CHAPTER 5

1755

The Role of Memory in Language and Communication

CHAPTER OUTLINE

Julie A. Van Dyke

Types of Memory
  Multiple Memory Systems
  Semantic Memory
  Episodic Memory
  Nondeclarative Memory
  Procedural Memory
  Working Memory
Working Memory and Language Comprehension
  Problems with the Capacity View
Implications for Language Processing
  Forgetting
Conclusions

Language comprehension requires the ability to construct linguistic dependencies between nonadjacent constituents. For example, a subject must agree with its verb, but the two are often separated by several words, phrases, or even clauses, as in *The athlete(s) in the training program run(s) every day* or *The athlete(s) in the training program that was designed by an Olympic gold-medal winner run(s) every day.* At the same time, research in verbal memory has long recognized that our ability to actively attend to and concurrently process information is severely limited. This constraint leads to a functional requirement for the language comprehension system: comprehenders must *retrieve* items that have already been processed in order to fully integrate new information into an evolving interpretation. Naturally then, limitations on memory storage and retrieval are important determinants of language performance. In addition, language systems must interact with linguistic and conceptual knowledge in order to create meaning. For example, native speakers of English will immediately perceive the unacceptability of **Sam uncrossed the street*, while *Sam uncrossed his arms* is perfectly acceptable. This points to a second functional requirement: the need to retrieve passively held lexical and conceptual knowledge about meanings of words and when particular grammatical devices (here "un") may be applied. These two requirements demonstrate the close dependence of language processes on memory, suggesting that a thorough understanding of language processing, and acquired language disorders, will benefit from an understanding of the healthy memory system. It is the goal of the current chapter to provide a brief review of this literature.

TYPES OF MEMORY

Ever since the early days of psychology, when the discipline was more akin to philosophy than science, thinkers who concerned themselves with the phenomenon of memory found cause to make distinctions based on the *type* of information held in that memory. For example, the French philosopher Maine de Biran proposed three distinct memories, which he referred to as mechanical, sensitive, and representative, each depending on different mechanisms and characterized by different properties (Maine de Biran, 1804/1929). According to Biran, mechanical memory involved the acquisition of motor and verbal habits and operates unconsciously; sensitive memory involves feelings and affect and also operates unconsciously; and representative memory involves the conscious recollection of ideas and events. A second early distinction was made by William James in his seminal text *Principles of Psychology* (1890), where he focused on temporal properties of particular memories, contrasting elementary memory (also called primary memory) and secondary memory. He wrote:

> Elementary memory makes us aware of . . . the just past. The objects we feel in this directly intuited past differ from properly recollected objects. An object which is recollected, in the proper sense of the term, is one which has been absent from consciousness

> altogether, and . . . is brought back . . . from a reservoir in which, with countless other objects, it lay buried and lost from view. But an object of primary memory is not thus brought back; it never was lost; its date was never cut off in consciousness from that of the immediately present moment. In fact it comes to us as belonging to the rearward portion of the present space of time, and not to the genuine past. (pp. 646–647)

Distinctions such as these have remained relevant even to the present day, with the field of memory research being divided into those that characterize memory based on separate *systems*, largely aligned to the type of information they contain, or else based on separate *processes*, which focuses on mechanisms of retrieval and forgetting. Both approaches are discussed later.

Multiple Memory Systems

The multiple memory systems approach focuses on identifying functionally and anatomically distinct systems, which differ in their "methods of acquisition, representation, and expression of knowledge" (Tulving, 1985, p. 3). There are a number of different versions of this approach. For example, Squire (2004; Squire & Zola-Morgan, 1988) suggested that the most fundamental distinction is between declarative and nondeclarative memories. Declarative memory is what is usually meant by the term *memory* in ordinary language, and is the kind of memory impaired in amnesia, that relating to the conscious recollection of facts and events. For this reason it has also been termed explicit memory. It provides a representational vocabulary for modeling the external world, and the resulting models can be evaluated as either true or false with respect to the world. It is typically assessed by tests of recall, recognition, or cued recall. In contrast, nondeclarative memory is actually a catch-all term referring to a variety of other memories, including most notably *procedural* memory. Nondeclarative (or *implicit*) memories have in common that they are expressed through action rather than recollection. As such, they are not true or false, but rather reflect qualities of the learning experience. Strong evidence in support of this distinction comes from studies of amnesic patients from as early as Milner (1962), who demonstrated that patient H. M. could learn a mirror drawing task (invoking procedural memory), but displayed no memory of actually having practiced the task before (a declarative memory). Additional demonstrations have shown normal rates of learning in a variety of skills without conscious awareness that the learning has taken place (cf. Squire, 1992, for a review).

Studies from brain damaged patients and animal models point to medial temporal lobe structures, including the hippocampal region and the adjacent entorhinal, perirhinal, and the parahippocampal cortices as crucial for establishing new declarative memories (Buckner & Wheeler, 2001; Squire, Stark, & Clark, 2004). These structures are significant because they receive multi-modal sensory input via reciprocal pathways from frontal, temporal, and parietal areas, enabling them to consolidate inputs from these regions (Alvarez & Squire, 1994; McClelland, McNaughton, & O'Reilly, 1995). A hallmark of damage to the medial temporal lobe is profound forgetfulness for any event occurring longer than 2 seconds in the past (Buffalo, Reber, & Squire, 1998), regardless of sensory modality (e.g., Levy, Manns, Hopkins, et al., 2003, Milner, 1972; Squire, Schmolck, & Stark, 2001). In addition, impairment in recollection of declarative memories can occur despite intact perceptual abilities and normal performance on intelligence tests (Schmolck, Kensinger, Corkin, & Squire, 2002; Schmolck, Stefanacci, & Squire, 2000), lending support to the idea that declarative memory may constitute a separable memory system. Over time, however, memories become largely independent of the medial structures, and more dependent on neocortical structures, especially in the temporal lobes.

In contrast, there is no specific brain system related to establishing nondeclarative memories, as the category includes a variety of different types of memories. For example, creation of memories via classic conditioning depends on the cerebellum and amygdala (e.g., Delgado, Jou, LeDoux, & Phelps, 2009; Thompson & Kim, 1996), while procedural learning depends on the basal ganglia, especially the striatum (e.g., Packard, Hirsh, & White, 1989; Poldrack, Clark, Pare-Blagoev, et al., 2001; Salmon & Butters, 1995; Ullman, 2004). Figure 5-1 provides a summary of the subtypes of memory falling under each of these two distinctions, together with the primary brain structures that have been shown to support these memories in humans and experimental animals.

A second frequently cited taxonomy of memory systems is that developed by Tulving and colleagues (Schacter & Tulving, 1994; Tulving, 1983). This approach is particularly concerned with establishing a distinction between two subtypes of declarative memory: semantic and episodic memory, distinguished by the relation of a particular piece of knowledge to a particular individual. For example, the knowledge that pizza is made with cheese and tomato sauce would reside in semantic memory, and would be shared by everyone, while the knowledge that Andrew had two slices of mushroom

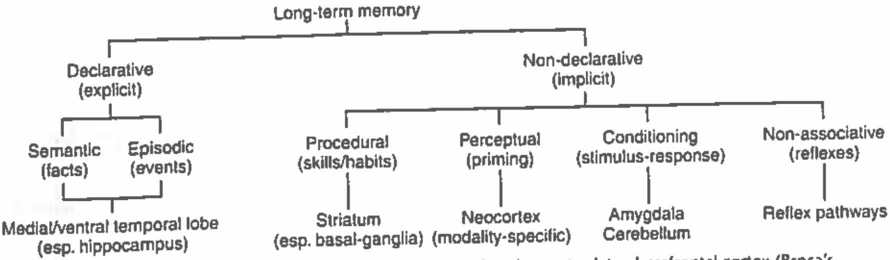

Figure 5-1 A taxonomy of long-term memory. Note that the ventro-lateral prefrontal cortex (Broca's area and its right homologue) have been implicated in semantic, episodic, and procedural memories). Brain regions supporting the perceptual representation system depend on the perceptual modality; the ventral occipital-temporal region is claimed to store word forms. [Adapted from Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory, 82*, 171–177.]

pizza and a soda for lunch would reside in episodic memory and be held only by Andrew and those who ate together with him. In addition to these two separate memory systems, these researchers have argued for the functional and neurological distinctness of three others: perceptual representation system (PRS), procedural memory, and working memory (WM). These first two would be considered nondeclarative memories under the taxonomy suggested by Squire and colleagues, and WM would be considered as a separate memory system all together according to this approach. We discuss the behavioral and neurological properties of each of these five systems in turn.

Semantic Memory

Semantic memory is generally assessed through object naming ("This is a picture of a ____"), and queries that require access to world knowledge (i.e., the Pyramids and Palm Trees test (Howard & Patterson, 1992) in which individuals must decide what type of tree is most associated with an Egyptian pyramid. Synonym generation tasks, in which patients are asked to name as many exemplars of a provided category in 1 minute, have also been used to evaluate fluency and speed of accessing categories of information. In addition to storing facts about the world, semantic memory is the repository for linguistic knowledge about words, including phonological (e.g., that the word *night* rhymes with *kite*), morphological (e.g., that *taught* is the past tense of *teach*), grammatical (e.g., that *hit* takes a direct object), and semantic properties (that *sleep* and *snooze* are synonyms). This knowledge has been referred to as the *mental lexicon* (Ullman, 2004; see also Chapter 6).

Cognitive psychology has long been interested in the organization of the mental lexicon, especially its semantic aspects, and the means through which it supports comprehension and communication (cf. Murphy, 2002, for a review). That knowledge is *organized* has been demonstrated experimentally in numerous studies observing correlations between reaction times to verify relationships between concepts and their degree of relation. For example, the early study of Collins and Quillian (1969) found that participants took less time to verify the statement "A canary is a bird" compared to "A canary is an animal." They interpreted this result as evidence for a hierarchical representation of concepts, such as that presented in Figure 5-2, where relationships between categories are represented by solid lines and properties of individual objects are represented by dashed lines. Since the concept *canary* is closer to *bird* than to *animal*, they reasoned that the faster reaction time was possible because there were fewer links to traverse in order to verify the statement. Later research revealed a situation not so simple as this, as statements about items that are more *typical* of a category, such as "A robin is a bird" were judged more quickly than atypical exemplars, such as "An ostrich is a bird," despite the fact that they are both located at the same level of the conceptual hierarchy (Rips, Shoben, & Smith, 1973; Rosch & Mervis, 1975). With this result, it became clear that knowledge organization reflects not just static or logical relationships between concepts, but also an individual's experience with the world.

There is now a substantial amount of evidence from neuroimaging techniques (e.g., PET, fMRI) that experience with the world determines how the mental lexicon is stored in the brain. For example, reading action words that are semantically related to different body parts (e.g., "kick," "pick," "lick") activates regions of the motor and premotor cortex responsible for controlling

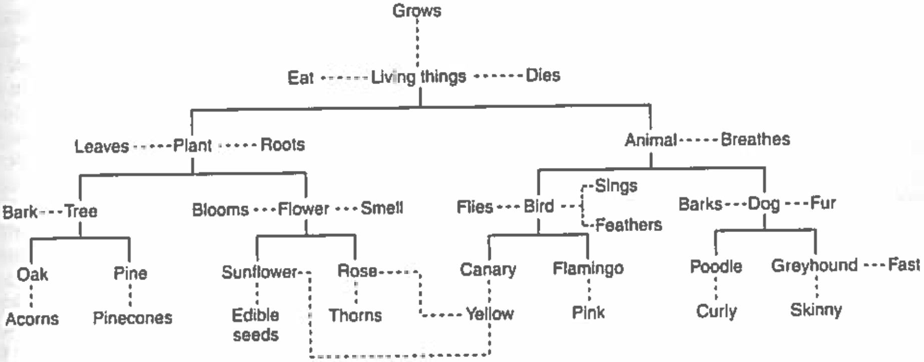

Figure 5-2 Hierarchical semantic network. Categorical relationships are depicted with solid lines between concepts, properties of individual concepts are depicted as dashed lines. Note that properties may apply to multiple concepts, and certain category members need not inherit all properties from its category (e.g., ostriches are birds but don't fly). Prototypical members of a category will inherit most of the properties of that category, however. [Adapted from Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*, 240–247.]

those body parts (Aziz-Zadeh, Wilson, Rizzolatti, & Iacoboni, 2006; Hauk, Johnsrude, & Pulvermuller, 2004; Pulvermuller, 2005; Tettamanti, Buccino, Saccuman, et al., 2005). Similarly, reading or naming words associated with tool actions (e.g., hammer) activate a network of sensorimotor regions also engaged when perceiving and using tools (Chao, Haxby, & Martin, 1999). In addition to the influence of embodiment, a variety of other properties of objects in the world appear to have dedicated temporal lobe regions in which they are processed, and accessing words associated with these properties activates adjacent brain regions. For example, color and motion perception are associated with separate regions in the left ventral and medial temporal lobe, respectively (Corbetta, Miezin, Dobmeyer, et al., 1990; Zeki, Watson, Lueck, et al., 1991). In a study in which participants were shown achromatic pictures of objects (e.g., line drawing of a pencil), and asked to generate color words (e.g., "yellow") and action words (e.g., "write") related to these objects, Martin, Haxby, Lalonde, et al. (1995) found that regions just anterior to the ventral and medial lobe regions just mentioned were active. Similar findings have been observed for size and sounds of objects (Kellenbach, Brett, & Patterson, 2001), as well as grammar related properties such as the animate/inanimate distinction (Chao, Haxby, & Martin, 1999). Taken together, these results suggest that knowledge in the mental lexicon is represented by a distributed network of features processed primarily in the temporal lobes, with different object categories eliciting different patterns of activation among relevant features (Martin & Chao, 2001; McClelland & Rogers, 2003). In addition to these temporal regions, neuroimaging suggests that the retrieval and selection of information in the mental lexicon is managed by the left ventrolateral prefrontal cortex, corresponding to the inferior frontal gyrus (including Broca's area) and Brodmann's areas 44, 45, and 47 (cf. Bookheimer, 2002; Thompson-Schill, 2003, for reviews). These areas will also become relevant later in the discussion of the interaction of memory and syntactic processing.

