E. Martin and McElree varied the distance between the ellipsis and the antecedent (10a, 10b), the length and complexity of the antecedent (10c, 10d), and the amount or proactive or retroactive interference (10e, 10f).

- 10a. The editor admired the author's writing but the critics did not.
- b. The editor admired the author's writing but everyone at the publishing house was shocked to hear that the critics did not.
- c. The history professor understood Roman mythology but the principal was displeased to learn that the over-worked students attending the summer school did not.
- d. The history professor understood Rome's swift and brutal destruction of Carthage but the principal knew that the over-worked students attending the summer school did not.
- e. Even though Claudia was not particularly angry, she filed a complaint. Ron did too.
- 10f. Claudia filed a complaint and she also wrote an angry letter. Ron did too.

None of these variables affected SAT dynamics, indicating that retrieval is unaffected by the amount of proactive or retroactive interference or the complexity of the retrieved item. This implies that retrieval uses a content-addressable mechanism, similar to findings in ST-WM reviewed above.

Capacity limits and temporal decay

The idea that parsing is capacity limited has been part of thinking about sentence comprehension for decades (Miller & Chomsky, 1963), and if it is correct that capacity limitations of the parser/interpreter are those found in the capacity-limited portion of ST-WM, this would tie a major feature of parsing and interpretation to a property of ST-WM. Lewis and his colleagues have proposed several such relations.

One is a limit on the number of items of the same type that can be maintained in a syntactic buffer. Working within the SOAR architecture (Newell, 1990),

Lewis (1996) developed a model in which lexical items can either be heads or dependents (or both) in structural relations. For instance, in the phrase *under the basket*, *under* can be the head of the relation *complement of a preposition* ([COMP P]), and *the basket* can be the dependent of the relation *complement of a preposition* ([COMP P]). When *under* is encountered, the structure [HEAD –COMP P–*under*] is placed in a buffer containing *head/dependent* sets; when *the basket* is encountered, the structure [DEPENDENT–COMP P–*the basket*] is created. Parsing consists of connecting heads and dependents under higher nodes. When both *under* and *the basket* are encountered, the two structures are combined into a single node [[HEAD–COMP P–*under*] [DEPENDENT–COMP P–*the basket*]].

Lewis (1996) argued that capacity limits on parsing arise because the buffer can contain only two constituents of the same type. For instance, in an object-relative clause such as (2b), repeated here for convenience,

- 2b. It was the manager that the clerk liked before the argument began,

when *the clerk* is encountered, the buffer contains the structure [DEPENDENT–SPEC–IP–*the manager; the clerk*], which is within the capacity of the buffer. However, in the doubly center-embedded structure (11),

11. It was the book that the editor who the secretary who quit married enjoyed, at *the secretary*

the buffer contains the structure [DEPENDENT–SPEC–IP–*the book; the editor; the secretary*], and the sentence becomes difficult because three NPs in the buffer are assigned the value [DEPENDENT–SPEC–IP].

A critical look at the evidence regarding features of retrieval-based parsing

As has been noted, results such as those presented above have been taken as evidence for links between ST-WM and the memory processes that apply in parsing and interpretation. The conclusion that has been suggested is that retrieval of information in parsing and interpretation makes use of a content-addressable retrieval process that derives from retrieval of information in ST-WM and that some parsing limitations are derived from capacity limitations of ST-WM. In this section, we critically consider these results and conclusions.

Empirical issues in results

We begin by noting several weaknesses in some of the empirical results reviewed above.

Empirical issues in studies demonstrating semantic interference effects

Several results cited above were not significant by conventional standards. In the Gordon et al. (2002) study (sentences 2a, 2b), the critical interaction did not approach significance, $F_1 = 2.4$, $p = .13$; $F_2 = 1.7$, $p = .19$. The interaction of description/proper-noun condition with structure in relative clauses in Gordon et al. (2001; sentences 5a, 5b) was not significant when the entire relative clause was considered, and differences in length of the clause-final NP in subject- and object-relative clauses could account for the differences when only clause-final words were compared (comparing *praised* with *the barber* would be expected to produce less of an effect than comparing *praised* with *Sue*). The interaction of noun type (matched, mismatched) with structure in clefts in Gordon et al. (2001; sentences 6a, 6b) was significant only in the analyses by participants. In Fedorenko et al. (2006; sentences 3a, 3b), all sentences contributed to online reading time analyses despite the fact that accuracy in answering the two questions that were asked about each sentence was only 55%.

In two studies, some results were unexpected. In Fedorenko et al. (2006), the difference between reading times for subject and object sentences in the segment that followed the relative clauses in the one-word list condition was larger in the mismatch than in the match conditions, the opposite of what retrieval interference would predict, and the object-relative cost was greater when participants recalled one-word than when they recalled three-word nonmatching lists, suggesting that reading times for the object-relative verb were reduced in the three-word nonmatching condition because of strategic choices made regarding performance on the list and sentence comprehension tasks. Van Dyke and McElree (2006) also reported an unusual effect of load (longer reading times in the no-load than in the load unintegrated condition), raising questions about how possible trade-offs between attending to the list and to the sentence might have influenced the results.³

Empirical issues in studies demonstrating syntactic interference effects

As has been noted, Van Dyke and Lewis (2003) found syntactic interference effects in unambiguous sentences. However, these effects were not affected by ambiguity; they were not greater in ambiguous than in unambiguous

³ In his review, Lewis suggested that the shorter reading times in the high-load condition in Van Dyke and McElree (2003) may have arisen because participants sped up to counteract decay of information in memory and that the critical finding is the interaction. However, there is no hint of a speedup prior to the critical point in Van Dyke and McElree's results.

versions of sentences (8) and (9). We believe that this is an unexpected finding, for the following reasons. In the ambiguous versions of (7)–(9), multiple factors induce *the student* to be attached as the object of *forgot* when it is encountered. When *was standing* is encountered, *the student* must be detached from *forgot* and attached as the subject of *was standing*. Van Dyke and Lewis said that this involves three steps: (1) retrieval of the feature of *forgot* that allows a sentential complement, (2) deactivation of the NP-complement frame, which makes *student* available to serve as the subject of *was₂*, and (3) retrieval of *the student* as the subject of *was standing*. There should be syntactic interference effects in the retrieval steps (1) and (3), for the following reasons. In step (1), *knew* in (9) shares allowing a sentential complement with *forgot*, while *waiting* in (8) does not. In step (3), *the exam* creates more interference with the retrieval of a subject in (9) than in (8), as is the case in the unambiguous versions of the sentences. Van Dyke and Lewis described these retrieval processes in the following terms:

when the disambiguating word *was_[2]* occurs, the parser must break the object link between *forgot* and *student* in order to attach *was_[2]* into the existing parse tree. In addition, the NP that was the object of *forgot* must now be reattached as the subject of the IP headed by *was_[2]* and the entire new clause must be attached as the complement of the alternative sentential complement-taking sense of *forgot* . . . two retrievals are required: one for the alternative sense of *forgot*, the assigner of the sentential complement, and one for the subject itself. (p. 287)

The prediction is thus that the syntactic interference effect (longer reading times for *was₂* in [9] than in [8]) will be greater in the ambiguous than in the unambiguous versions of these sentences. However, this was not the case. Van Dyke and Lewis (2003) concluded that the syntactic interference affects “the process that makes the correct attachment and not those functions related to repairing the incorrect parse” (p. 299). However, since the functions related to repair involve a second retrieval operation, we find these results inconsistent with the model.

Failure to find expected effects of potentially interfering items on retrieval are not restricted to resolution of ambiguities and are not simply null findings (see Phillips, Wagers, & Lau, 2011, for a review). They suggest that some items are not contacted during the retrieval process and, therefore, have implications for retrieval. We return to this issue below.

Empirical issues in studies using the SAT technique

A general issue pertaining to the SAT results is that both asymptotic d' and dynamics reflect both retrieval and the

construction of a syntactic and semantic interpretation, as well as a decision-making process. Equivalent dynamics in different sentence types do not, therefore, necessarily imply equivalent access in retrieval; different combinations of speed of retrieval, computation, and decision making could produce equivalent dynamics. In addition, participants see the same sentence structures many times, possibly promoting the use of particular retrieval strategies, despite the careful selection of stimuli. We will illustrate this second problem with an example.

McElree, Foraker, and Dyer (2003, Experiment 2) studied sentences with relative clauses, shown in (12). (Unacceptable versions of the sentences are indicated by parenthesis and *; the relative clause containing potentially interfering NPs is enclosed in square brackets [...].)

- 12a. The book ripped (/laughed).
 b. The book [that the editor admired (/amused)] ripped (/laughed).
 c. The book [from the prestigious press that the editor admired (/amused)] ripped (/laughed).
 d. The book [that the editor who quit the journal admired (/amused)] ripped (/laughed).
 e. The book [that the editor who the receptionist married admired] ripped (/laughed).

The d' dynamics were the same for sentences (12b), (12c), and (12d), which McElree et al. (2003) took as evidence for a content-addressable retrieval process triggered by a retrieval cue formed by the sentence-final verb. However, there is evidence that participants made judgments about the acceptability of the relative clause incrementally: Correct rejection rates for sentences that became unacceptable at the verb of the relative clause (*amused*) were at asymptote 50 ms after the end of the sentence (percent correct = 72.3, 76.4, 78.1, 75.3, 75.6, and 80.1 at 50, 300, 500, 800, 1,200, and 3,000 ms) but rose to asymptote over time for sentences that became unacceptable at the verb of the main clause (the sentence-final verb, the retrieval cue–*laughed*; percent correct = 42.3, 44.5, 63.4, 79.4, 85.1, and 87.7 at 50, 300, 500, 800, 1,200, and 3,000 ms). This indicates that participants made an acceptability judgment about the relative clause when they read the relative clause verb, which would have allowed them to consider only the sentence-initial noun for its plausibility as the subject of the sentence-final verb. In that case, the only item in the memory set at the point of reading the sentence-final verb would have been the sentence-initial NP in (12b), (12c), and (12d), making the d' dynamics equivalent in those sentences. The repetition of sentences may have been a factor that led to participants making judgments at the relative clause verb.

