Interactions between episodic and semantic memory

Neal W Morton

February 5, 2013

Contents

Introduction

Memory researchers have identified a number of distinct forms of memory. Tulving (1972) proposed that, among forms of declarative memory, there should be a distinction between semantic memory (memory for facts, free of context) and episodic memory (containing information about particular episodes in one's life). According to this view, episodic memory involves a process of "mental time travel" where the brain state associated with the original episode is reactivated. In contrast to semantic memory, where relatively abstract information is retrieved, episodic retrieval also involves reactivation of temporal and spatial context. These two forms of memory have distinct goals: Episodic memory requires rapid learning of arbitrary events, while the development of semantic memory involves slower learning that extracts information about statistical regularities in the environment (McClelland et al., 1995). Given their separate goals, it has been proposed that episodic and semantic memory rely on distinct brain regions with different properties (McClelland et al., 1995).

Models of semantic memory and episodic memory have mostly been developed in isolation. Models of semantic memory have often been constructed by hand (Collins & Loftus, 1975; Feldbaum, 1998) or from aggregate data (Landauer & Dumais, 1997; Steyvers et al., 2004), rather than by gradual extraction of information from individual episodes. Similarly, models of episodic memory have often ignored the influence of semantic knowledge on new learning, for the sake of simplicity (Howard & Kahana, 2002a; Raaijmakers & Shiffrin, 1980). However, recently, memory researchers have begun to develop models that attempt to characterize performance in both episodic and semantic memory tasks (McClelland et al., 1995; Howard et al., 2011; Socher et al., 2009), and determine the precise relation between episodic and semantic learning. This work is motivated at a basic theoretical level; similar to the effort in the physics community to find a unifying theory of classical physics and quantum mechanics, there is potentially great utility in the development of a unified theory of episodic and semantic memory. A unified theory would provide insight into how people develop sophisticated semantic knowledge based on scattered experience, and would allow better

understanding of the effects of brain damage on different forms of memory. Furthermore, models that attempt to characterize both episodic and semantic memory can be constrained by data in both types of tasks, providing further basis for selecting between different models. Finally, improved knowledge of the mechanisms by which pre-existing semantic knowledge influences new learning may be important for educational and legal applications.

Experimental data in episodic memory tasks reveals that memory performance in nominally episodic tasks is strongly influenced by semantic associations. Semantic influences have been studied extensively in the free recall paradigm, where participants study a list of words then are asked to recall them in any order. In free recall, participants tend to successively recall items that are in the same taxonomic category, even when items in a category are not presented adjacent to one another in the list (Bousfield, 1953). Even when there is no obvious semantic structure to a set of to-be-learned materials, subtle semantic associations affect the order in which items are remembered (Howard & Kahana, 2002b). Semantic associations can also exert a strong effect on what is remembered, in some cases inducing the formation of false memories. When participants learn a list of words that are each related to a common theme (e.g. in the SLEEP theme, participants might see BED, NIGHT, PILLOW, etc.), they have a strong tendency to recall the common associate, even though it was never presented on the list (Deese, 1959b). Performance in this type of paradigm, known as a false-memory task, provides important constraint for models of declarative memory (Kimball et al., 2007), as do studies of semantic organization in free recall (Raaijmakers & Shiffrin, 1980; Sirotin et al., 2005; Polyn et al., 2009).

Although many features of semantic memory appear to be developed over a long timescale of years (McClelland et al., 1995), there is evidence that generalization of knowledge also occurs on much shorter timescales (Howard et al., 2005; Kumaran & McClelland, 2012). Researchers have found that this type of rapid inference relies critically on the medial temporal lobe, which has generally been assumed to be involved in learning veridical memories of episodes. Research on rapid inference may provide clues to how semantic knowledge is

built up from individual episodes (Howard et al., 2005; Rao & Howard, 2008).

Section 1 reviews models that attempt to characterize the structure and access of semantic memory. Section 2 discusses findings in episodic memory, with an emphasis on semantic organization and false memory in the free recall paradigm, and discuss the major models that have been proposed to explain behavior in this task, including the influence of semantic associations. Section 3 discusses recent modeling work that attempts to account for performance in both episodic and semantic memory tasks.

1 Models of semantic memory

Researchers studying semantic memory have attempted to describe the rich structure of human knowledge using a number of different frameworks. This section focuses on frameworks that have been incorporated into models of episodic memory search. Network models treat individual concepts as nodes that are densely interconnected. The connections indicate relations such as the properties of an object (e.g. birds have feathers) or the membership of an object in a taxonomy (e.g. a canary is a bird). In network models, two concepts are related when few connections have to be traversed to move between them in the network. In contrast, spatial models represent individual concepts as points in a high-dimensional space; the dimensions represent individual features, and the distance between two concepts in the space indicates how similar they are. Non-metric models are similar to spatial models in that they assume that concepts can be broken down into individual features, but they do not assume that conceptual similarity can be well-described in terms of a spatial metaphor.

1.1 Network models

Network models of semantic memory, where concepts are represented as connected nodes, allow representation of a rich set of knowledge, including object properties and taxonomies of concepts. In an attempt to characterize how people might represent semantic information,

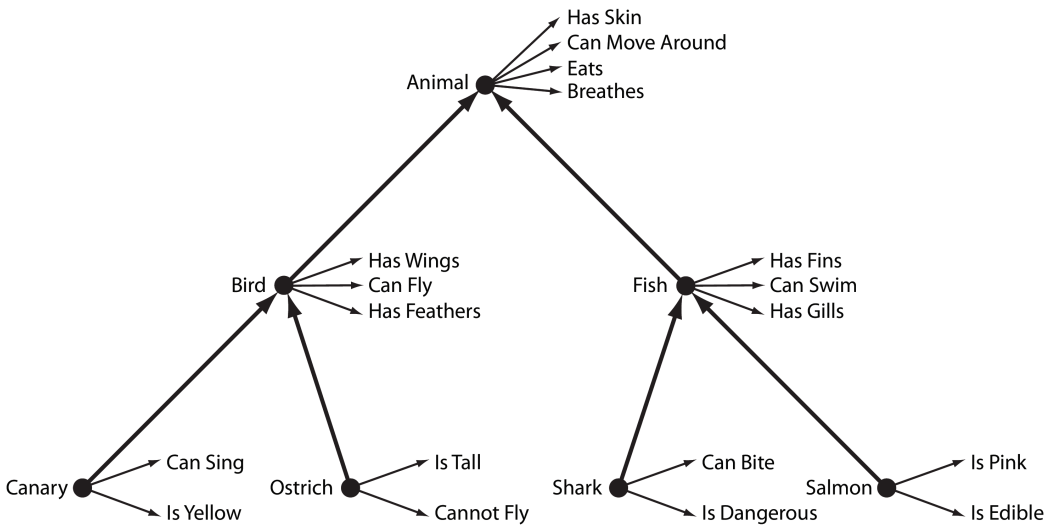


Figure 1. The semantic network model of Collins & Quillian (1969), as applied to a set of living things. Adapted from Collins & Quillian (1969).