Episodic Memory

The existence of a separate episodic memory system appears to receive strong motivation from data from amnesic patients, who have specific deficits in episodic memory with very few, if any, deficits in the other memory systems. Such pathology suggests that episodic memories should be dissociable from other types of memories, and may occupy a neurologically distinct region in the brain (Tulving, 2002, p. 12). Assessment of episodic memory proves difficult, however, since the personal nature of these memories limits the ability of experimenters to manipulate them and evaluate the correctness of responses. Consequently, many studies investigating episodic memory utilize list-learning paradigms, which give experimenters complete control over properties of the to-be-remembered stimuli. Participants are presented with a list of words (or visual items such as faces or patterns) and asked to report on arious incidental properties of them during an encoding phase (e.g., whether presented in upper or lower ase letters, in particular colors, with a particular other vord, or even whether it occurred at all). The subsejuent retrieval phase then asks them to make judgnents about whether items have been seen before recognition) or to produce the item or its associates recall), and sometimes to specify whether they conciously remember learning the word during the study hase or not (remember/know judgment). These paraligms enable experimenters to directly examine the onditions that lead to the successful creation of memries. For example, a group of studies have investigated subsequent memory effects" in which sets of items hat have been identified via post-hoc memory tests as aving been successfully remembered are contrasted vith those that have not been remembered (e.g., Rugg, tten, & Henson, 2002; Wagner, Koutstaal, & Schacter, 999). The goal was to uncover brain regions specifically nvolved in task-invariant episodic encoding, however uch a region has so far resisted identification. Instead, he main result from these studies is that the pattern of rain activation associated with a particular memory liffers depending on the type of processing engaged luring study (e.g., Kelley, Miezin, McDermott, et al., 998; McDermott, Buckner, Petersen, et al., 1999; Otten & Rugg, 2001; Wagner, Poldrack, Eldridge, et al., 1998). hus, words encoded via a semantic task (i.e., judging vhether a word is animate) activate areas of the medial refrontal cortex and in the dorsal part of the left infeior frontal gyrus, which has been linked to semantic VM (e.g., Buckner & Koutstaal, 1998; Gabrieli, Poldrack, & Desmond, 1998; Wagner et al., 1998). Words enoded via a syllable counting task, on the other hand, ailed to activate any prefrontal areas, and instead howed activations in bilateral parietal and fusiform egions and in the left occipital cortex, areas that ave been implicated in phonological processing tasks e.g., Mummery, Patterson, Hodges, & Price, 1998; oldrack, Wagner, Prull, et al., 1999; Price, Moore, Humphreys, & Wise, 1997).

Such task-specific activations are consonant with the dea that a memory for a particular stimulus includes a ariety of incidental information about the context in vhich it was remembered—even including subjective actors such as mood or cognitive state. Thus, episodic nemories—like semantic memories—are represented n the brain as distributed networks of activation, ointing to the need for a more refined explanation of pisodic amnesia than simply to look for the region hat houses them. One approach is the idea that damage must be specific to the mechanism through which hese ideas are reactivated (wherever they may be stored), and not the means through which they are stored. Indeed, Tulving and Pearlstone (1966) pointed out that much of what we commonly view as memory loss—a memory no longer being *available*—is in fact more properly viewed as a failure in *accessibility*. Subsequently, Tulving (1979) formulated the *encoding specificity principle*, which states "[t]he probability of successful retrieval of the target item is a monotonically increasing function of information overlap between the information present at retrieval and the information stored in memory" (p. 408). Indeed, a recent survey of neuroimaging research concludes that the same brain areas are active both at encoding and retrieval (Danker & Anderson, 2010). One demonstration of this idea is the classic study by Thomson and Tulving (1970), who observed the expected result when no associate for the target word *flower* was present during the study phase: a strong associate presented at test (*bloom*) elicited recall of *flower* better than no associate or than a weak associate (*fruit*) presented at test. But when the weak associate is presented during the study phase, the presence of this same weak associate at test produces markedly better recall than when the strong associate is presented (73% versus 33% correct recalls). Thus, the effectiveness of even a longstanding cue, drawn from semantic memory, depends crucially on the processes that occurred when particular episodic memories are created.

While the foregoing discussion has centered around studies of memory per se, evidence for the role of encoding context and its interaction with the information available at retrieval has also been observed in studies of language comprehension. In order to isolate the importance of encoding versus retrieval operations, Van Dyke and McElree (2006) manipulated the cues available at retrieval during sentence processing while keeping the encoding context constant. We tested grammatical constructions in which a direct object has been displaced from its verb by moving it to the front of the sentence (e.g., *It was the boat that the guy who lived by the sea sailed in two sunny days*). Here, when the verb *sailed* is processed, a retrieval must occur in order to restore the noun phrase *the boat* into active memory so that it can be integrated with the verb. We manipulated the encoding context by asking participants to remember a three-word memory list prior to reading the sentence (e.g., TABLE-SINK-TRUCK); this memory list was present for some trials (Load Condition) and not for others (No Load Condition). The manipulation of retrieval cues was accomplished by substituting the verb *fixed* for *sailed*, creating a situation where four nouns stored in memory (i.e., table, sink, truck, boat) are suitable direct objects for the verb *fixed* (Matched

Condition), while only one is suitable for the verb *sailed* (Unmatched Condition). The results of reading times on the manipulated verb are shown in Figure 5-3; when there was no load present, there was no difference in reading times, however the presence of the memory words led to increased reading times when the verb was Matched as compared to when it was Unmatched. Thus, as predicted by encoding specificity, the overlap between retrieval cues generated from the verb (e.g., cues that specify "find a direct object that is fixable/sailable") and contextual information was a strong determinant of reading performance. We note, however, that an important difference between this study and the Thomson and Tulving (1970) study is that here, the match between the cues available at retrieval and the encoding context produced a *detrimental* effect. This is because the similarity between the context words and the target word (i.e., table, sink, truck, and boat are all fixable) created *interference* at retrieval. We will discuss the role of interference in memory and language further in the section on forgetting. The important point here, however, is that encoding context has its effect *in conjunction with* the retrieval cues used to reaccess the encoded material.

An important unresolved question pertains to the relationship between episodic memories and semantic memories. From the perspective of the multiple memories approach, these two types of memories are considered to be separate *systems*; however, the criteria by which a system is determined had been criticized as indecisive (e.g., Surprenant & Neath, 2009). Much of the support for separate systems comes from functional and neurological *dissociations*, such that tasks that are diagnostic of System A, or brain regions implicated in the healthy functioning of System A are different from those tapping into the function of System B. While dissociations are common, a number of researchers have published papers questioning their logic as a means for identifying separate brain systems (e.g., Ryan & Cohen, 2003; Van Orden, Pennington, & Stone, 2001). For example, Parkin (2001) notes that apparent dissociations observed in amnesic patients, who show unimpaired performance on standardized tests of semantic memory, but intense difficulty recalling episodic events such as a recently presented word list or lunch menu, are confounded by test difficulty. In the face of the temporally graded nature of amnesia, in which more recently acquired memories are the most susceptible to loss, the key problem with these assessments of semantic memory is that they test information that was acquired by early adult life. When semantic memory tests are carefully controlled so as to test more recently acquired semantic memories, the relative sparing of one system over the other is less apparent.

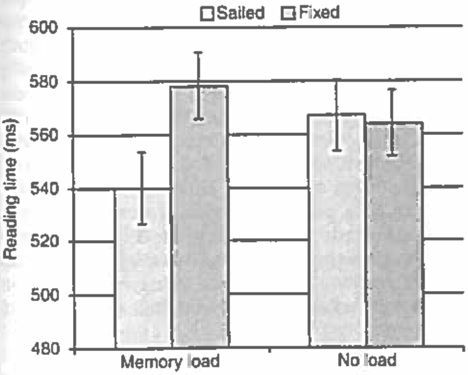

Figure 5-3 Participants took longer to read the verb *fixed* when it was preceded by a memory list of *fixable* words. The same memory list did not affect reading times for *sailed*. [Results from Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language, 55*, 157–166.]

While the debate on whether episodic and semantic memories are distinct systems will likely continue, from the point of view of language, it is at least helpful to distinguish autobiographical episodic memories, which are not fundamentally related to language processing, from other contextually anchored memories (e.g., Conway, 2001). The relationship between the latter type of episodic memories and semantic memory seems intrinsic—no scientist has ever claimed that individuals are *born* knowing the conceptual knowledge that comprises the meaning of words in the mental lexicon.[1] These must be learned through experience with the world and with language. There is now a considerable body of evidence suggesting that the meaning—and grammatical usage—of individual words is learned (even by infants) through repeated learning episodes (e.g., Harm & Seidenberg, 2004; Mirković, MacDonald, & Seidenberg, 2005; Sahni, Seidenberg, & Saffran, 2010). Fewer learning episodes appears to produce low quality lexical representations, characterized by variable and

---

[1] In contrast, there *have* been prominent proposals within linguistic theory that children are born with knowledge of grammar—a so-called Universal Grammar (e.g., Chomsky, 1986; Crain & Thornton, 1998), through which they can deduce the rules of grammar specific to their own native language. Some theories of language assign these rules to procedural memory (discussed later); the statistical learning approaches discussed here offer an important alternative to this approach.

inconsistent phonological forms, and more shallow meaning representations, incomplete specification of grammatical function, and (for reading) underspecified orthographic representations (Perfetti, 2007). The process of consolidating individual learning events into efficiently accessed long-term memory (LTM) representations has been attested in the domain of reading by neuroimaging studies showing that repeated exposures to a word results in *reduced* activation in reading-related brain regions, especially areas of the ventral occipital-temporal region thought to contain visual word forms (cf. McCandliss, Cohen, & Dehaene, 2003, for a review), and in the inferior frontal gyrus (e.g., Katz, Lee, Tabor, et al., 2005; Pugh, Frost, Sandak, et al., 2008). This reduction is consistent with studies of perceptual and motor skill learning in which initial (unskilled) performance is associated with increased activation in task-specific cortical areas, to be followed by task-specific *decreases* in activation in the same cortical regions after continued practice (e.g., Poldrack & Gabrieli, 2001; Ungerleider, Doyon, & Karni, 2002; Wang, Sereno, Jongman, & Hirsch, 2003).

Although episodic and semantic memories—both declarative memories—are generally characterized as explicit memories, in that they are accessible to conscious reporting, the process of learning that binds the two is not conscious. The ability to learn via repeated exposures engages the brain's ability to extract statistical regularities across examples, which occurs gradually over time without conscious awareness (Perruchet & Pacton, 2006; Reber, 1989; Reber, Stark, & Squire, 1998; Squire & Zola, 1996). A number of recent studies have shown that infants as young as 8 months old are sensitive to the statistical regularities that exist in natural languages and can use them, for example, to identify word boundaries in continuous speech (Saffran, Aslin, & Newport, 1996; Sahni, Seidenberg, & Saffran, 2010) and to learn grammatical and conceptual categories (Bhatt, Wilk, Hill, & Rovee-Collier, 2004; Gerken, Wilson, & Lewis, 2005; Shi, Werker, & Morgan, 1999). Computational models that implement this learning process over a distributed representation of neuronlike nodes (i.e., *connectionist* models) have demonstrated that the resulting networks produce humanlike performance in language acquisition and language comprehension (e.g., Seidenberg & MacDonald, 1999), including the same types of performance errors common to children learning language and adults processing ambiguous sentences.

Nondeclarative Memory

In contrast to the earlier discussion, the traditional taxonomy depicted in Figure 5-1, suggests a clear separation between explicit and implicit memories. Historically, this reflected the need to account for certain cases of amnesia (e.g., patient H.M., Scoville & Milner, 1957) in which patients displayed increasing improvement on complex cognitive skills (i.e., game playing) with no ability to recall ever having learned to play the game or even playing it previously. The explanation afforded was that while damage to the medial temporal lobe structures destroyed the ability to access declarative memory, these patients' nondeclarative memory (especially procedural memory), which does not depend on these brain regions was intact. This memory is characterized as *implicit* because patients are unaware of the learning that has taken place.

A second type of implicit memory that has observed in amnesic patients with an inability to access semantic memory is the preservation of priming effects. That is, these patients display improved performance in recognition tasks for target items following the previous presentation of the same object or some other object that is identical to the target on some perceptual dimension (e.g., sound, shape, etc.). Notably, patients need not be aware that the primed object occurred in order for these effects to occur, and in many cases primes are presented extremely quickly or extremely faintly, so as to be below the threshold of conscious perception.

Schacter and Tulving (1994) proposed two separate memory systems to account for these results: the procedural memory system (discussed later) and the PRS, comprised of a collection of domain-specific modules, which was responsible for priming results (Schacter, Wagner, & Buckner, 2000). The visual word form area, noted earlier, has been offered as one of the modules comprising the PRS (Schacter, 1992), based mainly on evidence from aphasics who show normal priming effects for the surface form of novel words, which consequently could not be stored in semantic memory (e.g., Cermak, Verfaellie, Milberg, et al., 1991; Gabrieli & Keane, 1988; Haist, Musen, & Squire, 1991; Bowers & Schacter, 1992). In addition, it has been observed that some amnesic patients can read irregularly spelled or unknown words, despite having no apparent contact with their meaning (e.g., Funnel, 1983; Schwartz, Saffran, & Marin, 1980). This has been interpreted as support for a separate word form representation independent of meaning.