Other results of this study pertain to the same issue. Rise rates were faster for (12a) and slower for (12e) than for the three other sentences. McElree et al. (2003) attributed the

results in (12a) to the immediacy of the NP to the sentence-final verb that served as the retrieval cue, consistent with the results showing that the list-final item, which immediately precedes the retrieval cue, is in a privileged position in cued recall studies of lists. McElree et al. attributed the slowed dynamics in (12e) to computational demands that arise after retrieval has occurred. However, only (12e) did not have a control sentence that became unacceptable at the relative clause verb, so the difference in dynamics for (12e) and the others could have been due to participants considering all NPs in (12e) and not in (12b), (12c), and (12d). Thus, the data from these sentences are consistent with a search process and effects of the task on sentence processing.

Interpretation of results

Despite issues such as those raised above, many results in the studies reviewed in the Mechanisms in ST-WM and Parsing and Interpretation: Empirical Results and Interpretation section and others in the literature demonstrate semantic and syntactic interference effects in online reading times, eye fixations, and SAT asymptotes and are consistent with content-addressable retrieval in many situations. As was noted, these results have generally been taken as providing consistent, converging evidence about the nature of memory processes in parsing and interpretation. There are, however, a number of areas in which there are alternate interpretations of results and others in which the results are hard to reconcile with one another and with other results in the literature.

Number of memory systems needed to account for effects of external load

Fedorenko et al. (2006) argued that an interaction of the semantic similarity of words in recall lists and words in sentences and sentence structure in online processing shows that both sets of words are maintained in a single memory system. However, Henkel and Franklin (1998) found that source memory confusions increased both when items in two contexts were conceptually similar and as the number of similar items in the nonsource list increased. Thus, the effect of match in the high-load condition only in Fedorenko et al. is consistent with separate memory systems for lists and parsing-interpretation, coupled with source memory confusions.

The nature of retrieval

Semantic similarity effects and retrieval The view that the semantic effects in the articles by Gordon, Fedorenko, and their collaborators are due to interference with retrieval of

noun phrases required by the verb of an object-relative clause raises many issues.

First, the semantic similarity effects in these studies cannot be due to the interference-generating mechanism outlined in the Retrieval-Based Parsing section, which maintains that only features that are specified in the retrieval cue affect retrieval. The verb of an object-relative clause may generate a cue to retrieve two noun phrases (and possibly their order), but it cannot generate a cue to retrieve two noun phrases that share a common semantic property or an aspect of form such as being a common or proper noun, because these features of the retrieved items are not present in many sentences. Thus, semantic features of the NPs that are in the memory set are not part of the retrieval cue and cannot give rise to interference by the mechanism outlined above.

In the passage quoted above, Gordon et al. (2001) connected semantic similarity effects to the finding in the ST-WM literature that similarity of items reduces recall of item and order information and did not claim that they were generated by the interference mechanism described in the Retrieval-Based Parsing section. However, suggesting that the semantic effects in these studies are due to the mechanisms that produce similarity effects in the recall of item and order information raises other questions.

One is whether the parser/interpreter makes use of order information; in some models (e.g., Lewis et al., 2006), it does not. Assuming that the parser does use order information, deriving the retrieval mechanism that produced similarity effects in item and order recall in Gordon and his colleagues' work from mechanisms in ST-WM poses a problem. Two basic mechanisms that could underlie retrieval of order information have been proposed: item-to-item associations (*chaining*) (Lewandowsky & Murdock, 1989; Wickelgren, 1965, 1966) and an *address* mechanism (Nairne, 1990a). Address models postulate that items in the recall list are retrieved on the basis of their properties, not on the basis of interitem associations, and negatively affect retrieval only if their features are part of the retrieval cue. Thus, given that semantic features were not part of the retrieval cue in the Gordon and Fedorenko studies, tying their results to models of ST-WM requires that the mechanism for retrieval of order information be some version of chaining. However, chaining models do not appear to be viable. Chaining fails to account for the *saw-tooth effect* in which similar items are less well remembered in lists of alternating phonologically similar and dissimilar items (Baddeley, 1968, Experiments 5 and 6), the fact that *first-in-report* errors for dissimilar items (the proportion of errors when responses to all previous items are correct) do not differ in mixed and pure lists (Henson, Page, Norris, & Baddeley, 1996, Experiments 2 and 3), and the fact that relative errors (an erroneous response in the correct position

relative to a preceding [necessarily erroneous] response) do not differ between mixed and pure lists (Henson et al., 1996, Experiment 1; see Henson, 1998, for a discussion).

Another concern about attributing semantic interference effects to retrieval is that, if features of items in the memory set that are not part of a retrieval cue interfere with retrieval, other shared properties of items in a sentence that are not part of retrieval cues should affect retrieval. Given the importance of phonological properties in immediate serial recall, which are believed to arise during retrieval, a test case would be phonological similarity of words. The evidence suggests that phonological similarity of words does not affect retrieval at points of syntactic disambiguation (Kennison, 2004) and by the verb of object-relative clauses (Obata, Lewis, Epstein, Bartek, & Boland, 2010; but see Acheson & Macdonald, 2011).

The comparison of the semantic similarity and SAT results also leads to a problem. Similarity of list items does not affect, and may slightly improve, retrieval of item information (Baddeley & Hitch, 1974; Nairne, 1990b; Nimmo & Roodenrys, 2004; Poirier & Saint-Aubin, 1996; Postman & Keppel, 1977, Experiment 2; Underwood & Ekstrand, 1967; Watkins, Watkins, & Crowder, 1974; Wickelgren, 1965, 1966). Therefore, if semantic similarity effects arise at retrieval, what must be retrieved is item and order information, as Gordon et al. (2001) suggested. However, as was discussed in our review of list memory studies, the content-addressable retrieval suggested by the SAT dynamics results is consistent only with retrieval of item information in ST-WM, not order information (Hockley, 1984, Experiment 3; McElree & Doshier, 1993). Since both the interference and SAT studies involve retrieval of NPs in the same structures (relative clauses), they must be characterizing the same retrieval process. If both sets of results arise during retrieval, the difference between the two studies reflects strategic effects of task or other factors on retrieval mechanisms.

An alternative, suggested by Lewis et al. (2006), is that the semantic similarity effects in the studies by Gordon and his colleagues arose during encoding. One test of an effect of a factor on encoding is whether the effect is increased by a concurrent task (cf. Baddeley, 1968, Experiment 4). Data from some of the studies reviewed above are consistent with this locus of semantic interference. For instance, in Gordon, Hendrick, and Levine (2002), there was an effect of the semantic match of the words in the lists and the sentences in reading times for the sentence-initial segment (but not for the sentence-final segment).

Two points arise if the locus of semantic interference in these studies is at encoding. One is that the results do not bear on retrieval mechanisms. The second is that encoding of words into memory during sentence processing appears to differ from encoding of words in ST-WM, where similarity effects appear to arise during retrieval. Baddeley (1968,

Experiment 4) found no effect of presentation in noise on the phonological similarity effect and concluded that the effect did not arise during encoding into ST-WM. Baddeley (1968, Experiments 1–3) also found that the recall of phonologically similar items declined less rapidly than that of phonologically dissimilar items in the Brown–Peterson paradigm, indicating that the phonological similarity effect does not arise during storage. In most models of ST-WM, including feature models in which similarity effects could arise at encoding due to overwriting of features of earlier items (e.g., Neath, 2000), the effect of phonological similarity arises at recall, not encoding. Assuming the principles underlying the effects of phonological and other forms of similarity are the same (which is reasonable in the case of the feature model), if the semantic similarity effects in the studies by Gordon, Fedorenko, and their collaborators arise during encoding, this suggests a difference from the mechanisms found in ST-WM.

Syntactic interference effects and retrieval cues Syntactic interference effects have been taken as evidence for content-addressable retrieval of items in sentences on the basis of features of retrieval cues created by incoming lexical items and features of their context. However, there is a developing literature that documents the absence of expected effects of potentially interfering items in situations where such effects would be expected if this retrieval mechanism applies.

One phenomenon that has been emphasized is illusions of grammaticality. An example is shown in (13), where the verb agrees with a plural NP that intervenes between it and a singular subject, and many readers initially find the sentence grammatical (Eberhard, Cutting, & Bock, 2005), suggesting that the intervening NP is sometimes retrieved by the agreement features of the verb.

*13. The key to the cabinets are on the table.

The critical finding is that illusions of grammaticality are selective. A context that contrasts with subject–verb agreement is a reflexive agreeing with a plural NP that intervenes between it and a singular subject, as in (14), where readers readily report the ungrammaticality (Dillon, Mishler, Sloggett, & Phillips, 2012; cited in Phillips et al., 2011). This argues that the intervening plural NP is not retrieved in (14).

*14. The diva that accompanied the harpists on stage presented themselves with lots of fanfare.

Phillips et al. (2011) suggested that the difference between (13) and (14) is due to the nature of the retrieval cues established by the agreement markers on verbs and reflexives. They proposed that reflexives retrieve their antecedents using only structural cues, while verbs create retrieval cues

that specify the features of the subject. In their view, this difference is due to the fact that a sentence subject reliably predicts a verb and its agreement features, whereas a reflexive cannot be reliably predicted.