Collins & Quillian (1969) described a method for storing and accessing semantic information in a computer program. In Quillian's model, concepts are represented as nodes in a network (see Fig. 1). Each concept is connected to properties; for example, the *canary* node might be connected to *sing* via a *can* connection. In the model, superordinate relations are represented through *isa* connections; for example, the *canary* node is connected to *bird* through an *isa* connection. Not all properties of a concept need to be stored with each node; some properties might only be connected to category representations. This allows the model to minimize the storage of redundant information. For example, in determining whether a canary has skin, one would likely have to infer this through the fact that a canary is a bird, a bird is an animal, and animals have skin. However, inference is only necessary when a relation has not been encountered and learned previously; for example, after one has learned that canaries have skin, the *has skin* property would be stored with the *canary* node; this property could then be accessed directly upon activation of *canary* (Collins & Loftus, 1975).

Collins & Quillian (1969) tested the predictions of this model by examining how quickly participants were able to evaluate the truth value of various statements. They contrasted properties that are unique to an object (e.g. canaries can sing), which the model assumes

are directly connected to that object, and more general properties (e.g. canaries have skin), which the model assumes are stored with superordinate representations (e.g. *has skin* is directly connected to *animal*). They predicted that statements relating to distant nodes (e.g. *canary* to *bird* to *animal* to *has skin*) should take more time to evaluate than statements relating to nearby nodes (e.g. *canary* to *can sing*). Consistent with the model, they found that reaction time was faster when evaluating unique properties, compared to when evaluating general properties (Collins & Quillian, 1969).

Quillian's network model was later developed into an elaborated model known as spreading activation theory (SAT; Collins & Loftus 1975). SAT assumed that semantic statements are evaluated by activating two or more nodes in the semantic network, and searching for an intersection between paths emanating from the two nodes. As nodes become activated, they also hold a tag indicating the source of the spreading activation. When activation at an intersection increases over a threshold, the path is evaluated for validity. This mechanism allowed the model to account for data in semantic priming experiments; activation of a node would spread from that node, allowing a quicker response when the probe was presented (Collins & Loftus, 1975). SAT also includes representations of superordinate relationships, which are necessary to account for some semantic decision strategies (Collins & Loftus, 1975). As discussed in Section 2, many authors have proposed that superordinate category representations of the type incorporated in network models are critical for guiding recall of strongly categorized lists; participants are assumed to recall some type of category label, which is used to cue for items from that category (e.g. Bousfield & Cohen 1953; Tulving & Pearlstone 1966; Pollio et al. 1969; Raaijmakers & Shiffrin 1980).

1.2 Spatial models

Although Quillian's (1969) network model can account for reaction time in semantic evaluation tasks, it is rather unconstrained, and must be coded by hand to represent the domain of knowledge being modeled. Landauer & Dumais (1997) took a different strategy to modeling

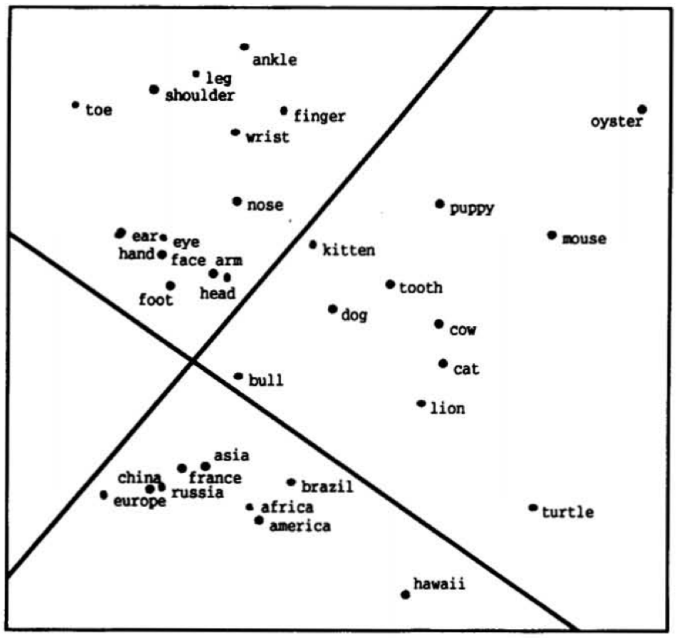


Figure 2. Example of a two-dimensional spatial representation of words derived from multidimensional scaling of co-occurrence information. The lines indicate rough division into taxonomic categories of body parts, countries, and animals. Adapted from Lund & Burgess (1996).

semantic knowledge. Rather than attempting to represent the richness of the different types of relations between words in semantic memory, they focused on determining the overall pairwise similarity between words in the lexicon. This pairwise similarity could then be applied to simple measures of semantic learning such as tests of synonyms. In order to correctly answer a synonym test (where a synonym of a cue word must be chosen from a list of candidates), one must have a good understanding of each of the words. However, children learn few words through direct instruction (Landauer & Dumais, 1997). Plato noted this problem of apparently insufficient learning material to explain people's level of knowledge, and proposed that some knowledge must be innate (Landauer & Dumais, 1997). Landauer & Dumais (1997) proposed an alternate solution to Plato's problem: They suggested that learning can occur in the absence of direct instruction, because humans are capable of performing inference on the reading material they are exposed to. They used text from Grolier's *Academic American Encyclopedia*, which contains 30,473 articles aimed at young students,

as an approximation of the written material that a student might encounter as they are learning new words, and attempted to extract the similarity structure of the words used in that text (Landauer & Dumais, 1997).