Procedural Memory

Of the implicit memory systems, procedural memory—memory for *how* to do something—has received the most attention, both in the memory and in the language domain. It has been claimed to support the learning of new, and the control of established, sensorimotor and cognitive habits and skills, including riding a bicycle and skilled game playing. As with all implicit memory systems, learning is gradual and unavailable to conscious description, however in the procedural system the outcome of learning is thought to be *rules*, which are rigid, inflexible, and not influenced by other mental systems (Mishkin, Malamut, & Bachevalier, 1984; Squire & Zola, 1996). Neurologically, the system is rooted in the frontal lobe and basal ganglia, with contributions from portions of the parietal cortex, superior temporal cortex and the cerebellum. The frontal lobe, especially Broca's area and its right homologue, is important for motor sequence learning (Conway & Christiansen, 2001; Doyon, Owen, Petrides, et al., 1996) and especially learning sequences with abstract and hierarchical structures (Dominey, Hoen, Blanc, & Lelekov-Boissard, 2003; Goschke, Friederici, Kotz, & van Kampen, 2001). The basal ganglia have been associated with probabilistic rule learning (Knowlton, Mangels, & Squire, 1996; Poldrack, Prabhakaran, Seger, & Gabrieli, 1999), stimulus-response learning (Packard & Knowlton, 2002), sequence learning (Aldridge & Berridge, 1998; Boecker, Dagher, Ceballos-Baumann, et al., 1998; Doyon, Gaudreau, Laforce, et al., 1997; Graybiel, 1995, Peigneux, Maquet, Meulemans, et al., 2000; Willingham, 1998), and real-time motor planning and control (Wise, Murray, & Gerfen, 1996).

From the perspective of language, one prominent proposal (Ullman, 2004) suggests that the procedural memory system should be understood as the memory system that subserves grammar acquisition and use. Implicit in this proposal is an understanding of grammar as fundamentally rule-based; an idea with a long (and controversial) history in linguistic theory (e.g., Chomsky, 1965, 1980; Marcus, 2001; Marcus, Brinkmann, Clahsen, et al., 1995; Marcus, Vijayan, Bandi, et al., 1999; Pinker, 1991). A frequently cited example is the rule that describes the past tense in English, namely, verb stem + *ed*. This rule allows for the inflection of novel words (e.g., *texted*) and accounts for the phenomenon of overgeneralizations in toddlers (e.g., Daddy goed to work). According to this view, the language-related functions of the neurological structures that support procedural memory are expected to be similar to their nonlanguage function. Thus, the basal ganglia and Broca's area (especially BA 44) are hypothesized to govern control of hierarchically structured elements in complex linguistic representations and assist in the learning of rules over those representations.

This approach is incompatible with the connectionist approach, discussed earlier, in which regularities in language are represented in distributed networks extracted through the process of statistical learning. In these models there are no rules, and indeed, one connectionist implementation specifically demonstrated that such a model could capture the rule-based behavior of past-tense assignment in a system without any rules (Rumelhart & McClelland, 1986). A number of heated exchanges between scientists on both sides of this debate have been published (e.g., Seidenberg, MacDonald, & Saffran [2002] versus Peña, Bonatti, Nespor, & Mehler [2002]; Seidenberg & Elman [1999] versus Marcus et al. [1999]; Keidel, Kluender, Jenison, & Seidenberg [2007] versus Bonatti, Peña, Nespor, & Mehler [2005]) with each side pointing to significant empirical results in support of their position. What is important for our current purpose is the conclusion that there need not be a separable declarative memory system to support grammar processing, as viable non-rule-bound systems have demonstrated that statistical learning over examples held in declarative memory can produce a network with the necessary knowledge held in a distributed representation. Even Ullman (2004) seems to acknowledge the difficulty of distinguishing between the separate declarative and procedural systems he proposes, as he states that the same or similar types of knowledge can in some cases by acquired by both systems. What appears to be more critical, as revealed by the statistical learning approach, and especially studies of language acquisition (e.g., Saffran et al., 1996), is the ability to identify and make use of *cues* in order to learn about the regularities in language. Indeed, a central claim of the connectionist approach is that the cues that facilitate language acquisition in infants become the *constraints* that govern language comprehension in adults (Seidenberg & MacDonald, 1999). As we discuss the memory mechanisms that support comprehension in the sections later, cues will again arise as an important determinant of successful language use.

Working Memory

The construct of WM as a separate store for temporarily held information is an outgrowth of the two-store memory taxonomy, which has been termed the Modal Model (Murdock, 1974) after the statistical term *mode*, because its influence became so pervasive during the last half of the twentieth century. Indeed, even in 2010 it figures prominently in many cognitive and introductory psychology textbooks. This model featured a short-term memory (STM) store characterized by a limited capacity in which verbal information could be held for very short durations, but only if constantly *rehearsed* via active articulation. This is in contrast to the LTM store, which corresponds roughly to the semantic, episodic and procedural memory systems discussed earlier, which is assumed to have an unlimited capacity and duration, so long as appropriate retrieval cues are present to restore passive memories into conscious awareness. The most frequently cited presentation of this model is that of Atkinson and Shiffrin (1968), illustrated in Figure 5-4, which also included a third store for sensory information, subdivided into separate registers for visual, auditory, and haptic information. The modal model emphasized both the qualitative differences between different memory types but also the *processing* mechanisms of each and the way they interact. Inspired by the nascent computer metaphor of the 1950s, this model embodied a specific algorithm through which fleeting sensory information was transformed into a lasting memory. In particular, research demonstrating the highly limited duration of sensory information (1–3 seconds; Sperling, 1960) suggested that it was necessary for information to be verbally *recoded*, and also *rehearsed*, in order to be maintained, and this occurred in the short-term store. Once information had received a sufficient amount of rehearsal in STM, it would move into LTM, where it would reside in a passive state until *retrieved* back into STM where it would be restored into consciousness. Thus, STM is the gateway to and from LTM—any information entering LTM must go through STM and whenever information is retrieved it must again enter STM. (It should be noted that original information is not really transferred, but rather *copied* from one store to another.) At the same time, STM represented a considerable bottleneck for cognitive activity, as it too was found to have a limited storage capacity, made memorable by George Miller's (1956) famous report entitled "The Magical Number Seven, Plus or Minus Two." Miller arrived at this estimate after reviewing data from a number of different paradigms in which individuals were presented with the task of learning new information, only to show highly limited recall on lists containing more than 8 items. Thus, as new information entered STM, some old information becomes lost through displacement—especially information that was not actively rehearsed. Further research revealed that it was possible to expand the capacity of STM via a process called *chunking*, in which meaningful pieces of information are grouped together into a single unit (i.e., the numbers 1, 4, 9, and 2 are remembered as the single unit 1492); however, a limit on the number of chunks that could be actively maintained remains restricted to 3–5 items (cf. Cowan, 2001, for a detailed review).

The centrality of STM motivated the development of models that more precisely articulate how information is brought in and out of consciousness during the performance of cognitive tasks. It is this workspace of active information that has been termed Working Memory—the most influential version of which is the model proposed by Alan Baddeley and colleagues (e.g., Baddeley & Hitch, 1974; reviewed in Baddeley, 2003), depicted in Figure 5-5.[2] The Working Memory model fractionated STM into a set of systems that separately characterized processing and storage; in fact, it was evidence from neuropsychological damage that emphasized the problems with a unitary STM, as patients with severe damage to STM nevertheless retained the ability to access LTM during complex cognitive tasks (Shallice & Warrington, 1970). The key and, ironically, least understood component of the Working Memory model is the Central Executive, which is the controlling mechanism through which information from three subsidiary "slave" storage systems (depicted as gray boxes in Figure 5-5) is brought in and out of the focus of attention (Baddeley, 2003). It is responsible for (at least) updating, shifting, and inhibiting information (Miyake et al., 2000) and has its neurological locus in the frontal lobes, especially dorsolateral prefrontal regions (BA 9/46) and inferior frontal regions (BA 6/44), with some parietal extension into (BA 7/44) (e.g., Braver et al., 1997; Cohen et al., 1997).

The three slave systems can be distinguished by their type of encoding, or the type of information they process. The visuospatial sketchpad is responsible for visuospatial information (e.g., images, spatial configuration, color, shape) and is fractionated into the visual cache (storage) and the inner scribe (rehearsal) components. The more recently postulated episodic buffer (Baddeley, 2000) is responsible for allowing information from LTM to interact with the other two slave systems to create multimodal *chunks* that are open to conscious examination. This buffer should not be confused with episodic memories, discussed earlier, as those are part of LTM while chunks created in Baddeley's episodic buffer are merely temporary associations between different types of information simultaneously manipulated by the central executive. A limit on the amount of information held in this buffer comes from the computational complexity of combining multiple types of codes into a single representation (Hummel, 1999).

---

[2] The Baddeley model is only one of many different formulations of working memory; however, it is the one that has received the most attention. The volume edited by Miyake and Shah (1999) provides a summary of 10 different models of working memory, including several with computational implementations, together with a compare and contrast discussion.

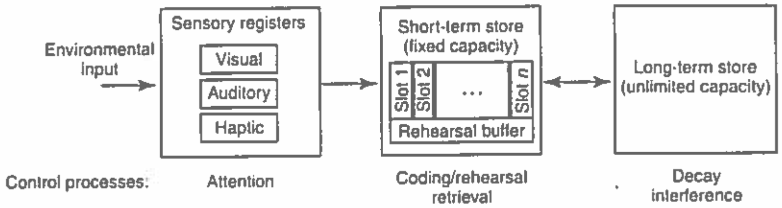

Figure 5-4 Modal Model based on Atkinson and Shiffrin, 1968. Information flow begins with processing information in sensory registers, which have an extremely short duration (<3 seconds). Attentional processes move information from sensory stores into short-term memory, where it is encoded and maintained via rehearsal. Related information may be brought out of the long-term store during encoding. Sufficiently encoded and rehearsed information transfers to long-term store and remains indefinitely, but may become inaccessible due to decay and/or interference. [From Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89–195). New York: Academic Press.]

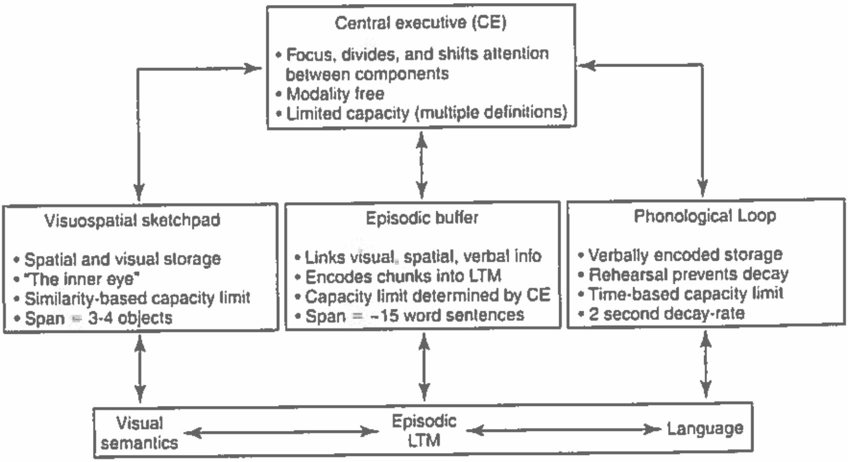

Figure 5-5 Working Memory Model. The model consists of the Central Executive and three "slave" systems, which it directs. The phonological loop is further fractionated into the "articulatory control system" which serves as the "inner voice" and the "phonological store" which serves as the "inner ear". Similarly, the visualspatial sketchpad is fractionated into the "inner scribe" and the "visual cache". The episodic buffer is a recent addition to the model (Baddeley, 2000) and is not as well developed as the other components. [From Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science, 4*, 417–423, and Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*, 829–839.]

The third "slave system," the phonological loop, is the most theoretically developed and experimentally attested. It is responsible for phonological encoding and rehearsal—the means through which verbal information is maintained in an active state. The psychological reality of this process was demonstrated in a number of important early experiments (e.g., Baddeley, 1966; Conrad, 1964; Wickelgren, 1965). For example, Murray (1967) developed a technique to prevent participants from utilizing inner speech to recode information, which became known as *articulatory suppression*. While given a list of words to remember, participants were required to say the word "the" over and over, out loud. When words in the list were similar sounding (i.e., man, mad, cap, can, map) recall errors in the memory condition *without* articulatory supression reflected acoustic confusions: participants were more likely to incorrectly recall items that sounded like the target items but that were not actually in the memory list. With articulatory suppression, on the other hand, acoustic errors were no longer more likely, suggesting that the speaking task prevented participants from recoding, or rehearsing, the memory words using inner speech. These results suggest that not only is information encoded acoustically, but that the amount of information that can be maintained is limited by the ability to actually articulate it—as the number of items to remember increases, some will be forgotten because they cannot be rehearsed. The exact capacity limit for the phonological loop has been quoted as being the amount of information that can be articulated in about 2 seconds (Baddeley, 1986; Baddeley, Thomson, & Buchanan, 1975). Neurologically, lesion studies and neuroimaging methods implicate the left temporoparietal region in the operation of the phonological loop, with BA 40 as the locus of the storage component of the loop and Broca's area (BA 6/44) supporting rehearsal (reviewed in Vallar & Papagno, 2002, and Smith & Jonides, 1997).