Items in certain syntactic positions appear not to be contacted during the retrieval process even when the retrieval cue specifies their features. For instance, NPs in certain positions do not appear to be contacted when a verb of an object-relative clause retrieves its object. Traxler and Pickering (1996) found evidence for a retrieval cue that specified semantic features of the object of the verb of a relative clause (similar to Van Dyke & McElree, 2006, discussed above) in (15) in the form of longer reading times for *shot* in the anomalous version of that sentence (with the NP *garage*), but no prolongation of reading times for *wrote* in the version of (16) with the NP *city*:

15. That's the pistol/garage with which the heartless killer shot the hapless man.
16. We liked the book/city that the author who wrote unceasingly saw while waiting for a contract.

This suggests that *the city/book* are not in the list of items contacted by the retrieval cue established by *wrote* in (16). It has been suggested that this is because of the syntactic environment in which *the city/book* and *wrote* occur—that *the book/city* is outside a syntactic “island” that contains *wrote*.

One common feature of sentences that might serve to restrict the items contacted by a retrieval cue is placing one or more items in linguistic focus. Further consideration of the results in McElree et al. (2003) suggests that this may be the case. Recall that McElree et al. found that SAT dynamics differed for (12e) and (12b–d), repeated here:

- 12b. 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.

In other studies, structures comparable to (12e) did not show different dynamics from others tested. There were no differences in *d'* dynamics in (17a–17c) (McElree, 2000) or in (18a–18c) (McElree et al., 2003):

- 17a. It was the book that the editor enjoyed.
- b. It was the book that the editor who the secretary married enjoyed.
- c. It was the book that the editor who the secretary who quit married enjoyed.

- 18a. It was the book that the editor enjoyed.
 b. It was the book that the secretary believed the editor enjoyed.
 c. It was the book that the secretary believed that the journalist reported that the editor enjoyed.

The clefted noun phrase in (17) and (18) is in linguistic focus, which could exclude the other nouns in those sentences from being contacted by the retrieval cue.

The selectivity of syntactic interference effects and grammaticality illusions point to highly detailed features of retrieval cues. Even with highly detailed cues, formulating retrieval cues in such a way that items in “invisible” environments are not contacted by a content-addressable retrieval mechanism has proven challenging. Alcocer & Phillips ([unpublished manuscript](#)) described several ways in which retrieval cues could include complex relations of syntactic nodes, such as those needed to account for the possible antecedents of reflexives (“c-command”; Reinhart, 1976). Although it is possible to formulate such cues, they suggested that the memory demands of maintaining the types of annotated syntactic structures that such cues require are too great to be realistic. They advocate hybrid retrieval processes that include both content-addressable retrieval and search mechanisms:

We might conclude that it is attractive to invoke a serial mechanism for accessing memory that explicitly uses the graph structure of a tree to access c-commanding nodes. Such mechanisms capture relational notions such as c-command with far greater ease than do mechanisms that implement parallel access. (ms. pp. 35–36)

On this view, there are two types of retrieval operations in parsing and interpretation: content-addressable retrieval and search. Within the set of search mechanisms, searches that use different retrieval cues would be expected to search a structure held in memory in different ways. Phillips et al. (2011) summarized the picture in the following terms:

We find many situations where on-line language processing is highly sensitive to detailed grammatical constraints and immune to interference from inappropriate material. . . . Yet in many other situations we find that on-line processes are susceptible to interference and to grammatical illusions. In order to explain this ‘selective fallibility’ profile we have argued that speakers build richly structured representations as they process a sentence, but that they have different ways of navigating these representations to form linguistic dependencies. The representations can be navigated using either structural information or using structure-insensitive retrieval cues. (p. 173)

The role of capacity and temporal limits in parsing and interpretation and their relation to ST-WM

As has been discussed, Lewis (1996) suggested that a capacity limit derived from ST-WM applies in parsing and interpretation. His proposal raises a variety of issues.

One is the number of items of the same type that the buffer can contain. Lewis (1996) set the limit of two on the number, on the grounds that two is the minimum number of items that must be connected in a relation, and it would be of interest to explore the possibility that it is also the maximum number that can be connected in a relation. However, this number does not correspond to any suggestion about the size of a fixed-capacity, capacity-limited portion of ST-WM, which, as we have seen, has been estimated as one by McElree and four to five by Cowan. It would be possible to set the number of items that can be maintained in a flexible-capacity ST-WM store such as that proposed by Verhaeghen to two, but to do so only to relate the capacity limit of a parsing store to one in ST-WM is obviously circular.

In other work, Lewis and his colleagues have proposed a second link between capacity limits in ST-WM and parsing and interpretation—namely, that the structure of their model, with three buffers (a control buffer, a problem state buffer, and a retrieval buffer), “has much in common with conceptions of working memory and short-term memory that posit an extremely limited focus of attention of one to three items, with retrieval processes required to bring items into focus for processing” (Lewis & Vasishth, 2005, p. 380). However, the existence of three buffers differs from the ST-WM concept of a single CS and triples the capacity limit of the CS.

Both these suggestions regarding ways that putative capacity limits on parsing and interpretation might be derived from capacity limits in some part of ST-WM raise broader issues.

One pertains to the nature of items in stores utilized in parsing and interpretation. In Lewis’s model, what is stored in the control and problem state buffers are phrase markers that are formed as the sentence is processed. As was noted above, these representations are created by a left-corner parser that expands the phrase marker associated with each new word according to stored procedures (rules). The contents of these phrase markers must remain accessible. In Fig. 1, the word *the* triggers the creation of an NP, a VP, and a dominating S node. If the next word is the adjective *big*, as in the phrase *the big dog*, it must be attached within the NP that is formed after the word *the* is encountered. Thus, in the terminology used to characterize the CS, the items in these buffers are *chunks* (Cowan, 2000; Halford, Cowan, & Andrews, 2007; Miller, 1956; Verhaeghen et al., 2004) that are not “opaque” (Halford et al., 2007).

The accessibility of the elements in a chunk in Lewis's model differs sharply from the conceptualization of chunks in the ST-WM literature. Halford et al. (2007) made an analogy between chunks and the representation of speed as a point on a line, as in a car dashboard display; the constituents of the concept of speed (distance/time) cannot be accessed from such a representation. The inability to access the contents of a chunk increases memory capacity (in some sense) at the expense of computational capacity. The distance that would be traveled in a given time cannot be calculated from the position of a needle on a car dashboard, but requires access to the constituent concepts. In Lewis's model, storage demands of memory are reduced because what is stored is a chunk, but the computational capacity of the system is not affected because the contents of the chunk are available. These are not the properties that have been attributed to a capacity-limited CS.

In addition, the transparency of chunks in Lewis's model differs from applications of models of capacity limits of the CS to other cognitive phenomena. For instance, Halford et al. (2007) argued that the limits of the CS restrict human reasoning to problems with relational complexity ("arity") of no more than four, but their model requires that the reasoning process cannot access items within a chunk. If the capacity limits in processing in various domains are to be explained because of capacity limits in the CS, the characterization of how chunks are processed must apply universally.⁴

A second issue is that capacity limits and content-addressable retrieval apply to different stores in models of ST-WM—capacity limits, if they exist, to the CS or an expanded focus of attention, and content-addressable retrieval to items outside the capacity-limited portion of ST-WM (i.e., to items in LTM). If capacity limits exist in parsing and interpretation, this would require that they are due to items being stored in a store from which they are not retrieved by a content-addressable mechanism, a noticeable discrepancy across the models we have been discussing.

A more general issue that arises regarding capacity limitations is whether there are any such limitations in ST-WM. As was reviewed above, in their work on memory, McElree and his colleagues have argued that there are no items in memory that are in an activated but unattended state and that the number of items that can be maintained in focal attention in memory is limited to one. This model does not contain a

capacity-limited store. Consistent with this model, McElree and his colleagues have argued that the SAT dynamics for retrieval are the same in sentences as in LTM, which is not capacity limited. As was noted, if this model of retrieval is correct, there are no storage capacity limits on parsing and interpretation. McElree (personal communication) has suggested that phenomena sometimes attributed to capacity limits are due to interference.

A second feature of Lewis's model is its reliance on a second limitation of ST-WM—temporal decay. Both locality and antilocality effects are attributed to how representations decay and are reactivated by retrieval processes (Vasishth & Lewis, 2006). However, evidence for temporal decay in ST-WM is controversial (Lewandowsky & Oberauer, 2008).

In summary, the effort to attribute parsing complexity to capacity (or temporal) limitations of ST-WM encounters a variety of challenges, some related to the details of the proposed limitations in parsing and interpretation and those that have been proposed in some models of ST-WM, some related to the nature of the items maintained in memory in parsing and interpretation, and some related to inconsistencies with proposals about the relation of retrieval dynamics in parsing and interpretation to various ST-WM and in stores.

Overlap and nonoverlap of retrieval mechanisms in ST-WM and in retrieval-based parsing

The lines of research reviewed in the Mechanisms in ST-WM and Parsing and Interpretation: Empirical Results and Interpretation section relate many mechanisms derived from the study of ST-WM to memory in parsing and interpretation: capacity limitations, content-addressable retrieval, retrieval of item and order information, temporal decay of representations, and so forth. In fact, the set of mechanisms attributed to ST-WM that have been appealed to in this literature is greater than our review has indicated. For instance, Johnson, Lowder, and Gordon (2011) suggested that encoding of item versus order information in sentences was affected by the nature of noun phrases in sentences and related their results to the attention-based mechanism postulated in the "the item-order theory of list composition effects" (McDaniel & Bugg, 2008). The invocation of a large number of mechanisms derived from models of ST-WM in models of memory utilized by parsing and interpretation leads to two questions: Are all the mechanisms that have been suggested to apply in parsing/interpretation reasonably well established in the ST-WM literature? Are there mechanisms that are reasonably well established in the ST-WM literature that are *not* used in parsing and interpretation?