In order to gain traction on the problem of semantic learning and inference, Landauer & Dumais (1997) relied on the assumption that words that appear in the same context have similar meanings. This general principle has been used by a number of methods for extracting the structure of semantic memory (Lund & Burgess, 1996; Landauer & Dumais, 1997; Jones & Mewhort, 2007; Rao & Howard, 2008; Howard et al., 2011); the co-occurrence of words in the same contexts provides a simple metric that can be easily extracted from written text, without having to make assumptions about pre-existing knowledge structures such as word type or sentence structure. Landauer & Dumais (1997) defined "context" at the level of documents; if two words appeared in the same document (in this case, in the same article in the encyclopedia), they were assumed to be more similar than words that did not co-occur in a document. However, even with the substantial text corpus used by Landauer & Dumais (1997), many word pairs never co-occurred in any document. Therefore, inference is necessary to determine the similarity of all words in the corpus. Landauer & Dumais (1997) used singular value decomposition (SVD) to reduce the dimensionality of the document co-occurrence values. This resulted in a set of high-dimensional vectors corresponding to each word in the corpus; the vectors of any two words could then be compared to determine their similarity. Landauer & Dumais (1997) used the cosine similarity between two vectors as a measure of similarity of the corresponding words. One can also calculate the dot product between two vectors for use as a similarity measure; this metric is sensitive to word frequency (high-frequency words are associated with a large number of associates), which is useful for predicting performance in some semantic tasks (Griffiths et al., 2007).

Landauer & Dumais (1997) tested the ability of LSA to extract semantic structure by applying it to a synonym test; the probe word was compared to all of the choices of possible synonyms, and the most similar item was chosen. Performance was poor if all 30,473 articles

were used, without dimensional reduction, and performance was also poor if the number of dimensions retained from the SVD was too small. They found that the optimal number of dimensions was around 300. The optimal dimensionality performed comparably to human participants on the test of English as a foreign language (TOEFL), suggesting that dimensionality reduction by SVD is able to produce the degree of inference exhibited by human subjects in vocabulary learning (Landauer & Dumais, 1997).

Although Landauer & Dumais (1997) emphasized the importance of inference through dimensionality reduction, recent work has demonstrated that dimensionality reduction is not necessary if a sufficient amount of co-occurrence data is used (Recchia & Jones, 2009). Using a simple measure of co-occurrence, pointwise mutual information (PMI), it is computationally feasible to process a much larger corpus, making up for the lack of inference. Furthermore, the less-computationally intensive calculation makes it more feasible to train on different corpora in order to create domain-specific similarity spaces. For example, in a medical setting, one might want to focus on the biological sense of the word "heart" rather than other senses (Recchia & Jones, 2009).

Although both LSA and PMI are able to automatically extract a great deal of information from text corpora, they have little to say about the process of learning, since they operate on complete co-occurrence matrices containing information aggregated over many individual documents. Jones & Mewhort (2007) proposed the BEAGLE (bound encoding of the aggregate learning environment) model of gradual semantic learning, which allows semantic item representations to be built up slowly over time. In BEAGLE, each word is assumed to be associated with an "environment vector", a random representation that corresponds to the perceptual correlates (e.g. orthography) of the word. Each word also has a second "context" representation that comes to reflect semantic similarity between words, through learning. When a word is presented in a given context (defined as a sentence), the environment vectors of the other words are added to its context vector. As a result, words that co-occur with the same words come to have similar representations since their context

vectors have incorporated the environment vectors of the same words. In contrast to LSA, which requires the calculation of matrix operations to learn indirect associations, the simple summing of environment vectors is sufficient to learn about indirect relationships between words; the pairwise similarity structure extracted by BEAGLE is similar to the similarity derived using LSA (Jones & Mewhort, 2007).

However, the use of pure co-occurrence information is insufficient to learn certain types of associations between words. For example, if the words "Martin Luther" occur in a text corpus, they are very likely to be followed by "King, Jr." Jones & Mewhort (2007) included a separate mechanism to extract order information as well as co-occurrence information. To learn associations between words, they employed a type of holographic learning; a similar mechanism was used in the TODAM model of episodic memory to learn interitem associations (Murdock, 1982; Jones & Mewhort, 2007). In BEAGLE, interitem associations are stored using the convolution operation; associated items can then be retrieved using correlation. In contrast to TODAM, BEAGLE distinguishes between forward and backward associations; because convolution is fundamentally symmetric, Jones & Mewhort (2007) first applied distinct transforms to forward and backward associations so that the stored representations of forward and backward associations are distinct. The advantage of this holographic approach is that a single representation can store both item representations (here, built up through summing of environment vectors) and order information (through convolving words in specific relative positions, e.g. two words ahead or one word back). The full BEAGLE model assumes that each item is represented by a composite vector that contains both co-occurrence and order information. These distinct components encode qualitatively different types of information; co-occurrence tends to capture the similarity between words that share the same topic (e.g. "bread", "baker", "oven"), while order information captures words that share similar roles, such as parts of speech (Jones & Mewhort, 2007). Importantly, this result shows that at least some aspects of sentence structure can be extracted from text corpora without making assumptions about the intrinsic structure of human semantic memory

(Jones & Mewhort, 2007).

Although LSA allows a great deal of semantic information to be extracted from large text corpora, and performs well on relatively direct tests of meaning such as synonym tests (Landauer & Dumais, 1997), it does less well at predicting the effect of semantic relatedness on performance in episodic memory tasks (Steyvers et al., 2004; Sirotin et al., 2005). Rather than examining a text corpus, Steyvers et al. (2004) examined free association data, which provide a different source of information about semantic associations. They examined data from a large study conducted at the University of South Florida, which recorded free association data for 72,000 word pairs, based on the responses of over 6,000 participants (Nelson et al., 2004). Even with this large dataset, many pairs of words were never given as associates of one another. Therefore, in addition to direct associations, Steyvers et al. (2004) examined secondary associations, which are mediated by two direct associations. Although free-association data is inherently asymmetric, for simplicity they summed over forward and backward associations between pairs of words to obtain an overall strength of association. SVD was then applied to the association strengths between each pair of words; similar to LSA, this allows calculation of similarity between words with no primary or secondary associations (Steyvers et al., 2004). The resulting high-dimensional word representations assign similar words nearby representations in a word association space (WAS). By deriving semantic similarity from free-association data, rather than co-occurrence in text, WAS performs better than LSA in predicting the effect of semantic similarity on behavior in recognition, free recall, and cued recall tasks (Steyvers et al., 2004).