WORKING MEMORY AND LANGUAGE COMPREHENSION

The notion that WM capacity is fixed has had a huge influence on theories of language processing. For example, it is a well-replicated finding that sentences in which grammatical heads are separated from their dependents are more difficult to process than when heads and dependents are adjacent (e.g., Grodner & Gibson, 2005; McElree, Foraker, & Dyer, 2003). This is true of unambiguous sentences (e.g., *The book ripped.* versus *The book that the editor admired ripped.*) and of ambiguous sentences (e.g., *The boy understood the man was afraid.* versus *The boy understood the man who was swimming near the dock was afraid.*), where reanalyses prove more difficult as the distance between the ambiguity and the disambiguating material is increased (e.g., Ferreira & Henderson, 1991; Van Dyke & Lewis, 2003). A number of prominent theories have attempted to account for these results by invoking WM capacity, with the common assumption being that capacity is exhausted by the need to simultaneously "hold on to" the unattached constituent (the grammatical subjects *book* and *man* in these examples) while processing the intervening material until the main verb (*ripped* or *was afraid*) occurs. The chief question is taken to be "how much is too much" intervening material before capacity is exhausted; some have suggested that the relevant metric is the number of words (Ferriera & Henderson, 1991; Warner & Glass, 1987) or discourse referents (Gibson, 1998; 2000). Others have focused on the hierarchical nature of dependencies, suggesting that difficulty depends on the number of embeddings (Miller & Chomsky, 1963), or the number of incomplete dependencies (Abney & Johnson, 1991; Gibson, 1998; Kimball, 1973).

This focus on capacity has also spawned a large body of research seeking to demonstrate that sentence comprehension suffers when capacity is reduced either experimentally through the use of dual-task procedures (e.g., Fedorenko, Gibson, & Rohde, 2006, 2007) or clinically, as when poorly performing participants also score poorly on tests of WM capacity, compared with those who do well. For example, King and Just (1991) found that college-level readers with "low" WM capacity showed worse comprehension and slower reading times on syntactically complex sentences than those with "high" or "middle" capacity levels. Similarly, MacDonald, Just, and Carpenter (1992) found that low capacity individuals from the same population had more difficulty interpreting temporarily ambiguous constructions than those with larger capacities. They suggested that this was because a larger WM capacity enabled readers to maintain all possible interpretations for longer, while the smaller capacity readers could only maintain the most likely interpretation. In cases where the ultimately correct interpretation was not the most likely one, low capacity readers would fail to comprehend because the correct interpretation had been "pushed out" of memory.

Studies of reading development also point to an association between low WM capacity and poor comprehension. In a longitudinal study of children with normal word-level (i.e., decoding) skills, Oakhill, Cain, and Bryant (2003) found that WM capacity predicted significant independent variance on standardized measures of reading comprehension at age 7–8 and again 1 year later. Further, Nation, Adams, Bowyer-Crane, and Snowling (1999) found that 10–11 year old poor comprehenders had significantly smaller verbal WM capacity (though not spatial WM capacity) than normal children matched for age, decoding skill, and nonverbal abilities. Likewise, reading disabled children have been found to score in the lowest range on tests of WM capacity (e.g., Gathercole, Alloway, Willis, & Adams, 2006; Swanson & Sachse-Lee, 2001), and these scores are significant predictors of standardized measures of both reading and mathematics attainment.

In all these studies, the standard means of measuring WM capacity is via tests referred to as complex span tasks (e.g., Turner & Engle, 1989; Daneman & Carpenter, 1980).[3] The Reading/Listening Span version of these tasks requires participants to read or listen to an increasingly large group of sentences, and report back only the last words of each sentence in the set. The task of processing the sentence (and in some cases answering questions about it) provides a processing component that, together with the requirement to store the last words, is thought to provide an assessment of the efficiency with which the central executive can allocate resources to both maintain and process linguistic information. Indeed, the task mirrors the functional demand of processing complex linguistic constructions (e.g., long-distance dependencies) mentioned earlier, where substantial information is situated in between two linguistic constituents that must be associated. A meta-analysis of 77 studies found that the Reading Span task predicted language comprehension better than simple span tasks (e.g., digit span) in which participants simply had to remember and report back lists of words (Daneman & Merikle, 1996).

While the impact of the Working Memory model on the study of language processing is undeniable, a close examination of the model reveals that it is not well matched to the functional demands of language comprehension (Lewis, Vasishth, & Van Dyke, 2006). For example, to process the types of sentences discussed earlier (*The book that the editor admired ripped.*), it is argued that the noun phrase *the book* must be held active in WM while the subsequent information is processed, and the difficulty associated with this is what makes the sentence difficult to process. Yet it seems clear that, even when not processing intervening information (*The book ripped.*), there would simply be no time to actively rehearse previously processed constituents during real-time comprehension, where grammatical associations must be made within a few hundred milliseconds (Rayner, 1998). In addition, it seems logical that language comprehension in patients with brain damage should be significantly limited when WM spans are reduced, yet such a relation has failed to materialize, whether span is measured in terms of traditional serial recall measures (Caplan & Hildebrandt, 1988; Martin & Feher, 1990) or in terms of reading span (Caplan & Waters, 1999). Moreover, the emphasis on Reading/Listening span as an index of WM capacity further complicates the issue, as the format of the task in which participants must switch between list maintenance and language comprehension evokes conscious executive processes that are not part of normal comprehension. Consequently, it is unclear whether a participant classified as having a "Low Working Memory Span" actually has a smaller memory capacity, a slower processing speed, difficulty with attention switching, or some combination of these. Next, we discuss further problems with the capacity view itself and then return to the issue of the type of memory model that might better support language processing.

Problems with the Capacity View

Despite its wide acceptance, the empirical support for a separate, fixed-capacity temporary storage system (either STM or WM) is weak. The main evidence in support of separable systems comes from neuropsychological double dissociations, where patients who show severely impaired LTM present with apparently normal STM, and vice versa (e.g., Cave & Squire, 1992; Scoville & Milner, 1957; Shallice & Warrington, 1970). At issue is the role of the medial temporal lobes (MTL) in STM tasks. Recall from our previous discussion that these structures are crucial for the creation and retrieval of long-term declarative memories, so if LTM were entirely distinct from STM, then the prediction is for no MTL involvement in creating STMs or in performing STM tasks. A number of studies have recently cast doubt on whether the double dissociation actually exists, however, showing MTL involvement in short-term tasks (Hannula, Tranel, & Cohen, 2006; Nichols, Kao, Verfaellie, & Gabrieli, 2006; Ranganath & Blumenfeld, 2005; Ranganath & D'Esposito, 2005).

Another source of evidence raising questions about the separability of the two types of memory is data suggesting that representations assumed to be in WM are not retrieved in a qualitatively different manner than those in LTM. Recent fMRI studies indicate that the retrieval of items argued to be within WM span recruit the same brain regions as retrieval from LTM, notably the left inferior frontal gyrus (LIFG) and regions of the medial temporal lobe (MTL) (Öztekin, Davachi, & McElree, 2010; Öztekin, McElree, Staresina, & Davachi, 2008). These imaging results align with behavioral investigations of experimental variables diagnostic of the nature of retrieval process, such as manipulations of recency and the size of the memory set (Box 5-1).

---

[3] One indication of how influential these tests have been is the number of citations they have received. The original Daneman and Carpenter (1980) paper describing the Reading/Listening span task has been cited 1712 times according to ISI Web of Knowledge. The article had 125 citations in 2009 and 74 as of July 2010. The Turner and Engle (1989) paper describing the nonlanguage version of the task (i.e., Operation span) has been cited 501 times since publication: 49 times in 2009 and 29 as of July in 2010. A second indication of their influence is their presence on the Web. A Google search for "individual differences and Sentence Span" received 2.2 million hits (103,000 on Google Scholar) and the same search for "individual differences and operation span" received 485,000 hits (366,000 on Google Scholar) as of July 2010.

Box 5-1

Important Concepts

Connectionism/connectionist models: A computational model of cognition in which knowledge is stored in connections among a set of "nodes" which are assumed to operate like neurons in the brain (i.e., propagating activation to other nodes when they themselves have attained a sufficient level of activation). Knowledge is acquired in these models through a process of supervised learning, wherein the strength of connections between nodes is adjusted over a series of learning episodes. These strengths modulate the rate at which activation is propagated throughout the system, allowing for certain nodes to be "tuned" to particular properties of a stimulus, yielding more activation in certain contexts.

Hierarchical embedding: Grammatical relationships may be either *linear* or *hierarchical*. Hierarchical relationships require retrieval of previous encountered material. For example, in a relative clause, such as *The teacher who gave the difficult test called the principal.* the noun phrase *the teacher* must be retrieved in order to be associated with *called*. In contrast, this retrieval is not required in the following construction, which contains linear relationships: *The teacher gave the difficult test and the teacher called the principal.*

Long-distance/nonadjacent dependency: Refers to grammatical constructions in which two elements that should be associated together are non-adjacent. For example, the simple sentence *The teacher called* becomes a long distance dependency when additional information is inserted between the subject and the verb, as in *The teacher who gave the very difficult test during English class called.* In such a case, the subject *the teacher* would need to be retrieved in order to be associated with the verb *called*. This retrieval would not be necessary in the simple case, when the two words are adjacent. A variety of constructions fall into this category, in addition to the relative clause example discussed earlier, viz.: wh-questions (*Which teacher did you say called our house yesterday?*) where *teacher* is retrieved to be associated with *called*; cleft constructions (*It was the phone that the startled lady realized was ringing.*) where *phone* must be retrieved to be associated with *ringing*; and verb-phrase ellipsis (*The lady heard the phone ring, and the toddler did too.*) where the verb associated with *toddler* has been omitted, and must be retrieved from the previous clause.

Proactive/Retroactive Interference: Two separate types of interference, distinguished by the position of the distracting information vis-à-vis the retrieval target, have been identified. For example in the series $[x_1\ x_2\ x_3\ A\ y_1\ y_2\ y_3\ B]$, if we consider that A is the retrieval target and B is the retrieval cue, then each of the x's create *proactive* interference for retrieving A, while each of the y's create *retroactive* interference for retrieving A. Recent research (Öztekin & McElree, 2007) in the memory domain suggests that proactive interference has its effect primarily on assessments of stimulus familiarity, such as those that yield "know" judgments in the Remember/Know task (cf. Box 3). Recent research in the language domain suggests that retroactive interference is more detrimental than proactive interference for resolving long-distance dependencies (Van Dyke & McElree, in press).

Pronoun resolution: The process of identifying the semantic content of a pronoun by matching it with elements from the previous discourse. For example, if the previous sentence in a text reads *The mother and the baby sat in the waiting room.* a following sentence like *She cried.* has two possible interpretations.

DIAGNOSING THE MECHANISMS OF RETRIEVAL

A number of retrieval mechanisms with quite different computational properties may be available to aid in the recovery of stored information, and empirical research is required to determine if and when each is employed. For example, retrieval may occur through a serial search process in which each item in memory must be checked until the desired item is found (Sternberg, 1966). An alternative process, discussed in the text, is content-addressable retrieval, which operates via direct association between the information available at retrieval-time (cues) and the content of stored memories. An easily understood example of this kind of retrieval is a search in a dictionary for the word "memory": A content-addressable mechanism could go directly to the page containing the words beginning "mem . . . ," while a serial search mechanism would have to begin at "A" and check each item. The chief diagnostic for distinguishing these mechanisms is the retrieval speed for a variety of set sizes and positions in the set. If retrieval occurs via serial search, then the time to access an item will depend on the number of items that must be examined prior to the target. For example, if the dictionary is quite large, then the time for a serial search mechanism to get to the *M*'s will be longer than if the dictionary is abridged. Similarly, a serial search mechanism will take less time to find a word beginning with "D" than it will to find one beginning with "M" because of their respective order in the alphabet. On the other hand,

Box 5-1

Important Concepts—cont'd

if retrieval is direct, then speed will be invariant across all set-sizes or serial positions, assuming the cues available at retrieval-time are sufficient to uniquely identify the target.

A considerable body of research has investigated set-size and serial position effects in the memory domain using a variety of methods (reviewed in McElree, 2006) and there is broad consensus over the conditions requiring direct access versus serial retrieval mechanisms. As reviewed in the text, direct access retrieval occurs when content must be retrieved, however if relational (or order) information is necessary, then serial search processes have been attested (e.g., Gronlund et al., 1997; McElree & Dosher, 1993). Research of this sort in the language domain is more recent, however the evidence points to direct access as the prominent retrieval mechanism. Manipulations that have attempted to duplicate the conditions of set-size by increasing the amount of information, and the amount of interference, between dependencies have consistently found no effects on retrieval speed (e.g., McElree, Foraker, & Dyer, 2003; Van Dyke & McElree, in press).

Contra long-standing claims that information in WM is retrieved with specialized operations (e.g., Sternberg, 1975), the retrieval profiles observed have consistently shown the signature pattern of a direct-access operation, the same type of retrieval operation thought to underlie LTM retrieval. In this type of operation, memory representations are "content-addressable," enabling cues in the retrieval context to make direct contact to representations with overlapping content, without the need to search through irrelevant representations. We take up this discussion further later.