The answer to the first question is almost certainly "no." We noted above that chaining models, needed to account for

⁴ In his review, Lewis pointed out that the transparency of chunks in his parsing model is only at the level of the features of the chunk; for example, an NP chunk might contain a feature for number that is accessible, but the structure of its determiner phrase in specifier position is not accessible without further retrievals. As we understand models such as Halford's, however, they would not allow the internal structure of an NP chunk to be accessible at all, not even by later retrievals.

semantic interference effects if they arise during retrieval, and decay in ST-WM have little support as features of ST-WM. The answer to the second question seems to be “yes.” Many mechanisms postulated to support memory functions in ST-WM, such as Weber-compressed temporal intervals from retrieved item to retrieval cue or response (Brown, Neath, & Chater, 2007), positional activation gradients measured from both list-initial and list-terminal items (Henson, 1998), temporal oscillators that estimate list length (Brown, Preece, & Hulme, 2000; Henson & Burgess, 1997), and others (for a review, see Henson, 1998), are not plausible candidates for mechanisms that support the memory requirements of parsing and interpretation.

This has consequences for the relation of memory in parsing and interpretation to ST-WM. If mechanisms such as temporal decay or chaining are found in parsing and interpretation and not in ST-WM, they are either specific to parsing and interpretation or derived from memory systems other than ST-WM. If there are mechanisms found in ST-WM and not in parsing and interpretation, parsing and interpretation, at most, select from memory mechanisms found in ST-WM.

The mechanisms listed above that are postulated to apply in ST-WM and that are not plausible candidates for mechanisms underlying memory in parsing and interpretation have been postulated in models of recall, not recognition. One might therefore consider whether parsing and interpretation use only mechanisms that support recognition. This has not been the assumption of researchers in this field, who have invoked mechanisms that apply in either recognition or recall. Uniform SAT dynamics have been found in recognition; they cannot be examined in recall. As has been noted, effects of phonological and semantic similarity occur in immediate serial recall Baddeley (1966), but not item recognition (Wickelgren, 1965). However, it might be possible to restrict consideration of mechanisms to those found in only one task.

This raises the question of which STM task, if any, has features that are comparable to those found in the memory processes that support parsing and interpretation. Neither recognition nor recall of items in lists is identical to the retrieval process that occurs in sentences. In recognition of items in lists, the retrieval cue is given to the participant, whereas it must be created in sentences. In this respect, retrieval in sentences resembles recall of items in lists. This raises questions about the relevance of the SAT results. However, retrieval in sentences also differs from recall of items in lists, because there are no features of lists that generate and constrain retrieval cues in recall in the same way that context and specific items generate and constrain retrieval cues in sentences. Retrieval in parsing and interpretation appears to be closest to what Humphreys, Bain, and Pike (1989) call “partial cued recall with a list associate

cue”—recall of one member of an associate pair, given the other—insofar as the cue is part of the study list, as is the case for items that generate retrieval cues and the items for which they search in sentences. However, the two tasks differ inasmuch as the cue and retrieved items are not associated in the sentence case.

These considerations show that retrieval in sentence comprehension is not easily related to retrieval in any laboratory STM task. Along with the other issues raised above, this invites us to consider whether ST-WM is the right memory system to be relating to the memory that supports parsing and interpretation. We conclude this article by considering an alternative.

An alternative framework for viewing the memory system for parsing and interpretation

A central feature of the memory system that supports parsing and interpretation that emerges from our discussion of the recent literature is the domain specificity of many aspects of memory in parsing and interpretation. We have argued that retrieval cues, aspects of retrieval processes, and any capacity limits that may apply in parsing and interpretation all have important domain-specific features. Domain-specific aspects of memory are characteristic features of cognitive skills that have been related to a different postulated memory system—*long-term working memory* (LT-WM). In this section, we develop the idea that the memory system that supports parsing and interpretation is LT-WM.

The concept of LT-WM was proposed and developed by Ericsson and Kintsch (1995) as a response to two phenomena: the greatly expanded working memory capacity of experts and skilled performers and the fact that skilled activities can be interrupted and later resumed without major effects on performance. Examples of activities supported by LT-WM in individuals with established expertise include using an abacus, mental calculation, waiters’ and waitresses’ memory for orders, and chess players’ memory for chessboard arrays, and examples of activities supported by LT-WM in individuals after training include expansion of span (to as much as 80 items). Considerable resistance of these memory performances to interruption has been documented in studies such as Chase and Ericsson (1982), which found only a small decrement in recall of 30-item lists due to proactive interference from other lists in a trained participant. Ericsson and Kintsch argued that several capacities are required for these memory performances. The individuals who show these memory skills must have a large body of relevant knowledge, which allows them to store the incoming information in LTM; the activity must be familiar; and the encoded information must be associated with a “retrieval structure”—a domain-specific complex set of retrieval cues.

Given the emphasis on retrieval in the literature we have been concerned with, the retrieval cues that support skilled memory are of particular interest. A well-studied example is the expansion of digit span. Chase and Ericsson (1981; cited in Ericsson & Kintsch, 1995) found that an individual with “an exceptional digit span” used “a hierarchical retrieval organization . . . to memorize lists of 30 digits. . . . At the lowest level, SF used mnemonic association to running times (3596 → 3 min and 59.6 s, or just below a 4-min time for a mile) and other numerical relations to group digits and encode them as units. SF then used spatial relations to encode digit groups into supergroups. At the time of recall, SF could easily regenerate any of the unique locations of the retrieval structure and use a given location as a cue to retrieve the corresponding digit group” (p. 216). As is illustrated by this example, the retrieval cues that support such performances are domain specific.

Ericsson and Kintsch (1995) applied this framework to discourse comprehension. They drew a distinction between “thoughts—the cognitive states—[which] are themselves the end products of complex generation processes,” and “intermediate results,” the “complex, not necessarily linear, sequence of processes involved in the generation of a cognitive state, including the necessary memory buffers for the temporary storage of intermediate processing results” (pp. 220–221). They argued in considerable detail that the “retrieval of cognitive end products” is not significantly affected by interruptions, favoring the LT-WM account.

It is not clear from Ericsson and Kintsch’s (1995) discussion whether they think LT-WM does not, or cannot, apply to sentence processing or whether they have simply focused on higher levels of language structure. As far as we can see, there is nothing in Ericsson and Kintsch’s model that limits the applicability of LT-WM to the relation of end products (“states”) to one another. We suggest that it applies to the generation of intermediate products of computation as well as to relations of “states.” The phenomena that led to the concept of LT-WM are seen in parsing and interpretation: a large capacity for memory of items, domain-specific knowledge, and a high degree of familiarity on the part of performers. The imperviousness of parsing and interpretation to temporal interruption is unknown; the few available results (e.g., Wanner & Maratsos, 1978) suggest that it is unexpectedly robust. Applying Ericsson and Kintsch’s model to parsing and interpretation would lead to a model in which input activates items in long-term memory; these items are related by knowledge-based associations, patterns, and schemas to establish an integrated memory representation, and what is maintained in a short-term memory system are retrieval cues for these items.

It is of interest to reexamine Lewis and his colleague’s model with this perspective in mind. The representation of syntactic knowledge in declarative memory in Lewis and his

colleague’s model corresponds to the “knowledge-based associations, patterns, and schemas” that are the basis for establishing an “integrated memory representation” in Ericsson and Kintsch’s (1995) model. Both Lewis and Ericsson and Kintsch maintain that the application of “knowledge-based associations, patterns, and schemas” to the creation of an “integrated memory representation” is a domain-specific process. As was noted above, the operations of Lewis’s model involve procedural memory and declarative LTM; ST-WM is invoked—questionably, we have argued—only to account for capacity limits on parsing. Van Dyke and Lewis (2003) said that “cues in the retrieval context are combined multiplicatively to produce a single retrieval probe . . . all cues are combined into a retrieval probe, which gives the strength of the relationship between each possible probe cue ($Q_1, . . . , Q_m$) and the memory trace.” This process has many similarities to the formation of cues in LT-WM as described by Ericsson and Kintsch.

The suggestion that parsing and interpretation rely on LT-WM has much in common with other suggestions in the literature. On the basis of neuropsychological and imaging data, Ullman (2004) suggested that parsing utilized procedural memory, while lexical access involved accessing declarative memory. Gordon et al. (2002) suggested that Ericsson and Kintsch’s (1995) characterization of working memory might apply to sentence processing. Johnson et al. (2011) noted the compatibility of LT-WM with cue-based parsing models. Misyak, Christiansen, and Tomblin (2010) argued for a role of implicit memory in comprehension (“implicit expectations of upcoming linguistic input and . . . the on-line engagement of such predictive skills comprises an important role in language acquisition and processing” [p. 139]). Fodor (1983) quoted Merrill Garrett as saying that parsing is a reflex. These formulations and others emphasize the idea that parsing and interpretation, or at least much of these processes, is a skill. The implication we draw is that the memory system that supports parsing and interpretation is the memory system that supports skilled cognitive performances—LT-WM.