Spatial models of semantic memory such as LSA and WAS allow automatic estimation of pairwise similarity for large word stimulus pools; as a result, they are useful for examining behavior in episodic memory tasks, which often involve large numbers of distinct stimuli (Howard & Kahana, 2002b). As discussed in Section 2, estimates of pairwise similarity between words can be used to measure the degree to which participants tend to group together semantically related words during episodic recall (Howard & Kahana, 2002b; Polyn

et al., 2009; Sederberg et al., 2010). Furthermore, pairwise similarity estimates derived from spatial models have been incorporated into models of episodic memory search in order to account for pre-existing semantic associations between words (Sirotin et al., 2005; Polyn et al., 2009). The semantic similarity structure assumed by a model can influence its account of data in episodic memory tasks; using realistic estimates of semantic similarity, such as those derived from LSA or WAS, allows a stronger test of a model than if random similarity values are used (Johns & Jones, 2010).

1.3 Non-metric models

Although estimated spatial representations of words have proven successful in predicting behavior in synonym tasks (Landauer & Dumais, 1997), as well as data from episodic memory tasks (Steyvers et al., 2004), there are theoretical issues with the assumption that words are represented in a high-dimensional space (Griffiths et al., 2007). When the similarity between words is represented by their distance in some common space, a property of their relations known as the *triangle inequality* must hold. If there are three words A, B, and C in the same space, then the distance between A and C must be greater than or equal to the sum of the distance between A and B and the distance between B and C (Shepard, 1980; Tversky, 1977). However, the triangle inequality appears to be violated by some word triplets; for example, a gas jet is similar to the moon in that both cast light, and the moon is similar to a ball in that they are both round, but a gas jet is not similar to a ball (Griffiths et al., 2007). Another issue with the LSA approach is that it assumes that each word is represented as a single point; this means that it cannot account for words with multiple senses (such as *bank*, referring to either a financial institution or the side of a river) without further elaboration (Griffiths et al., 2007).

Griffiths et al. (2007) developed an approach to automatic extraction of meaning from text corpuses that addresses these issues. They used the same type of corpus data as used by Landauer & Dumais (1997), a co-occurrence matrix indicating which words appeared in

which documents in the corpus. They defined a simple generative model that was assumed to have generated the set of words in each document in the corpus. The model represents meaning in terms of topics; each topic z is defined as a probability distribution over each word w, indicating words that are likely to be mentioned in the topic. Each document in a corpus has a *gist* g, which is represented in the model as a probability distribution over topics. Each word in the document is generated by choosing a topic (based on the probability distribution determined by the document's gist), then generating a word based on the probability distribution of the topic. Using the tally of which words occur in each document ($P(w|g)$, the probability of a word occurring given the gist of the document), it is possible to infer the properties of the generative model, using Bayes' rule. The probability distributions that must be estimated are $P(w|z)$, the probability of generating a word given a topic, and $P(z|g)$, the probability of each topic given the gist of the document (Griffiths et al., 2007).

In contrast to LSA, where each word corresponds to a single point in a semantic space, in the topic model, a given word may be included in multiple topics. For example, this allows the model to capture two meanings of the word "characters"; in a topic relating to printing, "characters" refers to printed symbols, whereas in a topic related to drama, "characters" refers to the fictional people appearing in a work of fiction. The representation of words in terms of topics also allows the topic model to capture violations in the triangle inequality. If two words w_1 and w_2 have high probabilities under the same topic, they will tend to be associated. If another word w_3 appears in one of the same topics as w_2, then w_1 and w_2 will be judged as similar, and w_2 and w_3 will be judged as similar. However, this does not imply that w_1 and w_3 are similar to one another; they do not necessarily appear in the same topics, and so may be unrelated even though they share a common associate.

In the topic model, a gist can be extracted from a sequence of words by estimating the probability of each topic given the combination of words observed; this distribution over topics can then be used to predict which words might come next (Griffiths et al., 2007). The

section below on models of semantic clustering in free recall discusses LDA-TCM, a model that takes advantage of this generative ability of the topic model to predict the sequence of recalls observed in a free-recall task (Socher et al., 2009). As also discussed below, there is evidence that an automatic process of extracting gist from a sequence of words can have a marked effect on episodic memory, causing recall of words that were never presented on a list (Kimball et al., 2007). LDA-TCM provides a theory of the nature of this gist representation, and its role in influencing recall (Socher et al., 2009).

2 Episodic memory tasks

In contrast to the semantic memory tasks described in the previous section, which focus on measuring semantic knowledge that is not tied to any particular spatial or temporal context, episodic memory tasks examine how people form memories for individual episodes that are associated with a specific context. Although episodic memory tasks generally do not explicitly require semantic processing, research has shown that semantic relations nonetheless can have a dramatic effect on performance. This section focuses on the free-recall paradigm, where the relatively unconstrained nature of the task allows examination of subtle effects of semantic associations. First, the major findings from the free-recall paradigm and the dominant models that have been proposed to explain these findings will be discussed. Then findings concerning the effects of semantic memory on the task, and attempts to extend models of free recall to account for these effects will be explored.

2.1 Free recall

In the free-recall paradigm, participants study a list of words, then are asked to recall the words on the list in any order. Behavior in this task can be decomposed into three phases: recall starting, transitions, and stopping (Kahana, 2012). When the recall test occurs immediately after the end of the list, participants demonstrate a strong tendency to

start by recalling an item from one of the last positions in the list; this is referred to as the *recency effect* (Murdock, 1962). Recent items are more likely to be recalled first, and are more likely to be recalled overall, than items in other positions (Howard & Kahana, 1999). Once recall has started, transitions from item to item exhibit *temporal organization*: Successive recalls tend to be from adjacent positions in the study list (Kahana, 1996). For example, if a participant studies the list *ABSENCE HOLLOW PUPIL PIANO FOUNTAIN*, and recalls *PUPIL*, his or her next recall will likely be *HOLLOW* or *PIANO*. This tendency is referred to as the *contiguity effect*. Participants are more likely to stop recall after an error than after a correct recall (J. F. Miller et al., 2012). Errors include repeating an item that has already been recalled, recalling an item from a prior list (prior-list intrusion [PLI]), or recalling an item not presented in any list so far (extra-list intrusion [ELI]). Both PLIs (Zaromb et al., 2006) and ELIs (Deese, 1959b) tend to be semantically related to the just-recalled item; this suggests that existing semantic associations can negatively affect episodic recall. The *False memory* section below discusses in more detail how semantic associations can distort episodic memories.