Thus, while it is well documented that our ability to concurrently process different types of information is extremely limited (e.g., Broadbent, 1958), the evidence noted sheds doubt on whether this necessitates the existence of a temporary storage system (be it STM or WM) distinct from LTM. Indeed, there is long strain of research that has challenged the multi-store view, in favor of a unitary-store model, where the information that multi-store models would ascribe to STM/WM is characterized as just the temporarily active portion of LTM (e.g., Anderson et al., 2004; Cowan, 1988, 1995, 2001; Crowder, 1976; McElree, 2001, 2006; Oberauer, 2002; Verhaeghen, Cerella, & Basak, 2004). While these models differ in a variety of details, Cowan's (2001) model can serve as an example (Figure 5-6).

This model suggests that there is only one representation of known information—that in LTM. These representations vary in activation strength, determined by such variables as recency and frequency of occurrence, and representations of increased strength are more available for retrieval when required, but remain in passive memory until such retrievals occur. One type of evidence in support of a unitary-store architecture comes from precise measures of retrieval speed: information in WM should have a privileged status compared to that in LTM, and so should be accessed more quickly. Based on this reasoning, it would be expected to find a "breakpoint" between the speed of accessing the items that have just been processed (i.e., that are in the focus of attention) and then another "breakpoint" between

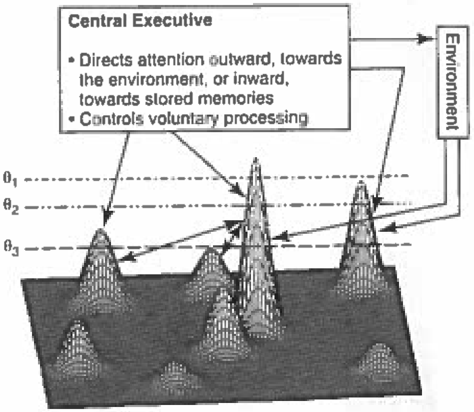

Figure 5-6 Unitary store model. Memories have various levels of activation within the same store. Activation is triggered by cues from the environment or from deliberate attentional processes. Activation may also increase due to associations between items in memory. Dashed lines represent possible threshold levels, θ, by which the size of activated memory would be determined. The most restrictive theories claim that active memory contains only the single item that is in the focus of attention, corresponding to the highest threshold, θ_1; others suppose a lower threshold, such that active memory may contain as many as 4 items.

items that are active in WM and those in LTM. This is not the pattern that has been observed, however. As reviewed in McElree (2006), direct measures of the speed and accuracy of memory retrieval across a broad range of tasks requiring the retention of sequentially presented information have consistently shown that items predicted to be within WM span do *not* exhibit privileged access, but rather are retrieved with the same speed as items well beyond the assumed WM span. These tasks include item recognition, paired-associate recognition, judgments of recency, rhyme and synonym judgments, and the n-back task (Box 5-2). Across all these tasks, there is unequivocal evidence that information being actively processed at test time—typically, the last item studied when there is no distracting activity between study and test—exhibits privileged access, with responses being 30%–50% faster than responses to items outside focal attention.

Box 5-2

Memory Research Methods

RECOGNITION TASKS

Recognition tasks can be distinguished by the overt presence of the very item participants are being asked to remember. As such, participants are generally better at recognition tasks compared to recall tasks (discussed later), making them a more sensitive test of the contents of memory. Participants will often do well in a recognition test even when they fail a recall test for the same item. Specific examples of these tasks are given:

Item recognition: Participants are presented with a memory set, typically containing letters or words, and asked to memorize them during a study phase. At test, participants are presented either with an item which occurred in the memory set, or with an item they had never seen before, and they are required to make a yes/no judgment about whether the item occurred in the memory set. Variable amounts of time may occur between the study and test phases, although studies using this method to examine short-term memory processes typically have the test immediately following the study phase.

Paired-associate recognition: A variation on the item recognition task in which the memory set contains two items (e.g., letters or words) that have been previously studied as a pair. When the pair is presented at test, the participant must verify whether they studied it previously (as a pair). In this paradigm, it may also be of interest to present participants with a pair containing one (or both) item(s) previously studied, but with different partners. This condition forces participants to distinguish specific learning episodes where the two items are paired, from a general feeling of familiarity with the individual members of the pair.

Recency Judgments: A variation on the test phase of the item recognition paradigm in which participants are asked to judge which of two items occurred more recently in the study list. Hence, this task requires participants to remember not only whether particular items were seen previously, but also in what order they occurred.

Rhyme/Synonym Judgments: Variations of the test phase of the item recognition paradigm in which participants are asked to judge whether two items rhyme or mean the same. These tests can be used to force participants to focus on content-related aspects of the studied items, in contrast to simple item recognition tests which may not require deep processing of the studied material.

Remember/Know Judgments: Variation of the test phase of the item recognition paradigm in which participants must indicate whether they actually have a conscious recollection of the item's occurrence in the study list, or whether they have a more diffuse intuition that it was there (they "just know it"). This procedure is useful for distinguishing memories that may be present as the result of explicit retrieval processes versus implicit memories based on quick assessments of familiarity based on sensory or perceptual features of the stimulus.

N-back task: A test of continuous working memory in which participants are presented with a stream of stimulus items and told to indicate (e.g., press a button) when the current stimulus item matches one that appeared n items earlier in the sequence. The variable n represents the load factor, which determines the task difficulty. For example, if n = 2, then participants report on every other stimulus; if n = 4 then participants report on the content of every fourth stimulus, remembering whether it matched the item occurring three trials previous.

Box 5-2

Memory Research Methods—cont'd

RECALL TASKS

Free recall: Participants reproduce material just learned, without any prompts are cues. As the task is quite difficult, it is common to encourage participants to recall as much of the information as possible. Guessing may also be encouraged, as a means of accessing subconscious (implicit) memory traces.

Serial recall: A variation of free recall with the added constraint that participants must recall the information learned in the same order it was learned. The addition of this constraint typically increases difficulty.

Cued recall: A variation of free recall in which partial information is given in order to aid memory. For example, when the memory list contained a list of words, a cued recall test may supply the first letter of each of the memory words. Another frequent example is for a single item of a studied pair to be presented, in order to prompt recall its associate.

IMPLICATIONS FOR LANGUAGE PROCESSING

In principle, language processing might use different memory operations than what has been observed in these basic memory tasks or even use a specialized memory system—for example, Caplan and Waters (1999) suggested that it might draw upon separate WM resources. However, studies investigating the real-time memory operations involved in the processing of linguistic dependencies have yielded results indicating that a dependant constituent is retrieved from memory with the same type of retrieval mechanism described earlier. A range of dependencies have been explored, including verb-argument dependencies (McElree, 2000; McElree, Foraker, & Dyer, 2003), subject-verb dependencies (McElree et al, 2003), verb-phrase ellipsis (Martin & McElree, 2008; 2009), and pronoun resolution (Foraker & McElree, 2007). The crucial studies have used adaptations of the speed-accuracy tradeoff procedure (Dosher, 1979; Wickelgren, 1977; Reed, 1973, 1976) to conjointly measure the speed and accuracy of interpreting an expression with a nonadjacent dependency as a function of 'distance,' viz., the amount of material interpolated between the dependant constituents.

For example, McElree et al. (2003) contrasted the speed of resolving subject-verb dependencies with no material intervening, such as *The editor laughed*, to sentences in which one or two subject- or object relative clauses intervened between the subject and verb. They found that interpretation of the subject-verb dependency occurred at an exceptionally fast rate when the dependent elements were adjacent to one another. However, the speed of accessing a distant noun phrase (NP) to bind as subject to the final verb was constant for each of the nonadjacent constructions, which contain varying numbers of intervening words, discourse items, and hierarchically embedded constituents. These results mirror those found in basic memory studies in two key respects. First, there was a "breakpoint" in processing speed for the most recent item processed and all other items, marking the distinction between items being actively processed and those that require retrieval to be restored to active processing. Second, retrieval speed was invariant across linear distance, as well as other types of metrics such as level of embedding or the number of incomplete dependencies. This is the signature pattern of a direct-access operation, in which associative retrieval cues provide direct access to the content of stored representations. It is not the pattern expected if retrieval required a search (either forward or backward) through the hierarchical parse-tree in a step-by-step fashion in order to identify the correct grammatical dependent (McElree, 2006).

At first blush, it might appear that a processing architecture eschewing a traditional 3–4 item WM storage buffer may be too restrictive to subserve sentence processing. However, Lewis, Vasishth, and Van Dyke (2006; see also Lewis & Vasishth, 2005) described a computational model of sentence processing that requires maintaining only the most recently parsed item in active memory. The model's memory consists of chunks representing the syntactic structure built so far, together with predictions for constituents licensed by the current state of the parse. These chunks are not actively held in memory and decay as a function of time and prior retrievals. The only access to these items is via a retrieval buffer with the capacity to hold a single chunk. This affords the model the minimum capacity required to create new linguistic relations—the item waiting to be integrated into the parse, and the chunk that licenses it. The item that is waiting is in the focus of attention and does not need to be retrieved. The chunk that licenses it is retrieved via the cues derived from the features of the waiting item. Critically, it is this cue-based retrieval process, which occurs via direct access, that provides the computational power necessary to create dependencies in real time. Mathematical analyses of reaction time distributions (Ratcliff, 1978) and evidence from the Speed-Accuracy Tradeoff (SAT) paradigm (McElree, 2001) suggest that humans can restore items into active memory in approximately 80–90 ms. Retrieval speeds that are this fast enable the parsing mechanism to compensate for the severe limit on the size of active memory, while still enabling parsing decisions to be made in about 200 ms, which is typical for real-time language processing.

Forgetting

Lost memory is perhaps the most vexing problem human beings face. Even in a nonclinical setting, the phenomenon is a constant reminder that even our highly evolved brains have inescapable limits. We noted earlier in our discussion of episodic memory that the primary account of forgetting long-held information relates to an inability to retrieve the information—that is, the information becomes *inaccessible*, but nevertheless remains in memory and can be reactivated if only suitable retrieval cues are supplied (Tulving, 1979). The usefulness of reminders and mnemonic devices seems to fit naturally with this account, and gives intuitional support to the body of evidence that weighs against other explanations based on failure to store memories in the first place (e.g., Crowder, 1982; Keppel, 1984; Quartermain, McEwen, & Azmitia, 1972).

From the perspective of language processing, the crucial question is what causes forgetting over the short term, since we are interested in the processing that occurs over the span of a paragraph, or even a sentence. This has been a question of great debate in the memory literature, centering around the role of *decay* (e.g., Nairne, 2002; Lewandowsky, Duncan, & Brown, 2004). As discussed earlier, limited-capacity multistore models have traditionally favored decay, or displacement, as the mechanism that controls forgetting: any information that is not maintained via some mechanism of active maintenance (e.g., rehearsal) will be lost. Alternative unitary-store models relinquish a separate maintenance mechanism for a fast cue-based retrieval mechanism that can restore information into active memory as needed. From this perspective, information is lost because retrieval cues are insufficient to uniquely identify the necessary information. This occurs when the presence of similar items in memory creates a condition of *cue-overload*, where retrieval cues are associated with multiple items in memory, making them inadequate discriminators (e.g., Öztekin & McElree, 2007; Nairne, 2002; Watkins & Watkins, 1975). The result is *interference*, where unwanted items are retrieved instead of the target item. Interference can come in two varieties: the case where similar items *precede* the target, creating *proactive interference*, and the case were similar items *follow* the target, creating *retroactive interference*.

Despite its popularity as a component of multistore models, and its intuitive appeal, the evidence supporting decay is weak. Even from the early days of memory theorizing, decay came under fire as a logically inadequate explanation: John McGeoch (1932) pointed out that just as iron rusts over time, memories are forgotten over time, but in neither case is time the causal agent. While oxidation is the mechanism through which rust forms, similarly, the mechanism through which forgetting arises must be stated. McGeoch proposed that interference is the most likely candidate.

One of the chief problems with evaluating the decay hypothesis is that it is nearly impossible to rule out interference as an alternative explanation. For example, the classic Brown-Peterson studies (Brown, 1958; Peterson & Peterson, 1959) used articulatory suppression to block rehearsal during a memory task where participants were supposed to remember a 3-consonant trigram (TWF). They found that correct recall was reduced as the length of the suppression task increased from 3 to 18 seconds (increasing the delay between study and test), until only about 10% of studied trigrams could be recalled. The apparent conclusion is that without the ability to rehearse, information will be almost completely lost within about 18 seconds. However, two follow-up studies make it clear that this conclusion is incorrect. Waugh & Norman (1965) varied the presentation rate for a study list of 16 digits, so that in the fast condition only 4 seconds passed during which decay could occur, while in the slow conditions (digits presented 1 per second) 16 seconds passed. Following the study list, a target digit was presented, and participants were asked to recall the digit that followed this target. Contrary to the prediction that a longer amount of time should produce more forgetting, they found no difference between the two presentation rate conditions. A further challenge to the decay account of the Brown-Peterson studies came from Keppel and Underwood (1962), who conducted a modified analysis of data produced via the Brown-Peterson method. The prediction of an interference account is that trials from the beginning of the experiment should be more easily recalled than those from the later part of the experiment because earlier trials will have less prior buildup of interfering material (i.e., less proactive interference).

By analyzing individual trials—something that was not done in the original studies—they found exactly this; accuracy on trial 1 was nearly 100% even after 18 minutes of delay time, but begins to reduce after this with each trial getting successively worse as more information in memory builds up. This finding directly contradicts the original study, suggesting that those results were obtained only because the experimenters aggregated data over individual trials, causing them to miss observing the buildup of proactive interference.