Placing the memory system that supports parsing and interpretation within the LT-WM/skilled-performance framework captures the facts that much of encoding, storage, and retrieval of information in parsing and interpretation is unconscious and subjectively undemanding and that the retrieval cues used in parsing and interpretation are domain specific. The domain specificity of retrieval cues and aspects of the retrieval process contributes toward the degree of independence of skills in individual domains from skills in others (great waiters are not particularly likely to be great chess players). This predicts that the skill of parsing and interpretation will not necessarily covary with other skills and accounts for the findings that measures of memory use in parsing and interpretation—such as online processing

times at points at which parsing and interpretation require retrieval from memory—correlate poorly or not at all with memory performance on ST-WM tasks (Caplan & Waters, 1999). On the other hand, overlap of components of skills, or their utilization of a common functional architecture for memory (see below), might lead to correlations between measures of online effects in parsing and interpretation and online implicit memory effects in selected implicit memory tasks, as has been shown for statistical learning tasks (Conway, Karpicke, & Pisoni, 2007; Conway & Pisoni, 2008; Misyak et al., 2010), and to correlations of measures of online effects in parsing and interpretation with elementary speed of processing (Caplan et al., 2011), which is common to many skills.

The major limitation that we see in testing the hypothesis that the memory system that supports parsing and interpretation is a form of procedural memory is the lack of detailed information about how processes such as encoding, maintenance, and retrieval operate online in skilled tasks that use some form of procedural memory. In the area closest to syntactic comprehension—implicit artificial grammar learning—there is little discussion of this topic. For instance, Pothos (2007) reviewed nine theoretical accounts of artificial grammar learning with respect to the types of structures that were learned but did not discuss the issue of how people who acquired an artificial grammar retrieved information about earlier items in the input string to recognize that a string was well-formed online. Researchers who have proposed that the memory system that supports parsing and interpretation is a form of procedural memory, such as Ullman (2004), have argued that the occurrence of deficits in parsing and interpretation in patients after lesions in areas known to affect skilled performance and the appearance of BOLD signal in parsing and interpretation tasks in these areas supports this view. However, neuropsychological and imaging data are both correlational findings. A rigorous test of the hypothesis that the memory system that supports parsing and interpretation is a form of procedural memory will require a characterization of online mechanisms that support encoding, storage, and retrieval in procedural memory that can be compared with memory utilization in online parsing and interpretation.

Roles of LT-WM and ST-WM in different aspects of parsing and interpretation

We have suggested that LT-WM supports skilled parsing and interpretation. However, skilled parsing and interpretation occasionally break down, leading to incremental comprehension failure. We have previously suggested that ST-WM plays a role in supporting memory demands of processes that occur at points of incremental comprehension

failure (Caplan & Waters, 1999). We here elaborate on this suggestion.

In Caplan and Waters (1999), we suggested that ST-WM supported retrieval of information when syntactic structures had to be revised. This suggests that one could identify points in sentences at which ST-WM supports retrieval on the basis of syntactic features and parsing considerations. The literature provides many hypotheses about structural and/or processing determinants of points of possible incremental parsing failure, such as very high surprisal values (Hale, 2001; Levy, 2008; Staub 2010), high integration costs (Gibson, 1998), the need for nonmonotonic parsing during revisions (Sturt & Crocker, 1996), memory capacity limits as discussed above, and others. However, we would now argue that no purely structural or processing theory can determine the points at which ST-WM may support retrieval, because these points are determined by incremental *comprehension* failure, not incremental *parsing and interpretation* failure. An important determinant of the success of sentence comprehension is intentional context—the comprehender’s task and the criterion he or she sets for accomplishing the task. Tasks that require superficial comprehension, or low criteria for task performance, can allow comprehension to succeed even if the parser/interpreter fails to generate a grammatically licensed meaning (e.g., if all that the comprehender needs to know is who participated in an action, and not what thematic roles different participants played), and, conversely, comprehenders may not have achieved adequate comprehension when the parser/interpreter generates a well-formed, semantically coherent representation if criteria for successful comprehension are set high (as seen in “wrap up” effects at clause and sentence boundaries). This view is closely related to Lewis’s idea of bounded optimality (Howes, Lewis, & Vera, 2009) and Ferreira’s views on “good enough parsing” (Ferreira, Engelhardt, & Jones, 2009; Ferreira & Patson, 2007). Models of structural and processing load can provide an upper bound to what can be comprehended by most individuals, but this only partially determines incremental comprehension failure.

To investigate the possible role of ST-WM in supporting processing that arises after incremental comprehension failure, it is therefore necessary to identify empirical markers of incremental comprehension success and failure. Since comprehenders have great freedom with respect to their behaviors, it is not clear that there are any behaviors that are inevitably triggered by either of these two states. However, the converse may be the case: Certain behaviors, if present, may reliably indicate incremental comprehension failure. We will suggest two such behaviors.

In whole-sentence reading, we suggest that regressive eye movements mark points of incremental comprehension failure (Reichle et al., 2009). Not all regressive eye movements are likely to involve use of verbal ST-WM, however. Within-word regressions do not require memory. Of

regressions to earlier words, some, such as those to words at line edges (Mitchell, Shen, Green, & Hodgson, 2008), those that move backward word by word (Mitchell et al., 2008), and those at the beginning of a sentence (Frazier & Rayner, 1982; von der Malsburg & Vasishth, 2011), are directed toward perceptually salient features of the print array and are most likely controlled by spatially directed attention. The regressive eye movements that are controlled to a significant extent by memory for the words in the sentence are regressions that are “targeted” to within one word of an item that is relevant to the clarification of the structure and meaning of the sentence (Frazier & Rayner, 1982; Inhoff & Weger, 2005; Meseguer et al., 2002; Mitchell et al., 2008; von der Malsburg & Vasishth, 2011; Weger & Inhoff, 2007). These have been said to require access to a representation in memory of both the words in a sentence and their spatial coordinates. We hypothesize that the memory system in which the words are maintained includes ST-WM. When a reader makes a targeted regression, we hypothesize that he or she accesses a word or phrase in an abstract representation of the structure and meaning of the sentence, which does not contain information about the linear order of words, and relates that item to a representation of the words in the sentence in verbal ST-WM, which does maintain order information. The verbal ST-WM representation is linked to a location in visual ST-WM.

In noncumulative self-paced reading and listening, we hypothesize that positive correlations between reading and listening times for segments and task performance mark points at which comprehension is incrementally unsuccessful and controlled processing occurs. The basis for this suggestion is as follows. Longer self-paced reading and listening times result from the reader doing more “work” at a point in a sentence. Applied to individual differences, this implies that, after lexical effects are removed, longer self-paced reading and listening times will occur in less efficient parsers. On this view, longer self-paced reading and listening times will be correlated with poorer comprehension across individuals if the extra work done by less efficient processors does not fully compensate for their inefficiency, or there will be no relation between self-paced reading and listening processing times and task performance if the extra work done by less efficient processors does fully compensate for their inefficiency. If longer self-paced reading and listening times are positively correlated with task performance across individuals, some process other than less efficient processors working harder to accomplish the comprehension task to the average level must have applied. We suggest that these processes include controlled problem-solving operations such as deliberate application of “rules of grammar” that a comprehender has learned, readers’ revisions of mental intonation contours to help determine constituent boundaries, and others that

require accessing representations of a presented sentence in memory.

To our knowledge, only one study has reported correlations across individuals between self-paced reading times and task performance. Caplan et al. (2011) found a positive correlation between accuracy in probe verification and self-paced reading times residualized on word length and frequency at the second verb (*saved*) in sentences such as *The sergeant who the soldier decorated by the lieutenant saved expressed his gratitude* and no correlation at seven other points in six sentence types. Caplan et al. (2011) hypothesized that, at this point of very high processing load, participants experienced incremental comprehension failure and engaged controlled comprehension mechanisms that resulted in both longer self-paced reading times and better comprehension. First-order correlations between ST-WM and self-paced reading times and between ST-WM and task performance suggested that the positive relation between self-paced reading times and comprehension was due to the performances of individuals with higher ST-WM scores. This suggests that individuals with higher ST-WM, who are better able to retrieve information from ST-WM, are able to make greater use of that information in controlled processing and supports the view that ST-WM supports retrieval from memory at these points.

The view that controlled processing involves the use of ST-WM to retrieve items in sentences is consistent with evidence that recall of sentences shows features that suggest the operation of mechanisms that apply in recall of lists. Marslen Wilson and Tyler (1976) reported serial position curves for immediate serial recall of words in sentences with two clauses, sentences with two clauses in which all content words were replaced by randomly selected words of the same grammatical categories (“syntactic prose”), and lists of randomly selected words pronounced with intonation contours designed to parallel two clauses. There were serial

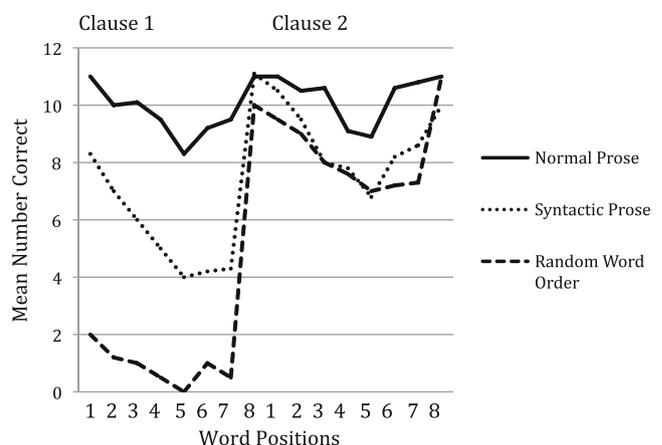


Fig. 2 Word recall as function of serial position in different prose types, adapted from Marslen Wilson and Tyler (1978), Fig. 1

position effects on word recall in each clause in sentences and syntactic prose (Fig. 2), pointing to an interaction of recall based on linguistic structures (reconstruction from a conceptual representation; Lombardi & Potter, 1992; Potter & Lombardi, 1990) and mechanisms that support recall of words in lists. When parsing and interpretation fail, the items in memory are not completely well-structured syntactically and/or semantically but are more like words in lists, increasing the use of mechanisms from ST-WM to retrieve words in a sentence that are needed to allow comprehension to proceed.