In order to account for serial position effects on recall, as well as other findings from short-term memory experiments, Atkinson & Shiffrin (1968) proposed a *dual-store* theory of memory. They proposed that memory is divided into a short-term store (STS) and a long-term store (LTS). The STS allows a few items to be held in active memory; in the absence of distraction, rehearsal can be used to maintain the items in the STS indefinitely. Items in the STS can be easily recalled. Unlike the STS, the LTS is assumed to have infinite capacity; however, recall from the LTS is cue-dependent: Retrieval requires an appropriate cue, such as an associated item or the context in which an item was learned. Associations between items, and associations from context to items are learned while items are being rehearsed in the STS.

Raaijmakers & Shiffrin (1980) implemented these assumptions in a computational model of memory search called search of associative memory (SAM). SAM assumes that participants

begin recall by recalling all of the items currently in STS. Because of the limited capacity of the STS, generally only the last few items on the list will still be active in memory; therefore, recall will tend to start with items at the end of the list, and those items are very likely to be recalled, since items in the STS can easily be recalled (Raaijmakers & Shiffrin, 1980). Consistent with this account, the recency effect is abolished if recall is preceded by a period of distraction, which is presumed to remove items from the rehearsal buffer (Glanzer & Cunitz 1966; Postman & Phillips 1965; but see Bjork & Whitten 1974). This variant of the free recall paradigm is known as *delayed free recall*.

SAM also provides an account of the contiguity effect in immediate and delayed free recall. Because interitem associations are formed among items that are in the rehearsal buffer at the same time, adjacent items on the list tend to become strongly associated to one another (Kahana, 1996). When an item is recalled, it therefore provides a good cue for other items that were near to it on the list, making nearby items likely to be recalled next.

Howard & Kahana (2002a) proposed an alternate account of recency and contiguity effects in free recall. They developed the temporal context model (TCM), which assumes that studied items become associated with a slowly changing representation of temporal context. Presentation of a studied item triggers retrieval of context that the item has been associated with in the past; as items on the list are presented, retrieved context is pushed into the current state of context to change it slowly over time. At the end of the list, the current state of context is used to cue memory to retrieve items on the list. The temporal context cue provides the most support for items that are associated with similar states of context. Because context changes slowly over time, the state of context provides a strong cue for recently studied items, thus giving rise to the recency effect (Howard, 2004). TCM can also account for the decrease in the recency effect observed in delayed free recall. The distraction task is assumed to cause context to change, so that the context at the time of test is less similar to items at the end of the list; this causes the recency items to lose their advantage over earlier items (Howard & Kahana, 2002a; Sederberg et al., 2008).

Critically, TCM assumes that recall of an item triggers retrieval of the contextual state associated with that item. The contextual retrieval mechanism in TCM allows it to account for the contiguity effect; the retrieved context provides a good cue for nearby items in the list, which were associated with similar states of context (Howard, 2004; Sederberg et al., 2008). The process of contextual reinstatement may correspond to the "mental time travel" that Tulving (2002) proposed takes place during episodic retrieval (Manning et al., 2011). Consistent with this account, brain activity related to the source context (e.g. visual vs. auditory presentation) of studied items is reactivated during episodic retrieval (Danker & Anderson, 2010; Polyn et al., 2012). TCM provides detailed predictions for the properties that a neural representation of temporal context should exhibit; there should be a brain region whose pattern of activity integrates information over time about presented stimuli and their source context, and the rate of this change should be related to subsequent recall performance (Polyn et al., 2012). Consistent with this prediction, evidence for integrative activity that influences subsequent recall performance has been found using multivariate pattern analysis of functional magnetic resonance imaging data (Polyn et al., 2012) and scalp electroencephalography data (Morton et al., in press).

TCM further predicts that when an item is recalled, there should be brain activity that reflects reinstatement of the temporal context associated with that item (Polyn & Kahana, 2008). The reinstated activity should be similar to activity observed during the initial presentation of the item, but should also be similar to activity observed during neighboring items, with similarity decreasing as a function of distance from the recalled item (Polyn & Kahana, 2008). A recent investigation of oscillatory neural activity measured in humans using electrocorticography (ECoG) during free recall found just this pattern of activity in the temporal lobe (Manning et al., 2011). Similar temporal context reinstatement has also been observed in human medial temporal lobe during a continuous recognition paradigm (Howard et al., 2012).

Although the basic versions of SAM and TCM provide good accounts of recency and

contiguity effects in immediate and delayed free recall, they contain no representation of semantic information, and therefore cannot account for effects of semantic relatedness on recall. The following sections review effects of semantic memory on free recall, the additions made to SAM and TCM in order to account for these findings, and other models that have been proposed to account for the data.

2.2 Semantic influences on recall performance

In addition to temporal influences exhibited in the recency and contiguity effects, the semantic relatedness of to-be-recalled items can exert a strong influence on behavior in free recall (Bousfield, 1953; Deese, 1959b; Howard & Kahana, 2002b; Zaromb et al., 2006). This influence of the semantic structure of the learned material exhibits itself in several measures of behavior, including the number of items successfully recalled, the order in which they are recalled, the number and type of incorrect responses, and the rate at which memory search progresses. This section outlines important findings relating to semantic organization in free and cued recall. These findings provide important constraints for models of episodic memory.

2.2.1 Measures of semantic relatedness

Researchers examining the influence of semantic associations in episodic memory tasks must first specify a model of semantic memory for their stimulus materials. This is an important issue, because assumptions about the semantic structure of stimuli can greatly affect conclusions about behavior (Manning & Kahana, 2012) and theories of the underlying memory system (Johns & Jones, 2010). Researchers studying episodic memory tasks have generally focused on simple models of semantic memory, either focusing on semantic relatedness between individual word pairs (e.g. Sirotin et al. 2005) or assuming a simple taxonomic structure (e.g. Raaijmakers & Shiffrin 1980), although some researchers have begun to incorporate more sophisticated semantic representations into their models (Socher et al.,

2009).