While it has proved difficult to disentangle decay and interference in memory studies, Van Dyke and Lewis (2003) presented data suggesting that both are at work in the domain of language processing. They manipulated the distance between two grammatically dependent constituents (man and was paranoid in this example) by comparing a sentence with no intervening distance (*1. The frightened boy understood that the man was paranoid about dying.*) with a sentence with an intervening clause (*2. The frightened boy understood that the man who was swimming near the dock was paranoid about dying.*) They also manipulated the amount of interference present in the intervening region by comparing the long sentence in (2) with the sentence (*3. The frightened boy understood that the man who said the townspeople were dangerous was paranoid about dying.*) The amount of interference is measured with respect to the retrieval cues set by the verb phrase *was paranoid*. This verb phrase is assumed to contain retrieval cues that will identify a grammatical subject with which it can be associated so that it can be integrated into a coherent interpretation of the sentence. Thus, sentence (3) is considered to have more interference than sentence (2) because the intervening noun phrase *the townspeople* shares its grammatical encoding with the target constituent; they are both grammatical subjects. The retrieval cues from the verb will therefore match to both *the townspeople* and *the man* as potential subjects. In contrast, sentence (2) is a low interference condition because it does not have a subject intervening between the verb phrase and the target noun phrase; the intervening noun phrase *the dock* is the object of a prepositional phrase. Note that sentences (2) and (3) are matched on the distance dimension; both have 6 intervening words. Thus, the contrast between (1) and (2) provides an estimate of the distance effect, while the contrast between (2) and (3) provides an estimate of the additional interference effect. The left panel of Figure 5-7 shows the results for acceptability judgments; an identical pattern of results was found for reading times on the verb phrase itself. Interference (2 versus 3) had a significant effect, but distance (1 versus 2) did not. This is consistent with the view that the critical factor for making constituents unavailable for retrieval is not the *amount* of information, but rather *how similar* the intervening information is to the target.

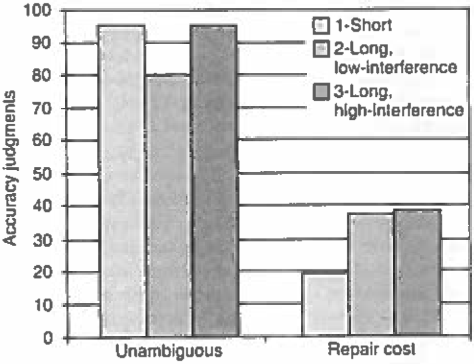

Figure 5-7 Participants had to judge whether sentences were grammatical. Repair cost was calculated by subtracting accuracy scores to unambiguous sentences from the ambiguous version of the same sentence. The interference manipulation (2 versus 3) had an effect on procedures associated with integrating the separated subject and verb, but this did not affect repair cost. Distance (1 versus 2), on the other hand, *did* affect repair cost. [Results from Van Dyke, J. A., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A retrieval interference theory of recovery from misanalyzed ambiguities. *Journal of Memory and Language, 49,* 285–413, Experiment 3.]

A further manipulation of ambiguity, created by removing the *that* in the earlier conditions, enabled Van Dyke and Lewis to investigate effects of decay because the less preferred interpretation is not pursued. For example, in the ambiguous version of (1), given here as (*4. The frightened boy understood the man was paranoid about dying.*) the verb *understood* can be interpreted either as a verb that takes a direct object (cf. *The boy understood the question and answered it*) or as a verb that takes a sentential complement (cf. *The boy understood the question was difficult*). Van Dyke and Lewis designed the experiment to include a large number of direct object sentences as filler items, so as to strongly bias the reader toward taking the direct object interpretation initially. The assumption was that the ultimately correct sentential complement interpretation of *understood* would not be pursued, causing the syntactic features licensing the sentential complement to decay because of disuse. Thus, in the ambiguous version of

(2), given here as (5. *The frightened boy understood the man who was swimming near the dock was paranoid about dying*), the initial interpretation could be consistent with the sentence *The boy understood the man who was swimming near the dock and smiled at him*.[4] A similar relationship between understood and the man would be adopted prior to the occurrence of was paranoid for the ambiguous version of the high interference sentence (3), given here as (6. *The frightened boy understood the man who said the townspeople were dangerous was paranoid about dying*.) Crucially, at the point when *was paranoid* must be processed, the sentential complement features must be reactivated in order to integrate the verb phrase into the sentence. The prediction was that any difficulty in reactivating the sentential complement features arises as a result of how much these features decayed while the incorrect interpretation was pursued. Consistent with this view, the distance effect on the ability reanalyze the ambiguous sentence was significant (cf. right panel of Figure 5-7), suggesting the decay of the less preferred interpretation. There was no additional effect of interference during reanalysis, however, consistent with the fact that the interfering material in the unambiguous sentences is identical to that in the ambiguous sentences.

These results have strong implications for the type of memory system thought to underlie sentence comprehension. As discussed earlier, the dominant capacity approach has suggested that sentences such as (2) and (3) are difficult to process because *the man* must be "held" in WM while processing the intervening material, which expends memory resources because of its length (e.g., Gibson, 1998, 2000). Contra this, Van Dyke and Lewis found that only particular types of intervening constructions—those containing syntactically similar material—produced difficulty. Thus, the distance effects that were previously thought to occur because of decay, or because of a lack of memory resources, can be attributed to retrieval interference. Decay, on the other hand, seems to have its effect only on the ability to re-retrieve information after it has been completely abandoned—and notably, without any new retrieval cues that would guide the retrieval mechanism in doing so.

While the Van Dyke and Lewis study investigated interference arising from syntactically similar distractors, other types of interference effects have also been observed in sentence comprehension. For example, in an extension of the study just described, Van Dyke (2007) showed that interference could arise from *semantically* similar distractors, even when not in a syntactically similar position. Thus, (2) was easier than the same sentence with the word *dock* replaced with the word *girl*, which fits the semantic cues of the verb phrase (i.e., a *girl* can be paranoid but a *dock* cannot). Still another type of interference—referential interference—was observed by Gordon and colleagues (Gordon, Hendrick, & Johnson, 2001; 2004) who investigated the role of various noun phrase types appearing as the second (underlined) noun in subject relative clauses (e.g., The banker that praised the barber climbed the mountain) and object relative clauses (e.g., The banker that the barber praised climbed the mountain). The greater difficulty of the object relative as compared with the subject relative construction has been repeatedly documented (e.g., King & Just, 1991; Staub, 2010; Traxler, Morris, & Seely, 2002), with the dominant explanation focusing on different demands each construction makes on memory. Gordon et al. sought to pinpoint the contribution of interference to this contrast by manipulating the referential status of the second noun phrase. In several experiments, they contrasted the sentences earlier with identical sentences except for substituting a pronoun (*you* or *everyone*) or a proper name (*Joe*) for barber, and found that the advantage for subject-relative clauses over object-relative clauses was reduced or eliminated. Common nouns like *barber* and *banker* refer indirectly by virtue of their description, while pronouns and proper names refer directly, singling out specific entities in the current discourse context. Thus, similarity-based interference arises in a variety of linguistic contexts in the presence of syntactic, semantic, and referentially similar distractors.

The appearance of interference effects—a classic memory phenomenon—in language comprehension weighs against the proposal of a language-specific memory capacity (Caplan & Waters, 1999) and points to a unification of memory mechanisms operating both over the short- and long-term temporal periods, and both in the memory and language domains. Support for this parsimonious approach is apparent in neuroimaging research that has attempted to identify the brain regions responsible for memory retrieval. We have already noted fMRI evidence suggesting that retrieval of recent items recruits the same brain regions as retrieval from LTM, notably the left inferior frontal gyrus (LIFG) (Öztekin et al., 2008, 2010). The LIFG has also been repeatedly implicated in neuroimaging studies of memory interference resolution (reviewed in Jonides & Nee, 2006).

---

[4] The continuation *and smiled at him* is included here only to emphasize that *the man* is interpreted as the direct object of *understood*. The experiment did not include continuations such as these.

Additionally, patient work (e.g., Thompson-Schill et al., 2002) and repetitive transcranial magnetic stimulation investigations (e.g., Feredoes, Tononi, & Postle, 2006) have provided converging evidence for a direct role of LIFG in successful interference resolution. This is exciting because this same region, which includes Broca's area, has also had a long history of being associated with language (especially syntactic) processing (Rogalsky & Hickock, 2010, for a review). In particular, the subregions of BA 44 and 45 in LIFG have been been repeatedly implicated in the processing of syntactically interfering sentence constructions as in (3) earlier (e.g., Cooke et al., 2001; Fiebach, Vos, & Friederici, 2004; Makuuchi, Bahlmann, Anwander, & Friederici, 2009; Stowe et al., 1999). Likewise, a recent fMRI study following on Van Dyke (2007) has found semantic interference effects in the pars triangularis region of BA 45 (Guo, Martin, Van Dyke, & Hamilton, 2010). Recent attempts to further specify the functional role of the subregions of LIFG during memory retrieval comport well with the language processing results, as they point to a unique role of the pars triangularis region (BA 45) in tasks requiring selection among competing alternatives (Badre & Wagner, 2007; Badre, Poldrack, Paré-Blagoev, et al., 2005). Taken together, these separate streams of research in the memory and language domains appear to converge on the idea that the ability to manage retrieval interference may be at the root of memory and language deficits in both clinical and nonclinical populations. Indeed, a number of researchers have already suggested that differences in susceptibility to interference (Hasher & Zacks, 1988; Stoltzfus, Hasher, & Zacks, 1996) provide more veridical characterizations of age-related changes and individual differences in memory ability. Current approaches to language deficits in clinical populations have also moved toward explanations that implicate interference. For example, comprehension deficits in patients with Parkinson's disease have been linked to deficits in cognitive flexibility and the ability to inhibit irrelevant information (Hochstadt, Nakano, Lieberman, & Friedman, 2006).

CONCLUSIONS

A long history of research in neuropsychology, psycholinguistics, and cognitive psychology has attempted to characterize the relationship between memory and language. While advances have been made, a number of stumbling blocks have been encountered due to the adoption of memory models that were developed to account for memory phenomenon unrelated to the task of language processing. The advent of connectionism and statistical learning theory has led to a number of important advances, but there is still much to understand about how the two systems interact. This review has emphasized areas where it would be fruitful to examine the extent to which the memory system and the language processing system rely on (at least functionally) the same mechanisms. Namely, a growing body of evidence now suggests that language processing is supported by a memory architecture that emphasizes a unitary store and a fast cue-based retrieval mechanism, which is susceptible to retrieval interference (e.g., Lewis, Vasishth, & Van Dyke, 2006). The central issue in determining how clinical variables resulting from brain damage and aging affect this system will be to develop a further understanding into the mechanisms necessary for identifying and using cues, both as a means through which new linguistic knowledge is learned and as the engine that drives comprehension. Although this issue is understudied at present, it has gained increased attention in recent years. The data available so far suggest that individuals do vary in their capacity for statistical learning and that these differences are correlated with differences in language and reading performance (Ahissar et al., 2006, Ahissar, 2007; Conway, Bauernschmidt, Huang, & Pisoni, in press; Evans, Saffran, & Robe-Torres, 2009). Additional research into the neural basis for cue-based learning and retrieval will be important in order to gain a more complete understanding of the interaction of memory and language processes.

ACKNOWLEDGMENTS

Preparation of this chapter was supported by NIH/NICHD grant R21-HD-058944 to Haskins Laboratories (Van Dyke, PI), NIH/NICHD grant R01-HD-040353 to Haskins Laboratories (Shankweiler, PI), and by NIH/NICHD grant R01-HD-056200 to New York University (McElree, PI).

REFERENCES

Abney, S. P., & Johnson, M. (1991). Memory requirements and local ambiguities of parsing strategies. *Journal of Psycholinguistic Research, 20*, 233–250.

Ahissar, M. (2007). Dyslexia and the anchoring-deficit hypothesis. *Trends in Cognitive Sciences, 11*(11), 458–465.

Ahissar, M., Lubin, Y., Putter-Katz, H., & Banai, K. (2006). Dyslexia and the failure to form a perceptual anchor. *Nature Neuroscience, 9*, 1558–1564.

Aldridge, J. W., & Berridge, K. C. (1998). Coding of serial order by neostriatal neurons: A "natural action" approach to movement sequence. *Journal of Neuroscience, 18*(7), 2777–2787.

Alvarez, P., & Squire, L. R. (1994). Memory consolidation and the medial temporal lobe: A simple network model. *Proceedings of the National Academy of Sciences USA, 91*, 7041–7045.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of mind. *Psychological Review, 111*, 1036–1060.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89–195). New York: Academic Press.

Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., & Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology, 16*(18), 1818–1823.

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. (1986). *Working memory*. Oxford, UK: Clarendon Press.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science, 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.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*, 575–589.

Badre, D., Poldrack, R. A., Paré-Blagoev, E. J., Insler, R., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. Neuron, 47, 907–918.

Badre D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the control of memory. *Neuropsychologia, 45*, 2883–2901.

Bhatt, R. S., Wilk, A., Hill, D., & Rovee-Collier, C. (2004). Correlated attributes and categorization in the first half-year of life. *Developmental Psychobiology, 44*, 103–115.

Boecker, H., Dagher, A., Ceballos-Baumann, A. O., Passingham, R. E., Samuel, M., Friston, K. J., Poline, J., . . . Brooks, D. J. (1998). Role of the human rostral supplementary motor area and the basal ganglia in motor sequence control: Investigations with H2 150 PET. *Journal of Neurophysiology, 79*(2), 1070–1080.

Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience, 25*, 151–188.

Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations: The role of consonants and vowels in continuous speech processing. *Psychological Science, 16*, 451–459.

Bowers, J. S., & Schacter, D. L. (1992). Priming of novel information in amnesia: Issues and data. In P. Graf & M. E. J. Masson (Eds.), *Implicit memory: New directions in cognition, neuropsychology, and development*. New York: Academic Press.

Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., & Noll, D. C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *NeuroImage, 5*, 49–62.

Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon Press.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology, 10*, 12–21.

Buckner, R. L., & Koutstaal, W. (1998). Functional neuroimaging studies of encoding, priming, and explicit memory retrieval. *Proceedings of the National Academy of Sciences USA, 95*, 891–898.

Buckner, R. L., & Wheeler, M. E. (2001). The cognitive neuroscience of remembering. *Nature Reviews Neuroscience, 2*(9), 624–634.

Buffalo, E. A., Reber, P. J., & Squire, L. R. (1998). The human perirhinal cortex and recognition memory. *Hippocampus, 8*(4), 330–339.

Caplan, D., & Hildebrandt, N. (1988). *Disorders of syntactic comprehension*. Cambridge, MA: MIT Press/Bradford Books.

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

Cave, C. B., & Squire, L. R. (1992). Intact and long-lasting repetition priming in amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(3), 509–520.

Cermak, L. S., Verfaellie, M., Milberg, W., Letourneau, L., & Blackford, S. (1991). A further analysis of perceptual identification priming in alcoholic Korsakoff patients, *Neuropsychologia, 29*, 725–736.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience, 2*(10), 913–919.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York: Praeger.

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., & Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature, 386*, 604–608.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*(2), 240–247.

Conrad, R. (1964). Acoustic confusion in immediate memory. *British Journal of Psychology, 55*, 75–84.

Conway, C., & Christiansen, M. (2001). Sequential learning in non-human primates. *Trends in Cognitive Sciences, 5*(12), 539–546.

Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition, 114*(3), 356–371.

Conway, M. A. (2001). Sensory-perceptual episodic memory and its context: Autobiographical memory. *Philosophical Transactions of the Royal Society of London, B356*, 1375–1384.

Cooke, A., Zurif, E. B., DeVita, C., Alsop, D., Koenig, P., Detre, J., Gee, J., . . . Grossman, M. (2001). Neural basis for sentence comprehension: Grammatical and short-term memory components. *Human Brain Mapping, 15*(2), 80–94.

Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., & Petersen, S. E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science, 248*, 1556–1559.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin, 104*, 163–191.

Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford, UK: Oxford University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87–185.

Crain, S., & Thornton, R. (1998). *Investigations in universal grammar*. Cambridge, MA: MIT Press.

Crowder, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.

Crowder, R. G. (1982). General forgetting theory and the locus of amnesia. In L. S. Cermak (Ed.), *Human memory and amnesia*. Hillsdale, NJ: Erlbaum.

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

Daneman, M. E., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*(4), 422–433.

Danker, J. F., & Anderson, J. R. (2010). The ghosts of brain states past: Remembering reactivates the brain regions engaged during encoding. *Psychological Bulletin, 136*(1), 87–102.

Delgado, M. R., Jou, R. L., LeDoux, J. E., & Phelps, E. A. (2009). Avoiding negative outcomes: Tracking the mechanisms of avoidance learning in humans during fear conditioning. *Frontiers in Behavioral Neuroscience, 3*, 1–9. doi:10.3389/neuro.08.033.2009

Dominey, P. F., Hoen, M., Blanc, J. M., & Lelekov-Boissard, T. (2003). Neurological basis of language and sequential cognition: Evidence from simulation, aphasia, and ERP studies. *Brain and Language, 83*, 207–225.

Dosher, B. A. (1979). Empirical approaches to information processing: Speed-accuracy tradeoff or reaction time. *Acta Psychologica, 43*, 347–359.

Doyon, J., Gaudreau, D., Laforce, R., Castonguay, M., Bedard, P. J., Bedard, F., & Bouchard, J. P. (1997). Role of the striatum, cerebellum, and frontal lobes in the learning of a visuomotor sequence. *Brain and Cognition, 34*(2), 218–245.

Doyon, J., Owen, A. M., Petrides, M., Sziklas, V., & Evans, A. C. (1996). Functional anatomy of visuomotor skill learning in human subjects examined with positron emission tomography. *European Journal of Neuroscience, 8*(4), 637–648.

Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 52, 321–335.

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.

Fedorenko, E., Gibson, E., & Rohde, D. (2007). The nature of working memory in linguistic, arithmetic and spatial integration processes. *Journal of Memory and Language, 56*(2) 246–269.

Feredoes, E., Tononi, G., and Postle, B. R. (2006). Direct evidence for a prefrontal contribution to the control of proactive interference in verbal working memory. *Proceedings of the National Academy of Sciences USA, 103*(51), 19530–19534.

Ferreira, F., & Henderson, J. M. (1991). Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language, 30*(6), 725–745.

Fiebach, C. 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.

Foraker, S., & McElree, B. (2007). The role of prominence in pronoun resolution: Availability versus accessibility. *Journal of Memory and Language, 56*(3), 357–383.

Funnell, E. (1983). Phonological processes in reading: New evidence from acquired dyslexia. *British Journal of Psychology, 74*, 159–180.

Gabrieli, J. D. E., & Keane, M. M. (1988). Priming in the patient H. M.: New findings and a theory of intact and impaired priming in patients with memory disorders. *Society for Neuroscience Abstracts, 14*, 1290.

Gabrieli, J. D. E., Poldrack, R. A., & Desmond, J. E. (1998). The role of left prefrontal cortex in language and memory. *Proceedings of the National Academy of Sciences USA, 95*, 906–913.

Gathercole, S. E., Alloway, T. P., Willis, C., & Adams, A. M. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology, 93*, 265–281.

Gerken, L., Wilson, R., & Lewis, W. (2005). Infant can use distributional cues to form syntactic categories. *Journal of Child Language, 32*, 249–268.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz (Ed.), *Image, language, brain: Papers from the first mind articulation project symposium* (pp. 94–126). Cambridge, MA: MIT Press.

Gibson, E. A. (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, 51*, 97–114.

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science, 13*, 425–430.

Goschke, T., Friederici, A., Kotz, S. A., & van Kampen, A. (2001). Procedural learning in Broca's aphasia: Dissociation between the implicit acquisition of spatio-motor and phoneme sequences. *Journal of Cognitive Neuroscience, 13*(3), 370–388.

Graybiel, A. M. (1995). Building action repertoires: Memory and learning functions of the basal ganglia. *Current Opinion in Neurobiology, 5*, 733–741.

Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input. *Cognitive Science, 29*, 261–290.

Gronlund, S. D., Edwards, M. B., & Ohrt, D. D. (1997). Comparison of retrieval of item versus spatial position information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1261–1274.

Guo, Y., Martin, R., Van Dyke, J., & Hamilton, C. (2010). Interference effects in sentence comprehension: An fMRI study. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Haist, F., Musen, G., & Squire, L. R. (1991). Intact priming of words and nonwords in amnesia. *Psychobiology, 19*, 275–285.

Hannula, D. E., Tranel, D., & Cohen, N. J. (2006). The long and the short of it: Relational memory impairments in amnesia, even at short lags. *The Journal of Neuroscience, 26*(32), 8352–8359.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111*, 662–720.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193–225). New York: Academic Press.

Hauk, O., Johnsrude, I., & Pulvermuller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*, 301–307.

Hochstadt, J., Nakano, H., Lieberman, P., & Friedman, J. (2006). The roles of sequencing and verbal working memory in sentence comprehension deficits in Parkinson's disease. *Brain & Language, 97*, 243–257

Howard, D., & Patterson, K. (1992). *Pyramids and palm trees: A test of semantic access from pictures and words.* Thames Valley: Bury St Edmunds.

Hummel, J. (1999). The binding problem. In R. A. Wilson & F. C. Keil (Eds.), *The MIT encyclopedia of cognitive sciences* (pp. 85–86). Cambridge, MA: MIT Press.

James, W. (1890). *The principles of psychology.* New York: Henry Holt and Company.

Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience, 139*, 181–193.

Katz, L., Lee, C. H., Tabor, W., Frost, S. J., Mencl, W. E., Sandak, R., Rueckl, J., & Pugh, K. R. (2005). Behavioral and neurobiological effects of printed word repetition in lexical decision and naming. *Neuropsychologia, 43*, 2068–2083.

Keidel, J., Kluender, K., Jenison, R., & Seidenberg, M. (2007). Does grammar constrain statistical learning? *Psychological Science, 18*, 922–923.

Kellenbach, M. L., Brett, M., & Patterson, K. (2001). Larger, colorful, or noisy? Attribute- and modality-specific activations during retrieval of perceptual attribute knowledge. *Cognitive and Affective Behavioral Neuroscience, 1*, 207–221.

Kelley, W. M., Miezin, F. M., McDermott, K. B., Buckner, R. L., Raichle, M. E., Cohen, N. J., Ollinger, J. M., . . . Petersen, S. E. (1998). Hemispheric specialization in human dorsal frontal cortex. *Neuron, 20*, 927–936.

Keppel, G. (1984). Consolidation and forgetting theory. In H. Weingartner & E. S. Parker (Eds.), *Memory consolidation: Psychobiology of cognition.* Hillsdale, NJ: Erlbaum.

Keppel, G., & Underwood, B. J. (1962). Proactive-inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior, 1*, 153–161.

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition, 2*, 15–47.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30*(5), 580–602.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273*, 1399–1402.

Levy, D. A., Manns, J. R., Hopkins, R. O., Gold, J. J., Broadbent, N. J., & Squire, L. R. (2003). Impaired visual and odor recognition memory span in patients with hippocampal lesions. *Learning and Memory, 10*, 531–536.

Lewandowsky, S., Duncan, M., & Brown, G. D. A. (2004). Time does not cause forgetting in short-term serial recall. *Psychonomics Bulletin Review, 11*, 771–790.

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 Science, 10*(10), 447–454.

MacDonald, M. C., Just, M. A., & Carpenter, P. C. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology, 24*, 56–98.

Maine de Biran, F. P. G. (1929/1804). *The influence of habit on the faculty of thinking.* Baltimore: Williams & Wilkins.

Makuuchi, M., Bahlmann, J., Anwander, A., Friederici, A. D. (2009). Segregating the core computational faculty of human language from working memory. *Proceedings of the National Academy of Sciences USA, 106*(20), 8362–8367.

Marcus, G. F. (2001). *The algebraic mind: Integrating connectionism and cognitive science.* Cambridge, MA: MIT Press.

Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., & Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology, 29*, 189–256.

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science, 283*, 77–80.

Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: Structure and processes. *Current Opinion in Neurobiology, 11*, 194–201.

Martin, A., Haxby, J. V., Lalonde, F. M., Wiggs, C. L., & Ungerleider, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science, 270*, 102–105.

Martin, A., & McElree, B. (2008). A content-addressable pointer mechanism underlies comprehension of verb-phrase ellipsis. *Journal of Memory and Language, 58*(3), 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 & Cognition*, *35*, 1231–1239.

Martin, R. C., & Feher, E. (1990). The consequences of reduced memory span for the comprehension of semantic versus syntactic information. *Brain and Language*, *38*, 1–20.

McCandliss, B., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, *7(7)*, 293–299.

McDermott, K. B., Buckner, R. L., Petersen, S. E., Kelley, W. M., & Sanders, A. L. (1999). Set- and code-specific activation in frontal cortex: an fMRI study of encoding and retrieval of faces and words. *Journal of Cognitive Neuroscience*, *11*, 631–640.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–457.

McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, *4*, 1–13.

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

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

McGeoch, J. (1932). Forgetting and the law of disuse. *Psychological Review*, 39, 352–370.

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 D. R. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. II). New York: John Wiley.

Milner, B. (1962). Les troubles de la memoire accompagnant des lesions hippocampiques bilaterales. In *Physiologie de l'hippocampe* (pp. 257–272). Paris: Centre National de la Recherche Scientifique. English translation: B. Milner & S. Glickman (Eds.). Princeton: Van Nostrand, 1965 (pp. 97–111).

Milner, B. (1972). Disorders of learning and memory after temporal lobe lesions in man. *Clinical Neurosurgery*, *19*, 421–466.

Mirković, J., MacDonald, M. C., & Seidenberg, M. S. (2005). Where does gender come from? Evidence from a complex inflectional system. *Language and Cognitive Processes*, *20*, 139–168.

Mishkin, M., Malamut, B., Bachevalier, J. (1984). Memories and habits: Two neural systems. In G. Lynch, J. L. McGaugh, & N. W. Weinburger (Eds.), *Neurobiology of learning and memory* (pp. 65–77). New York: Guilford Press.

Miyake, A., Friedman, N., Emerson, M., Witzki, A., Howerter, A., & Wager, T. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.

Miyake, A., & Shah, P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge University Press: Cambridge.

Mummery, C. J., Patterson, K., Hodges, J. R., & Price, C. J. (1998). Functional neuroanatomy of the semantic system: divisible by what? *Journal of Cognitive Neuroscience*, *10*, *766–777*.

Murdock, B. B., Jr. (1974). *Human memory: Theory and data*. Hillsdale, NJ: Erlbaum.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Murray, D. J. (1967). The role of speech responses in short-term memory. *Canadian Journal of Psychology*, *21*, 263–276.

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology*, *53*, 53–81.

Nation, K., Adams, J. W., Bowyer-Crane, C. A., & Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology*, *73*, 139–158.