Modularity and specialization

We have repeatedly emphasized the *domain specificity* of many important features of memory in parsing and interpretation, and the reader may have formed the view that we endorse a strongly modular view of memory, in which ST-WM and LT-WM/procedural memory are separate memory systems that utilize different memory mechanisms and in which there are different LT-WM/procedural memory “systems.” This formulation, however, oversimplifies specialization and modularity. We close this article with a brief discussion of these topics.

Many researchers have argued for common mechanisms in different memory tasks; for example, Brown et al. (2007) emphasized “qualitative similarity in the retrieval principles involved in memory performance at all timescales” (p. 539). Our arguments are consistent with this view. If incrementally successful parsing and interpretation utilize LT-WM, and if there are memory mechanisms such as content-addressable retrieval that apply in both parsing and interpretation and ST-WM, then these memory mechanisms apply in different memory “systems.”

Rick Lewis (personal communication) has suggested that it might be most profitable to view the memory mechanisms that support parsing and interpretation as being tokens of types of mechanisms used more generally, pointing to the many memory phenomena to which the ACT-R architecture has been applied (list memory: Anderson & Mattessa, 1997; sentence memory: Budiu & Anderson, 2004; expertise: Anderson & Fincham, 1994; Anderson, Fincham, & Douglass, 1999). The idea that memory mechanisms that support different domain-specific skills are variants and specializations of a common functional architecture is appealing because of the successful application of such frameworks and, biologically, because of the conservative nature of evolutionary change.

Nonetheless, regardless of any shared mechanisms or a common functional architecture for the memory processes found in many tasks and skills, there are substantial

differences in the memory systems that support different tasks and skills. In Lewis’s terms, the “tokens” of a common memory system “type” differ significantly. We suggest that several factors contribute to the specialization of memory systems. One is the existence of domain-specific representations in the set of to-be-recalled items. For instance, sentences have syntactic structure not present in lists, which serves to group items in storage and becomes part of retrieval cues. A second is the nature of the task that creates the need for retrieval, which differs with respect to whether a cue is provided (recognition) or not (recall), whether retrieval is consciously initiated (list memory studies) or not (parsing and interpretation), the temporal dimension of memory (which applies not only to the duration over which items are stored, but also to the speed with which they are encoded and retrieved before new material is presented), the extent to which encoding, storage, and rehearsal are performed in isolation or while other operations (such as comprehension or applying the products of comprehension to a task) are performed, and along other dimensions (Humphreys et al., 1989). On this view, domain-specific memory systems arise because different sets of mechanisms are used to satisfy the memory needs of performance under the conditions found in different domains. We suggest that skilled, incrementally successful parsing and interpretation have a unique set of properties that are supported by one such domain-specific adaptation of memory mechanisms.

Conclusion

Recent studies that develop models of retrieval-based parsing have begun to characterize the memory mechanisms that support parsing and interpretation. In much of the literature, these memory mechanisms have been related to ones found in ST-WM. We have argued that trying to attribute the memory mechanisms that support parsing and interpretation to features of ST-WM encounters many problems. We have suggested that the memory system that supports incrementally successful parsing and interpretation is a form of LT-WM and that ST-WM supports retrieval of information by a second processing mode that applies when comprehension is not incrementally successful. The memory system that supports incrementally successful skilled parsing and interpretation may be a specialization of a functional architecture for memory that is common to the memory processes found in many tasks and skills, and this system may utilize mechanisms found in other types of memory. However, as is the case for the memory mechanisms that support other skills, the aggregate of memory mechanisms that support incrementally successful skilled parsing and interpretation is unique.

Acknowledgments This work was supported by grants from NIDCD (DC00942, DC011032) to Drs. Caplan and Waters. We wish to thank Will Evans, Merrill Garrett, David Gow, and reviewers of an earlier version of the manuscript, including Chuck Clifton, Rick Lewis, and Walter Kinsch, for many helpful comments.

References

- Acheson, D. J., & MacDonald, M. C. (2011). The rhymes that the reader perused confused the meaning: Phonological effects during on-line sentence comprehension. *Journal of Memory and Language*, *65*, 193–207.
- Alcoer, P., & Phillips, C. (unpublished manuscript). Using relational syntactic constraints in content-addressable memory architectures for sentence processing. Available at <http://ling.umd.edu/~colin/research/>.
- Anderson, J. R. (1976). *Language, memory, and thought*. Mahwah: Lawrence Erlbaum Associates.
- Anderson, J. R., & Fincham, J. M. (1994). Acquisition of Procedural skills from examples. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1322–1340.
- Anderson, J. R., Fincham, J. M., & Douglass, S. (1999). Practice and retention: A unifying analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1120–1136.
- Anderson, J. R., & Matessa, M. P. (1997). A production system theory of serial memory. *Psychological Review*, *104*, 728–748.
- Ayres, T. J. (1984). Silent reading time for tongue-twister paragraphs. *The American Journal of Psychology*, *97*, 605–609.
- Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *The Quarterly Journal of Experimental Psychology*, *18*, 362–365.
- Baddeley, A. D. (1968). How does acoustic similarity influence short-term memory? *The Quarterly Journal of Experimental Psychology*, *20*, 249–263.
- Baddeley, A. D. (1986). *Working Memory*. New York: Oxford University Press.
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423.
- Baddeley, A. (2003) Working memory: Looking back and looking forward *Nature Reviews Neuroscience*, *4*, 829–839.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory*, vol. 8 (pp. 47–89). New York: Academic Press.
- Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General*, *133*, 83–100.
- Basak, C. (2005). *Capacity limits of the focus of attention and dynamics of the focus switch cost in working memory*. Unpublished doctoral dissertation, Syracuse University, Syracuse, NY.
- Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, *114*, 539–576.
- Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, *107*, 127–181.
- Budiu, R., & Anderson, J. R. (2004). Interpretation-based processing: A unified theory of semantic sentence processing. *Cognitive Science*, *28*, 1–44.
- Burgess, N., & Hitch, G. J. (1992). Towards a network model of the articulatory loop. *Journal of Memory and Language*, *31*, 429–460.
- Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, *106*, 551–581.
- Caplan, D., & Waters, G. S. (1990). Short-term memory and language comprehension: A critical review of the neurophysiological literature. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory* (pp. 337–389). Cambridge: Cambridge University Press.
- Caplan, D., & Waters, G. (1996). Syntactic processing in sentence comprehension under dual-task conditions in aphasic patients. *Language & Cognitive Processes*, *11*, 525–551.
- Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *The Behavioral and Brain Sciences*, *22*(1), 77–94.
- Caplan, D., DeDe, G., Waters, G., Michaud, J., & Tripodis, Y. (2011). Effects of Age, Speed of Processing and Working Memory on Comprehension of Sentences with Relative Clauses. *Psychology and Aging*, *26*, 439–450.
- Caplan, D., Waters, G. S., & Howard, D. (2012). Slave Systems in Verbal Working Memory. *Aphasiology: Special Issue on Neuropsychology of Short Term Memory*, *26*, 279–316.
- Chase, W. G., & Ericsson, K. A. (1981). Skilled memory. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 141–189). Hillsdale: Lawrence Erlbaum Associates.
- Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation*, vol. 16 (pp. 1–58). New York: Academic Press.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge: MIT Press.
- Colle, H. A. (1980). Auditory encoding in visual short-term recall: Effects of noise intensity and spatial location. *Journal of Verbal Learning and Verbal Behavior*, *19*, 722–735.
- Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior*, *15*, 17–32.
- Coltheart, V., Patterson, K., & Leahy, J. (1994). When a ROWS is a ROSE: Phonological effects in written word comprehension. *Quarterly Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 824–845.
- Conrad, R. (1963). Acoustic confusions and memory span for words. *Nature*, *197*, 1029–1030.
- Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology*, *55*, 75–84.
- Conway, C. M., Karpicke, J., & Pisoni, D. B. (2007). Contribution of implicit sequence learning to spoken language processing: Some preliminary findings from normal-hearing adults. *Journal of Deaf Studies and Deaf Education*, *12*, 317–334.
- Conway, C. M., & Pisoni, D. B. (2008). Neurocognitive basis of implicit learning of sequential structure and its relation to language processing. *Annals of the New York Academy of Sciences*, *1145*, 113–131.
- Cowan, N. (1995). *Attention and memory: An integrated framework*. *Oxford Psychology Series (No. 26)*. New York: Oxford University Press.
- Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *The Behavioral and Brain Sciences*, *24*, 87–185.
- Craig, S., & Lewandowsky, S. (in press). Whichever way you choose to categorize, working memory helps you learn. *Quarterly Journal of Experimental Psychology*.
- Craik, F. I. M. (1970). The fate of primary memory items in free recall. *Journal of Verbal Learning and Verbal Behavior*, *9*, 143–148.
- Craik, F. I. M., & Levy, B. A. (1970). Semantic and acoustic information in primary memory. *Journal of Experimental Psychology*, *86*, 77–82.