In order to determine the semantic relatedness of stimuli, one approach is to use stimuli that are generated from norming studies. For example, Deese (1959b) used lists of words that were related to a critical stimulus. The related words were generated by presenting participants with the critical stimulus and asking them to produce the first word that comes to mind. Using this method, researchers may use a critical word of the same type as the associated stimuli (e.g., Deese 1959b), or they may use category labels such as "a four-legged animal" or "a piece of furniture", prompting participants to name exemplars of that category (e.g., Battig & Montague 1969; Van Overschelde et al. 2004). Typically, all members of a given category are assumed to be equally semantically related to one another, and items in different categories are assumed to be unrelated. However, some investigators have drawn a finer distinction based on *taxonomic frequency*. Taxonomic frequency is the frequency with which participants in a norming study responded with a given item when prompted with a category name (e.g. Battig & Montague 1969). Recall behavior can then be contrasted in conditions where the to-be-recalled items are high- or low-taxonomic frequency associates of a given category (Bousfield et al., 1958). Free-association experiments can also provide more comprehensive data on associations between each pair of words in a stimulus pool (Nelson et al., 2004).

The advantage of using free-association data to estimate semantic relatedness is that it is relatively free of assumptions about the underlying memory system. However, the use of raw free-association data is impractical for experiments that rely on the use of a large number of unique stimuli (Howard & Kahana, 2002b). To address this issue, researchers have applied dimensionality reduction techniques such as singular value decomposition (SVD) to allow inference of the relatedness of word pairs, even when direct evidence of their association is not available (Steyvers et al., 2004). The raw data may be provided by free-association data (Steyvers et al., 2004) or from large corpuses of text that are representative of the reading materials that participants may have exposure to (Landauer & Dumais, 1997; Griffiths et

al., 2007).

2.2.2 Semantic clustering

The "standard" free-recall paradigm uses randomly chosen "unrelated" words as stimuli, with the primary manipulation being the serial position in which a word was presented (e.g., Murdock 1962). Bousfield (1953) examined how recall is affected when to-be-remembered stimuli have a strong semantic structure. He examined free recall of stimuli taken from several distinct taxonomic categories. Bousfield (1953) found that, even when items in the same category were never presented adjacent to one another, participants tend to successively recall, or cluster, items from the same category. He introduced the ratio of repetition, the proportion of transitions in recall that were between items in the same category, as a measure of category clustering strength. Participants exhibited a higher ratio of repetition than would be expected if items were randomly sampled from the list (Bousfield, 1953). The phenomenon of category clustering has been replicated in many variations of categorized free recall, using a number of different measures of clustering (Shuell, 1969; Roenker et al., 1971).

Tulving (1962) extended these results by demonstrating that participants show consistent organization even in the absence of experimental manipulations designed to impose semantic structure on a list. Participants studied and recalled a random list of words multiple times, with the order of the words randomized for each trial. Over trials, the order of each participant's recalls developed regularities where certain pairs of items tended to be recalled together. Different participants developed somewhat similar orderings of recalled words, suggesting that they discovered similar relations between the items on the list to use to guide recall (Tulving, 1962).

With the later development of techniques for measuring pairwise similarity between large sets of word stimuli (e.g. Landauer & Dumais 1997), researchers were able to test whether participants organize recall according to specific semantic similarity metrics, even when there is no obvious semantic structure to the materials. Using latent semantic analysis (LSA;

Landauer & Dumais 1997) of a large text corpus, Howard & Kahana (2002b) estimated the similarity between each pair of words in a given list. They then divided each pair of words into bins reflecting the strength of their semantic similarity, based on the cosine similarity between the words in LSA space (see section 1.2 for details of LSA). For each recalled item, they determined what words on the list were still available (assuming that any words already recalled would not be repeated), and calculated the semantic similarity between the just-recalled word and each of the remaining items. They recorded both what transitions were possible, and which transition the participant actually made. Over all the transitions between recalled items, they calculated the probability of making a transition within a given semantic bin, conditional on a transition of that degree of semantic similarity being available. They found that the conditional response probability increased as a function of semantic similarity. Conditional response probability differences were observed even between pairs of items with no obvious similarity and pairs with even less similarity (Howard & Kahana, 2002b), suggesting that semantic clustering is a graded and pervasive phenomenon rather than an idiosyncratic mnemonic strategy. Semantic clustering of random words has been demonstrated using a number of different measures of semantic similarity, including ones based on triad similarity judgments (Romney et al., 1993), free-association norms (Manning & Kahana, 2012), and co-occurrence of word pairs in a large text corpus (Howard & Kahana, 2002b).

2.2.3 Memory availability versus accessibility

G. A. Miller (1956) reviewed evidence that human information processing capacity is limited to a small number of "chunks" of information. He also discussed evidence that the amount of information that can be held in immediate memory can be increased through a process of recoding. For example, a person can increase their capacity for memorizing strings of binary digits by recoding groups of digits into codes that they have previously memorized (e.g., "10100" is recoded as "20", and "01001" is recoded as "9"). Using this technique,

only the codes are stored, rather than the original raw input. During recall, one can then translate the codes back to obtain the original sequence (G. A. Miller, 1956). In essence, this technique increases the span of immediate memory (for specific materials) by offloading much of the processing to long-term memory.

Cohen (1963a) suggested that a similar process might allow participants in a categorized free recall study to increase their recall performance. Category clustering might reflect a process of recoding, where each category would define a chunk of information containing the items in that category. He ran participants in a free-recall paradigm with word stimuli drawn from 20 different categories, with 3 or 4 items in each category. If, as he hypothesized, recall of individual items on the list is mediated by recall of chunks of items from the same category, then there are two measures of interest: the number of categories recalled, and the number of items recalled given that their category was recalled. A category was considered recalled if there was at least one recall from that category. According to Cohen (1963a), the number of categories recalled reflects the success of participants in recalling different chunks, and the number of items recalled per category (conditional on category recall) reflects the degree of success in recalling individual items from a chunk.

To compare performance in his categorized free recall study to performance in standard free recall, Cohen (1963a) used a formula developed by Murdock (1960) to predict performance in free recall of unrelated words. Murdock (1960) based this formula on a series of experiments examining recall performance in the standard free recall paradigm with a range of list lengths and presentation times. Cohen (1963a) found that the number of the 20 distinct categories recalled was about the same as the number of items that would be expected to be recalled in a list of 20 unrelated items; this suggests that participants were successfully able to use recoding, storing 20 chunks rather than storing each of the 70 items on the list independently. Cohen (1963b) verified this relation between recall of categorized lists and recall of unrelated words. When total presentation time is controlled for, the number of items recalled from a list of unrelated words is approximately equivalent to the number of

categories recalled from a categorized list (Cohen, 1963b). Furthermore, Cohen (1966) found that participants tend to either recall no items from a given category, or recall most of the items from that category; he referred to this tendency as *some-or-none* recall. He argued that the tendency to recall most of the items in a category, given that the category was accessed at some point during recall, is consistent with recoding on the basis of taxonomic category.