Nichols, E. A., Kao, Y-C., Verfaellie, M., & Gabrieli, J. D. E. (2006). Working memory and long-term memory for faces: Evidence from fMRI and global amnesia for involvement of the medial temporal lobes. *Hippocampus*, *16*, 604–616.

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes*, *18*(4), 443–468.

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

Otten, L. J., & Rugg, M. D. (2001). Task-dependency of the neural correlates of episodic encoding as measured by fMRI. *Cerebral Cortex*, *11*(12), 1150–1160.

Öztekin, I., Davachi, L., & McElree, B. (2010). Are representations in working memory distinct from those in long-term memory? Neural evidence in support of a single store. *Psychological Science*, *21*(8), 1123–1133.

Öztekin, I., & McElree, B. (2007). Retrieval dynamics of proactive interference: PI slows retrieval by eliminating fast assessments of familiarity. *Journal of Memory and Language*, *57*, 126–149.

Öztekin, I., McElree, B., Staresina, B. P., & Davachi, L. (2008). Working memory retrieval: Contributions of left prefrontal cortex, left posterior parietal cortex and hippocampus. *Journal of Cognitive Neuroscience*, 21, 581–593.

Packard, M. G., Hirsh, R., & White, N. M. (1989). Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: Evidence for multiple memory systems. *Journal of Neuroscience*, *9*, 1465–1472.

Packard, M., & Knowlton, B. (2002). Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience, 25*, 563–593.

Parkin, A. J. (2001). The structure and mechanisms of memory. In B. Rapp (Ed.), *The handbook of cognitive neuropsychology: What deficits reveal about the human mind*. (pp. 399–422). Philadelphia: Psychology Press.

Peigneux, P., Maquet, P., Meulemans, T., Destrebecqz, A., Laureys, S., Degueldre, C., Delfiore, G., . . . Cleeremans, A. (2000). Striatum forever, despite sequence learning variability: A random effect analysis of PET data. *Human Brain Mapping, 10*, (4), 179–194.

Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science, 298*, 604–607.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*:357–383.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences, 10*, 233–238.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental. Psychology, 58*, 193–198.

Pinker, S. (1991). Rules of language, *Science, 253*, 530–535.

Poldrack, R. A., Clark, J., Pare-Blagoev, J., Shohamy, D., Creso Moyano, J., Myers, C., & Gluck, M. A. (2001). Interactive memory systems in the human brain. *Nature, 414*, 546–550.

Poldrack, R. A., & Gabrieli, J. D. E. (2001). Characterizing the neural mechanisms of skill learning and repetition priming: Evidence from mirror reading. *Brain, 124*, 67–82.

Poldrack, R. A., Prabhakaran, V., Seger, C. A., & Gabrieli, J. D. (1999). Striatial activation during acquisition of a cognitive skill. *Neuropsychology, 13*, 564–574.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *NeuroImage, 10*, 5–35.

Price, C. J., Moore, C. J., Humphreys, G. W., & Wise, R. J. S. (1997). Segregating semantic from phonological processes during reading. *Journal of Cognitive Neuroscience, 9*, 727–733.

Pugh, K. R., Frost, S. J., Sandak, R., Landi, N., Rueckl, J. G., Constable, R. T., Seidenberg, M. S., . . . Mencl, W. E. (2008). Effects of stimulus difficulty and repetition on printed word identification: A comparison of nonimpaired and reading-disabled adolescent cohorts. *Journal of Cognitive Neuroscience, 20*, 1146–1160.

Pulvermuller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience, 6*, 576–582.

Quartermain, D., McEwen, B. S., & Azmitia, E. C., Jr. (1972). Recovery of memory following amnesia in the rat and mouse. *Journal of Comparative and Physiological Psychology, 76*, 521–529.

Ranganath, C., & Blumenfeld, R. S. (2005). Doubts about double dissociations between short- and long-term memory. *Trends in Cognitive Science, 9*, 374–380.

Ranganath, C., & D'Esposito, M. (2005). Directing the mind's eye: Prefrontal, inferior and medial temporal mechanisms for visual working memory. *Current Opinion in Neurobiology, 15*(2), 175–182.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59–108.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, 118, 219–235.

Reber, P. J., Stark, C., & Squire, L. R. (1998). Contrasting cortical activity associated with category memory and recognition memory. *Learning and Memory, 5*, 420–428.

Reed, A. V. (1973). Speed-accuracy trade-off in recognition memory. *Science, 181*, 574–576.

Reed, A. V. (1976). The time course of recognition in human memory. *Memory and Cognition, 4*, 16–30.

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Language and Verbal Behavior, 12*(1), 1–20.

Roediger, H. L., Buckner, R. L., & McDermott, K. B. (1999). Components of processing. In J. K. Foster & M. Jelicic (Eds.), *Memory: Systems, process, or function?* (pp. 31–65). Oxford, UK: Oxford University Press.

Rogalsky, C., & Hickock, G. (2010). The role of Broca's area in sentence comprehension. *Journal of Cognitive Neuroscience, 23*(7), 1664–1680.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*(4), 573–605.

Rugg, M. D., Otten, L. J., & Henson, R. N. A. (2002). The neural basis of episodic memory: Evidence from functional neuroimaging. *Philosophical Transactions: Biological Sciences, 357*, 1097–1110.

Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tenses of English verbs. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 2: Psychological and biological models* (pp. 216–271). Cambridge, MA: MIT Press.

Ryan, J. D., & Cohen, N. J. (2003). Evaluating the neuropsychological dissociation evidence for multiple memory systems. *Cognitive, Affective & Behavioral Neuroscience, 3*, 168–185.

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

Sahni, S. D., Seidenberg, M. S., & Saffran, J. R. (2010). Connecting cues: Overlapping regularities support cue discovery in infancy. *Child Development, 81*, 727–736.

Salmon, D. P., & Butters, N. (1995). Neurobiology of skill and habit learning. *Current Opinion in Neurobiology, 5*, 184–190.

Schacter, D. L. (1992). Priming and multiple memory systems: Perceptual mechanisms of implicit memory. *Journal of Cognitive Neuroscience, 4*(3), 244–256.

Schacter, D. L., & Tulving, E. (1994). What are the memory systems of 1994? In D. L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 1–38). Cambridge, MA: MIT Press.

Schacter, D. L., Wagner, A. D., & Buckner, R. L. (2000). Memory systems of 1999. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 627–643). New York: Oxford University Press.

Schmolck, H., Kensinger, E., Corkin, S., & Squire, L. R. (2002). Semantic knowledge in patient H. M. and other patients with bilateral medial and lateral temporal lobe lesions. *Hippocampus, 12*, 520–533.

Schmolck, H., Stefanacci, L., & Squire, L. R. (2000). Detection and explanation of sentence ambiguity are unaffected by hippocampal lesions but are impaired by larger temporal lobe lesions. *Hippocampus, 10*, 759–770.

Schwartz, M. F., Saffran, E. M., & Marin, O. S. M. (1980). Fractionating the reading process in dementia: Evidence for word-specific print-to-sound associations. In M. Coltheart, K. Patterson, & J. C. Marshall, (Eds.), *Deep dyslexia* (pp. 259–269). London: Routledge and Kegan Paul.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry, 20*, 11–21.

Seidenberg, M. S., & Elman, J. L. (1999). Networks are not "hidden rules." *Trends in Cognitive Science, 3*, 288–289.

Seidenberg, M. S., & MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science, 23*(4), 569–588.

Seidenberg, M. S., MacDonald, M. C., & Saffran, J. R. (2002). Does grammar start where statistics stop? *Science, 298*, 553–554.

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stories: A neuropsychological study. *The Quarterly Journal of Experimental Psychology, 22*, 261–273.

Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition, 72*, B11–B21.

Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology, 33*, 5–42.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied, 74*(11), (Issue 498), 1–29.

Squire, L. R. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience, 4*(3), 232–243.

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory, 82*, 171–177.

Squire, L. R., Schmolck, H., & Stark, S. (2001). Impaired auditory recognition memory in amnesic patients with medial temporal lobe lesions. *Learning and Memory, 8*, 252–256.

Squire, L. R., Stark, C. E. L., & Clark, R. E. (2004). The medial temporal lobe. *Annual Review of Neuroscience, 27*, 279–306.

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences of the USA*, 93, 13515–13522.

Squire, L. R., & Zola-Morgan, S. (1988). Memory: Brain systems and behavior. *Trends in Neurosciences, 11*(4), 170–175.

Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition, 116*, 71–86.

Sternberg, S. (1975). Memory-scanning: New findings and current controversies. *Quarterly Journal of Experimental Psychology, 27*, 1–32.

Stoltzfus, E. R., Hasher, L., & Zacks, R. T. (1996). Working memory and aging: Current status of the inhibitory view. In J. T. E. Richardson, W. W. Engle, L. Hasher, R. H. Logie, E. R. Stoltzfus, & R. T. Zacks (Eds.), *Working memory and human cognition* (pp. 66–88). Oxford, UK: Oxford University Press.

Stowe, L. A., Paans, A. M. J., Wijers, A. A., Zwarts, F., & Vaalburg, G. M. W. (1999). Sentence comprehension and word repetition: a positron emission tomography investigation. *Psychophysiology, 36*, 786–801.

Surprenant, A., & Neath, I. (2009). *Principles of memory*. New York: Psychology Press.

Swanson, H. L., & Sachse-Lee, C. (2001). Mathematical problem solving and working memory in children with learning disabilities: Both executive and phonological processes are important. *Journal of Experimental Child Psychology, 79*, 294–321.

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., Ferruccio, F., . . . Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience, 17*(2), 273–281.

Thompson, R. F., & Kim, J. J. (1996). Memory systems in the brain and localization of a memory. *Proceedings of the National Academy of Sciences USA, 93*(24), 13438–13444.

Thompson-Schill, S. (2003). Neuroimaging studies of semantic memory: Inferring "how" from "where." *Neuropsychologia, 41*, 280–292.

Thompson-Schill, S. L., Swick, D., Farah, M. J., D'Esposito, M., Kan, I. P., & Knight, R. T. (1998). Verb generation in patients with focal frontal lesions: A neuropsychological test of neuroimaging findings. *Proceedings of the National Academy of Sciences, 95*, 15855–15860.

Thomson, D. M., & Tulving, E. (1970). Associative encoding and retrieval: Weak and strong cues. *Journal of Experimental Psychology, 86*(2), 255–262.

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

Tulving, E. (1979). Relation between encoding specificity and levels of processing. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 405–428). Hillsdale, NJ: Erlbaum.

Tulving, E. (1983). *Elements of episodic memory*. Oxford, UK: Clarendon Press.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology, 26*, 1–12.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53*, 1–25.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5*(4), 381–391.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*, 27–154

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*, 231–270.

Ungerleider, L. G., Doyon, J., & Karni, A. (2002). Imaging brain plasticity during motor skill learning. *Neurobiology of Learning and Memory, 78*, 553–564.

Vallar, G., & Papagno, C. (2002). Neuropsychological impairments of verbal short-term memory. In A. D. Baddeley, M. D. Kopelman, & B. A. Wilson (Eds.), *Handbook of memory disorders* (pp. 249–270). Chichester: Wiley.

Van Dyke, J. A. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(2), 407–430.

Van Dyke, J. A., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A retrieval interference theory of recovery from misanalyzed ambiguities. *Journal of Memory and Language, 49,* 285–413.

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. (in press). Cue-dependent interference in comprehension. *Journal of Memory and Language,* doi:10.1016/j.jml.2011.05.002

Van Orden, C. G., Pennington, B. F., & Stone, G. O. (2001). What do double dissociations prove? *Cognitive Science, 25,* 111–172.

Verhaeghen, P., Cerella, J., & Basak, C. (2004). A working memory workout: How to expand the focus of serial attention from one to four items in 10 hours or less. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30,* 1322–1337.

Wagner, A. D., Koutstaal, W., & Schacter, D. L. (1999). When encoding yields remembering: Insights from event-related neuroimaging. *Philosophical Transactions: Biological Sciences, 354,* 1307–1324.

Wagner, A. D., Poldrack, R. A., Eldridge, L. L., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1998). Material-specific lateralization of prefrontal activation during episodic encoding and retrieval. *NeuroReport, 9,* 3711–3717.

Wang, Y., Sereno, J. A., Jongman, A., & Hirsch, J. (2003). fMRI evidence for cortical modification during learning of Mandarin lexical tone. *Journal of Cognitive Neuroscience, 15,* 1019–1027.

Warner, J., & Glass, A. L. (1987). Context and distance-to-disambiguation effects in ambiguity resolution: Evidence from grammaticality judgments of garden path sentences. *Journal of Memory and Language, 26,* 714–738.

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain, 107*(3), 829–853.

Watkins, O. C., & Watkins, M. J. (1975). Build-up of proactive inhibition as a cue overload effect. *Journal of Experimental Psychology: Human Learning and Memory, 104,* 442–452.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review, 72,* 89–104.

Wickelgren, W. A. (1965). Short-term memory for phonemically similar lists. *The American Journal of Psychology, 78,* 567–574.

Wickelgren, W. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica, 41,* 67–85.

Willingham, D. B. (1998). A neuropsychological theory of motor skill learning. *Psychological Review, 105*(3), 558–584.

Wise, S. P., Murray, E. A., & Gerfen, C. R. (1996). The frontal cortex-basal ganglia system in primates. *Critical Reviews in Neurobiology, 10*(3–4), 317–356.

Zeki, S., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, R. S. (1991). A direct demonstration of functional specialization in human visual cortex. *Journal of Neuroscience, 11,* 641–649.