- Crowder, R. G. (1982). The demise of short term memory. *Acta Psychologica*, *50*, 291–323.
- Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2012, under review). Contrasting interference profiles for agreement and anaphora: Experimental and modeling evidence.
- Eberhard, K. M., Cutting, J. C., & Bock, J. K. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review*, *112*, 531–559.
- Emery, O. B. (1985). Language and aging. *Experimental Aging Research (Monograph)*, *11*, 3–60.
- Engle, R., Kane, M., & Tuholski, S. (1999a). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). Cambridge: Cambridge University Press.
- Engle, R., Tuholski, S., Laughlin, J., & Conway, A. R. A. (1999b). Working memory, short-term memory, and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General*, *128*, 309–331.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*(2), 211–245.
- Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain-specific resources. *Journal of Memory and Language*, *54* (4), 541–553.
- Ferreira, F., & Patson, N. (2007). The good enough approach to language comprehension. *Language and Linguistics Compass*, *1*, 71–83.
- Ferreira, F., Engelhardt, P. E., & Jones, M. W. (2009). Good enough language processing: A satisficing approach. In N. Taatgen, H. Rijn, J. Nerbonne, & L. Schomaker (Eds.), *Proceedings of the 31st Annual conference of the Cognitive Science Society*. Austin: Cognitive Science Society.
- Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2001). Syntactic working memory and the establishment of filler-gap dependencies: Insights from ERPs and fMRI. *Journal of Psycholinguistic Research*, *30*, 321–338.
- Fiebach, J., Vos, S. H., & Friederici, A. D. (2004). Neural correlates of syntactic ambiguity in sentence comprehension for low and high span readers. *Journal of Cognitive Neuroscience*, *16*, 1562–1575.
- Fodor, J. A. (1983). *The Modularity of Mind*, MIT Press.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*, 178–210.
- Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*, 1–76.
- Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1411–1423.
- Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, *30*, 97–114.
- Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science*, *13*, 425–430.
- Grossman, M., Carvell, S., Gollomp, S., Stern, M. B., Vernon, G., & Hurtig, H. I. (1991). Sentence comprehension and praxis deficits in Parkinson's disease. *Neurology*, *41*, 1620–1628.
- Grossman, M., Carvell, S., Stern, M. B., Gollomp, S., & Hurtig, H. (1992). Sentence comprehension in Parkinson's disease: The role of attention and memory. *Brain and Language*, *42*, 347–384.
- Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of NAACL (Vol.2, pp. 159–166)*.
- Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. *Trends in Cognitive Sciences*, *11*(6), 236–242.
- Henkel, L. A., & Franklin, N. (1998). Reality monitoring of physically similar and conceptually related objects. *Memory & Cognition*, *26*, 659–673.
- Henson, R. N. A. (1998). Short-term memory for serial order. The start-end model. *Cognitive Psychology*, *36*, 73–137.
- Henson, R. N., & Burgess, N. (1997). Representations of serial order. In J. A. Bullinaria, D. W. Glasspool, & G. Houghton (Eds.), *4th Neural Computation and Psychology Workshop* (pp. 283–300). London: Springer.
- Henson, R. N. A., Page, M. P., Norris, D., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *The Quarterly Journal of Experimental Psychology*, *49A*(8).
- Hockley, W. E. (1984). Analysis of response time distributions in the study of cognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 598–615.
- Howes, A., Lewis, R. L., & Vera, A. H. (2009). Rational adaptation under task and processing constraints: Implications for testing theories of cognition and action. *Psychological Review*, *116*(4), 717–751.
- Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic and procedural tasks. *Psychological Review*, *96*, 208–233.
- Inhoff A.W., & Weger, U.W. (2005) Memory for word location during reading: Eye movements to previously read words are spatially selective but not precise. *Memory & Cognition*, *33*, 447–461.
- Jalbert, A., Neath, I., & Surprenant, A. M. (2011). Does length or neighborhood size cause the word length effect? *Memory & Cognition*, *39*, 1198–1210.
- James, W. (1890). *The Principles of Psychology*. N.Y: Holt.
- Johnson, M. L., Lowder, M. W., & Gordon, P. C. (2011). The sentence composition effect: Processing of complex sentences depends on the configuration of common versus unusual noun phrases. *Journal of Experimental Psychology: General*, *140*, 707–724.
- Jones, D., Hughes, R. W., & Macken, W. J. (2006). Perceptual organization masquerading as phonological storage: Further support for a perceptual-gestural view of short-term memory. *Journal of Memory and Language*, *54*, 265–281.
- Jones, D. M., Hughes, R. W., & Macken, W. J. (2007). The phonological store abandoned. *Quarterly Journal of Experimental Psychology*, *60*, 505–511.
- Jones, D., Macken, W. J., & Nicholls, A. P. (2004). The phonological store of working memory: Is it phonological and is it a store? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 656–674.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122–149.
- Kane, M. J., & Engle, R. W. (2003). Working memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*, 47–70.
- Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in Working Memory* (pp. 21–48). NY: Oxford University Press.
- Keller, T. A., Carpenter, P. A., & Just, M. A. (2003). Brain imaging of tongue twister sentence comprehension: Twisting the tongue and the brain. *Brain and Language*, *84*(2), 189–203.
- Kennison, S. M. (2004). The effect of phonemic repetition on syntactic ambiguity resolution: Implications for models of working memory. *Journal of Psycholinguistic Research*, *33*, 493–516.

- King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, *30*, 580–602.
- Kontiola, P., Laaksonen, R., Sulkava, R., & Erkinjuntti, T. (1990). Pattern of language impairment is different in Alzheimer's disease and multi-infarct dementia. *Brain and Language*, *38*, 364–383.
- Lalami, L., Marblestone, K., Schuster, S., Andersen, E., Kempler, D., Tyler, L., & MacDonald, M. (1996). On-line vs. off-line sentence processing in Alzheimer's disease. Poster presented at the Meeting of Cognitive Aging, Atlanta.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–117.
- Lewandowsky, S., & Murdock, B. B., Jr. (1989). Memory for serial order. *Psychological Review*, *96*(1), 25–57.
- Lewandowsky, S., & Oberauer, K. (2008). The word-length effect provides no evidence for decay in short-term memory. *Psychonomic Bulletin & Review*, *15*, 875–888.
- Lewis, R. L. (1996). Interference in short-term memory: The magical number two (or three) in sentence processing. *Journal of Psycholinguistic Research*, *25*(1), 93–115.
- Lewis, R. L. (2000). Specifying architectures for language processing: Process, control, and memory in parsing and interpretation. In M. W. Crocker, M. Pickering, & C. Clifton Jr. (Eds.), *Architectures and Mechanisms for Language Processing*. Cambridge: Cambridge University Press.
- Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*, 375–419.
- Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, *10*, 44–54.
- Lieberman, P., Friedman, J., & Feldman, L. S. (1990). Syntax comprehension in Parkinson's disease. *Journal of Nervous and Mental Diseases*, *178*, 360–366.
- Lombardi, L., & Potter, M. C. (1992). The regeneration of syntax in short term memory. *Journal of Memory and Language*, *31*, 713–733.
- MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, *24*, 56–98.
- Marslen Wilson, W., & Tyler, L. (1976). Memory and Levels of Processing in a Psycholinguistic Context. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 112–119.
- Martin, A. E., & McElree, B. (2008). A content-addressable pointer mechanism underlies comprehension of verb-phrase ellipsis. *Journal of Memory and Language*, *58*, 879–906.
- Martin, A. E., & McElree, B. (2009). Memory operations that support language comprehension: Evidence from verb-phrase ellipsis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1231–1239.
- Martin, N., & Ayala, J. (2004). Measurements of auditory-verbal STM in aphasia: Effects of task, item and word processing impairment. *Brain and Language*, *89*, 464–483.
- Martin, R. (1990). The consequences of reduced memory span for the comprehension of semantic versus syntactic information. *Brain and Language*, *38*, 1–20.
- Martin, R. C., & He, T. (2004). Semantic short-term memory and its role in sentence processing: A replication. *Brain and Language*, *89*, 76–82.
- Martin, R. C., Shelton, J., & Yaffee, L. (1994). Language processing and working memory: Neuropsychological evidence for separate phonological and semantic capacities. *Journal of Memory and Language*, *33*, 83–111.
- McCutchen, D., & Perfetti, C. A. (1982). The visual tongue-twister effect: Phonological activation in silent reading. *Journal of Verbal Learning and Verbal Behavior*, *21*, 672–687.
- McCutchen, D., Bell, L. C., France, I. M., & Perfetti, C. A. (1991). Phoneme-specific interference in reading: The visual tongue-twister effect revisited. *Reading Research Quarterly*, *26*, 87–103.
- McDaniel, M. A., & Bugg, J. M. (2008). Instability in memory phenomena: A common puzzle and a unifying explanation. *Psychonomic Bulletin & Review*, *15*, 237–255.
- McElree, B. (1996). Accessing short-term memory with semantic and phonological information: A time-course analysis. *Memory & Cognition*, *24*, 173–187.
- McElree, B. (1998). Attended and non-attended states in working memory: Accessing categorized structures. *Journal of Memory and Language*, *38*, 225–252.
- McElree, B. (2000). Sentence comprehension is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research*, *29*, 111–123.
- McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 817–835.
- McElree, B. (2006). Accessing recent events. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 46). San Diego: Academic Press.
- McElree, B., & Doshier, B. A. (1989). Serial position and set size in short-term memory: Time course of recognition. *Journal of Experimental Psychology: General*, *118*, 346–373.
- McElree, B., & Doshier, B. A. (1993). Serial retrieval processes in the recovery of order information. *Journal of Experimental Psychology: General*, *122*, 291–315.
- McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subservise sentence comprehension. *Journal of Memory and Language*, *4*(1), 67–91.
- Meseguer, E., Carreiras, M., & Clifton, C. (2002). Overt reanalysis strategies during the reading of mild garden path sentences. *Memory & Cognition*, *30*: 551–561.
- Miller, G. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.
- Miller, G. A., & Chomsky, N. (1963). Finitary models of language users. In R. Duncan Luce, R. R. Bush, & G. Eugene (Eds.), *Handbook of mathematical psychology*, vol. 2 (pp. 419–491). New York: Wiley.
- Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology*, *1*, 31.
- Mitchell, D. C., Shen, X., Green, M. J., & Hodgson, T. (2008). Accounting for regressive eye-movements in models of sentence processing: A reappraisal of the Selective Reanalysis hypothesis. *Journal of Memory and Language*, *59*, 266–293.
- Miyake, A., Carpenter, P., & Just, M. (1994). A capacity approach to syntactic comprehension disorders: Making normal adults perform like aphasic patients. *Cognitive Neuropsychology*, *11*, 671–717.
- Murdock, B. B., Jr., & Vom Saal, W. (1967). Transpositions in short-term memory. *Journal of Experimental Psychology*, *74*, 137–143.
- Nairne, J. S. (1990a). A feature model of immediate memory. *Memory & Cognition*, *18*, 251–269.
- Nairne, J. S. (1990b). Similarity and long-term memory for order. *Journal of Memory and Language*, *29*, 733–746.
- Natsopoulos, D., Katsarou, Z., Bostantzopoulou, S., Grouios, G., Mentanopoulos, G., & Logothetis, J. (1991). Strategies in comprehension of relative clauses in Parkinsonian patients. *Cortex*, *27*, 255–268.
- Neath, I. (2000). Modeling the effects of irrelevant speech on memory. *Psychonomic Bulletin and Review*, *7*, 403–423.
- Newell, A. (1990). *Unified Theories of Cognition*. Cambridge: Harvard University Press.