Later researchers expanded on the work of Cohen by investigating experimental variables that differentially affect category recall and items recalled per category. Tulving & Pearlstone (1966) contrasted free recall of a categorized list to a cued recall paradigm, where participants were provided with the names of the categories on the list, and asked to recall the items belonging to each category. They used blocked presentation of items in each category, and varied the category size (1, 2, or 4 items) and the list length (12, 24, or 48 items). The list was followed by either free recall or cued recall, where the category names were provided as a cue to recall the items from that category. In the free recall condition, the percentage of categories with at least one recalled item decreased as a function of list length. This finding parallels the list-length effect observed in recall of unrelated words (Murdock, 1960), again suggesting that category recall is similar to recall of individual items in a list of unrelated words. In contrast, the number of items recalled per category did not vary with list length, suggesting that recall of items within a category is not subject to interference from items in other categories (Tulving & Pearlstone, 1966). They also found that providing a category retrieval cue can have a strong influence on recall performance. Overall recall was greater in the cued-recall condition than in free recall (Tulving & Pearlstone, 1966). After the first test (which was either free or cued recall, depending on the participant group), all participants were given a cued recall test. Interestingly, they found that participants who first performed free recall were able to recall more items on the later cued-recall test. To explain this striking result, Tulving & Pearlstone (1966) drew a distinction between *availability* and *accessibility*. They argued that items may be stored in memory, that is, available, but still

not be accessible due to a lack of sufficient memory cues. In the case of categorized free recall, they argued that the category label provides a powerful cue to make items from that category be accessible. In their experiment, some available items were not recalled during free recall, but were recalled during cued recall when the category label made those items accessible (Tulving & Pearlstone, 1966).

A study by Cohen (1966) also supports the hypothesis that recall of items within a given category is effectively restricted to that category, without interference from words in other categories on the list. Participants studied lists with 10, 15, or 20 categories, with 3 or 4 exemplars per category. To examine how well items in each category are recalled, given that the category itself was remembered, Cohen (1966) examined mean item recall per category, conditional on at least one item from that category having been recalled. Mean recall per category did not vary with the number of categories on the list, suggesting that items in other categories do not interfere with search within a category (Cohen, 1966).

However, there is also evidence that memories of items from different categories do interfere with each other during retrieval. Smith (1971) had participants learn a list with words drawn from 7 distinct taxonomic categories, with the items within each category presented together; then they were then given the name of each category and asked to recall items from the list in that category. The order in which the categories were tested was balanced across participants using a 7x7 Latin square. Smith (1971) found that recall declined with output position; recall was decreased for categories tested later. The effect was significant even when recall from the last input category was excluded, suggesting that the results cannot be explained by recall from short-term memory (Smith, 1971). In a separate experiment, Smith (1971) examined the effect of varying the number of items from each category. He found that recall performance declined more quickly in conditions with more items in each category (Smith, 1971); this is consistent with the hypothesis that the amount of output interference is directly related to the number of items recalled in each recall period (Roediger, 1973). Raaijmakers & Shiffrin (1980) proposed an explanation for this output interference effect.

Based on the search of associative memory (SAM) framework, they assumed that, during encoding, items become associated with a representation of the list context. During recall, this context serves as a cue to retrieve items from the list. When an item is recalled, the association between it and the list context is incremented. As a result, categories probed at later recall tests will face increased competition from any items from other categories that were recalled during previous tests, and will be less likely to be recalled (Raaijmakers & Shiffrin, 1980).

2.2.4 Inter-response times

In addition to examining the number of items recalled and the order in which items are recalled (i.e., semantic clustering), response times provide an important source of information about the processes underlying memory search. Most classic free recall studies used written recall to collect participant responses (e.g. Murdock 1962). In order to accurately measure the timing of each recall, studies of response timing in free recall have instead used vocal recall, which is recorded and scored by the experimenter after the session (Pollio et al., 1968; Solway et al., 2010). There have been no systematic comparisons of written and vocal recall, but informal comparison of published studies using the two different methods shows roughly similar results in terms of serial position effects on recall (Murdock, 1962; Murdock & Okada, 1970), temporal clustering (Howard & Kahana, 1999), and category clustering (Bousfield, 1953; Morton et al., in press).

Inter-response times (IRTs) have been alternately measured as the time from offset of one vocalization to the onset of the next vocalization (e.g. Pollio et al. 1968; Murdock & Okada 1970), or the time between the onsets of adjacent recalls (e.g. Patterson et al. 1971; Polyn et al. 2009). These different scoring measures should provide similar results, unless the duration of each vocalization varies between the conditions of interest. Patterson et al. (1971) argued that onset–onset scoring is generally preferable, because onsets are easier to detect than offsets, and participants may be engaging processes related to ongoing memory

search even as they are vocalizing a recalled word.

Before discussing IRTs observed in categorized free recall paradigms, it is useful to understand the basic pattern of IRTs observed in standard free recall. Murdock & Okada (1970) examined IRTs in a standard free recall paradigm with lists of 20 "unrelated" (that is, without obvious semantic structure) words from the Toronto word pool (Friendly et al., 1982). They found that IRTs increased exponentially with output position, with the first transition (between the first and second recalled words) taking less than 1 s on average, and the last transition taking about 10 s. This pattern of results suggests that the time to recall a word varies inversely with the number of words remaining to be recalled. One explanation for the increase in IRTs with output position is that recall involves sampling with replacement (Murdock & Okada, 1970). According to random sampling theories of free recall, even after a given word has been recalled, that word might still be retrieved during subsequent search of memory. Given that participants rarely repeat themselves (Kahana, 2012), these retrievals of already-recalled items must generally be recognized as repetitions, and therefore not vocalized (Patterson et al., 1971). Nevertheless, failed attempts to retrieve new items from the list may slow down the recall process, causing IRTs to increase as recall of already-retrieved items becomes more frequent as recall progresses (Murdock & Okada, 1970). Raaijmakers & Shiffrin (1980) demonstrated that the SAM model, which incorporates the assumption that already-recalled items compete with items that have not yet been recalled, produces a positively accelerated increase in IRTs with output position, similar to the results of Murdock & Okada (1970).