- Nimmo, L., & Roodenrys, S. (2004). Investigating the phonological similarity effect: Syllable structure and the position of common phonemes. *Journal of Memory and Language*, *50*, 245–258.
- Obata, M., Lewis, R. L., Epstein, S., Bartek, B., and Boland, J. (2010). Featural analysis and short-term memory retrieval in on-line parsing: Evidence for syntactic, but not phonological, similarity-based interference. In *Proceedings of NELS 41: Conference of the North East Linguistics Society*, Philadelphia.
- Oberauer, K., & Lewandowsky, S. (2011). Modelling working memory: A computational implementation of the Time-Based Resource-Sharing theory. *Psychonomic Bulletin & Review*, *18*, 10–45.
- Page, M., & Henson, R. N. (2001). Models of short-term memory: Modelling immediate serial recall of verbal material. In Andrade, J. (Ed.), *Working Memory in Perspective* (pp. 177–198). Hove, UK: Psychology Press Ltd.
- Perfetti, C. A., & McCutchen, D. (1982). Speech processes in reading. In N. Lass (Ed.), *Speech and language: Advances in basic research and practice* (pp. 237–269). New York: Academic Press.
- Phillips, C., Wagers, M. W., & Lau, E. F. (2011). Grammatical illusions and selective fallibility in real-time language comprehension. In J. Runner (Ed.), *Experiments at the Interfaces, Syntax & Semantics, vol 37* (pp. 153–186). Bingley: Emerald Publications.
- Poirier, M., & Saint-Aubin, J. (1996). Immediate serial recall, word frequency, item identity and item position. *Canadian Journal of Experimental Psychology*, *50*(4), 408–412.
- Postman, L., & Keppel, G. (1977). Conditions of cumulative proactive inhibition. *Journal of Experimental Psychology. General*, *106*, 376–403.
- Pothos, E. M. (2007). Theories of Artificial Grammar Learning. *Psychological Bulletin*, *133*, 227–244.
- Potter, M. C., & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Learning*, *29*, 633–654.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.
- Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model effects of higher level language processing on eye-movements in reading. *Psychonomic Bulletin & Review*, *16*(1), 1–20.
- Reinhart, T. M. (1976). *The Syntactic Domain of Anaphora*. (Doctoral dissertation, Massachusetts Institute of Technology). (Available online at <http://dspace.mit.edu/handle/1721.1/16400>).
- Ricker, T., AuBuchon, A. M., & Cowan, N. (2010). Working memory. *Wiley interdisciplinary reviews: Cognitive science*, *1*:573–585.
- Rochon, E., & Saffran, E. M. (1995). *A semantic contribution to sentence comprehension impairments in Alzheimer's disease*. San Diego: Poster presented at the Academy of Aphasia.
- Rochon, E., Waters, G. S., & Caplan, D. (1994). Sentence comprehension in patients with Alzheimer's Disease. *Brain and Language*, *46*, 329–349.
- Saults, J. S., & Cowan, N. (2007). A central capacity limit to the simultaneous storage of visual and auditory arrays in working memory. *Journal of Experimental Psychology. General*, *136*(4), 663–684.
- Shallice, T. (1975). On the contents of primary memory. In P. M. A. Rabbit & S. Domic (Eds.), *Attention and performance, vol. 5* (pp. 269–280). London: Academic Press.
- Shallice, T., & Vallar, G. (1990). The impairment of auditory-verbal short-term storage. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory* (pp. 11–53). Cambridge: Cambridge University Press.
- Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition*, *116*, 71–86.
- Stevens, K., Kempler, D., Andersen, E., & MacDonald, M. C. (1996). Preserved use of semantic and syntactic context in Alzheimer's disease. Poster presented at the Meeting of Cognitive Aging, Atlanta.
- Sturt, P., & Crocker, M. (1996). Monotonic syntactic processing: A cross-linguistic study of attachment. *Language & Cognitive Processes*, *11*(5), 449–494.
- Tomoeda, C. K., Bayles, K. A., Boone, D. R., Kaszniak, A. W., & Slauson, T. J. (1990). Speech rate and syntactic complexity effects on the auditory comprehension of Alzheimer patients. *Journal of Computer Disorders*, *23*, 151–161.
- Traxler, M. J., & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language*, *35*, 454–475.
- Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, *101*, 80–102.
- Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*, 231–270.
- Underwood, B. J., & Ekstrand, B. R. (1967). Effect of distributed practice on paired-associate learning. *Journal of Experimental Psychology*, *73*, 1–21.
- Van Dyke, J., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalysed ambiguities. *Journal of Memory and Language*, *49*, 285–316.
- Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language*, *55*, 157–166.
- Van Dyke, J. A., & McElree B. (2011) Cue-dependent interference in comprehension, *Journal of Memory and Language*, *65*, 247–263.
- Vasishth, S., & Lewis, R. L. (2006). Argument-head distance and processing complexity: Explaining both locality and anti-locality effects. *Language*, *82*, 767–794.
- Verhaeghen, P., & Basak, C. (2005). Aging and switching of the focus of attention in working memory: Results from a modified n-back task. *The Quarterly Journal of Experimental Psychology. A*, *58*, 134–154.
- Verhaeghen, P., Cerella, J., & Basak, C. (2004). A working memory workout: How to change to size of the focus of attention from one to four in ten hours or less. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *30*, 1322–1337.
- Verhaeghen, P., Cerella, J., Basak, C., Bopp, K. L., Zhang, Y., & Hoyer, W. J. (2007). The ins and outs of working memory: Dynamic processes associated with focus switching and search. In N. Osaka, R. Logie, & M. D'Esposito (Eds.), *Working memory: Behavioural & neural correlates* (pp. 81–98). Oxford: Oxford University Press.
- von der Malsburg, T. and Vasishth, S. (2011). What is the scanpath signature of syntactic reanalysis? *Journal of Memory and Language*, *65*(2):109–127.
- Wanner, E., & Maratsos, M. P. (1978). An ATN approach to comprehension. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality Cambridge*. MA: MIT Press.
- Waters, G., & Caplan, D. (1996). Processing resource capacity and the comprehension of garden path sentences. *Memory & Cognition*, *24*, 342–355.
- Waters, G., Caplan, D., & Hildebrandt, N. (1987). Working memory and written sentence comprehension. In M. Coltheart (Ed.), *Attention and Performance XII: The Psychology of Reading* (pp. 531–555). London: Lawrence Erlbaum.
- Waters, G. S., Caplan, D., & Hildebrandt, N. (1991). On the structure of verbal short-term memory and its functional role in sentence

- comprehension: Evidence from neuropsychology. *Cognitive Neuropsychology*, 8(2), 81–126.
- Waters, G. S., Caplan, D., & Leonard, C. (1992). The role of phonology in reading comprehension: Implications of the effects of homophones on processing sentences with referentially dependent categories. *The Quarterly Journal of Experimental Psychology. A*, 44, 343–372.
- Waters, G. S., Caplan, D., & Rochon, E. (1995). Processing capacity and sentence comprehension in patients with Alzheimer's Disease. *Cognitive Neuropsychology*, 12, 1–30.
- Watkins, M. J., Watkins, O. C., & Crowder, R. G. (1974). The modality effect in free and serial recall as a function of phonological similarity. *Journal of Verbal Learning and Verbal Behavior*, 13, 430–447.
- Weger, U.W. & Inhoff, A.W. (2007). Long-range regressions to previously read words are guided by spatial and verbal memory. *Memory & Cognition*, 35, 1293–1306.
- Wickelgren, W. A. (1965). Short-term memory for phonemically similar lists. *The American Journal of Psychology*, 78, 567–574.
- Wickelgren, W. A. (1966). Associative intrusions in short-term recall. *Journal of Experimental Psychology*, 72, 853–858.
- Zhang, S., & Perfetti, C. A. (1993). The tongue twister effect in reading Chinese. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 19, 1082–1093.