Pollio et al. (1968) examined whether IRTs are modulated by semantic similarity of the word pairs being transitioned between. Participants performed free recall of lists composed either of associates to MUSIC or to COMMAND. Pollio et al. (1968) first examined average IRTs as a function of output position, and found a similar exponential increase to that found by Murdock & Okada (1970). When examining the timing of individual recall sequences, however, they found that participants alternated between fast and slow sequences of recalls.

They first defined *rate alternations*, where the IRT was at least five times greater, or less than, the previous IRT. Then, for each recall period, they identified the fastest sequence (in terms of mean IRT) and slowest sequence. They found that the pairwise semantic similarity of items in fast sequences was greater than in the slow sequences[1]. These results suggest that transitions between strongly related items tend to be faster than transitions between weakly related words (Pollio et al., 1968). Howard & Kahana (2002b) found similar results in a free-recall experiment with lists of random words. They found that IRT decreased as a function of the relatedness of the words being transitioned between (as measured by cosine distance in LSA space).

While Pollio et al. (1968) and Howard & Kahana (2002b) investigated the effect of subtle variations in semantic similarity on IRTs, Pollio et al. (1969) examined how IRTs vary in free recall of strongly categorized materials. They had participants study a list of 25 words, with 5 words from each of 5 taxonomic categories from the norms of Cohen et al. (1957). They found that participants had a strong tendency to first attempt recall of items from each category, and only then to revisit previously recalled categories to attempt to recall more items from each category. To simplify their analysis of IRTs, they focused on the IRTs during the initial recalls from each category. They separated recall transitions into *between-category* and *within-category* transitions, and found that between-category transitions had consistently longer IRTs. They also examined how IRTs changed over successive recalled categories. Between-category IRTs increased exponentially with category output position, while the average within-category IRTs only increased slightly with output position. They also found that, within a given category, the IRTs of successive transitions between items increased with output position. Pollio et al. (1969) proposed that, similar to the increase in IRTs with output position in recall of unrelated words (Murdock & Okada, 1970), the exponential increase in between-category IRTs reflects a search that becomes more difficult as more categories are recalled. The slight increase in within-category transition IRTs with

[1]The details of the semantic similarity metric are unclear; the authors refer to a technical report, which presumably explains how the measure was derived, but I have been unable to obtain that report.

category output position may reflect output interference between categories, similar to the effects that Smith (1971) observed on recall performance. Finally, the increase in IRTs with successive recalls within a given category suggests that search within a category is also driven by a process that samples items in the category with replacement.

Patterson et al. (1971) followed up on the results of Pollio et al. (1969) using a similar paradigm. Patterson et al. (1971) argued that between-category IRTs, which Pollio et al. (1969) treated as simply reflecting the time to search for an untapped category, are actually more complex. Between-category IRTs are measured as the time between recall of the last item from a category and the recall of the first item from the next category. They argued that during this time, there are actually three steps taking place: (1) a failed retrieval attempt from the previous category (assuming that the category was not exhausted, which it usually was not), (2) search for an untapped category, and (3) search for an item within the new category. In an attempt to understand how these steps contribute to the time it takes to switch between categories, Patterson et al. (1971) gave participants in one condition an index card with the category names printed on it; this card was available throughout the trial. They hypothesized that this manipulation would eliminate any component of between-category IRTs related to search for a new category. In their control condition, where participants were not given the category names, they replicated the finding of Pollio et al. (1969) that between-category IRTs increase exponentially with category output position. In contrast, between-category IRTs did not increase with category output position when participants were given the category names during recall. Patterson et al. (1971) concluded that the exponential increase in between-category IRTs is due to participants searching for a new category, rather than the other components of between-category IRTs.

2.2.5 Semantic organization and temporal contiguity

Tulving (2002) argued that a hallmark of episodic memory is the subjective experience of "mental time travel." Retrieved-context models such as the temporal context model (Howard

& Kahana, 2002a) and the context maintenance and retrieval model (Polyn et al., 2009) assume that episodic memories are bound to a representation of temporal context; when an item is retrieved in free recall, it also retrieves its associated context. This contextual retrieval mechanism has been proposed as the cause of the subjective experience of mental time travel that Tulving (2002) described (Polyn & Kahana, 2008). Furthermore, Polyn & Kahana (2008) argue that temporal organization in free recall (i.e. the contiguity effect) is a result of this contextual retrieval. Given the importance of temporal contiguity in informing our understanding of episodic memory in general, it is important to understand how semantic and temporal influences on recall interact; these interactions may provide important clues to how pre-existing semantic associations affect new episodic learning.

Based on the framework of Atkinson & Shiffrin (1968), which proposes that there is a short-term store (STS), which is involved in elaborative processing of stimuli that facilitates transfer to the long-term store (LTS), Glanzer (1969) hypothesized that semantic associations between a pair of items in a list will only affect encoding if the items are in the STS at the same time. The STS has a limited capacity and only holds the most recently presented items. In simulating free recall with their SAM model, Raaijmakers & Shiffrin (1980) found that setting STS capacity to 4 provided a good fit to serial position effects. Glanzer (1969) assumed that associative relations between items in a list will only be discovered if they co-occur in the STS, and that discovery of associations between items will enhance the strength of their encoding. To test this hypothesis, he examined free recall of lists composed of 8 related pairs of words, with pairs presented with 0, 1, 3, or 7 words in between the words in the pair. For example, distance 0 lists might include the sequence: *stove*, *coal*, *ankle*, *foot*, etc., while distance 3 lists might include *stove*, *ankle*, *parent*, *alcohol*, *coal*, *foot*, *child*, *whiskey*, etc. Consistent with his hypothesis, Glanzer (1969) found that recall performance decreased as a function of the distance between related pairs. He did not report any information about the order of recalls, so it is unclear whether semantic organization (in this case, the tendency to successively recall items from a pair) was also modulated by the distance between pairs.

A number of researchers have examined a similar manipulation of categorized free recall paradigms, comparing lists where the items in each category were *blocked* (i.e. adjacent to one another) with *random* lists where the spacing between items in a category varied randomly, generally with the constraint that no two same-category items appeared adjacent to one another on the list. Dallet (1964) contrasted blocked and random presentation of categorized lists with 12 items each, with the number of categories represented in each list varied from 1 to 6. He found that recall was consistently greater in the blocked lists. He also measured the strength of category clustering, by calculating a deviation score indicating the number of category repetitions minus the number of repetitions based on random sampling of the number of items recalled. Clustering was also decreased in the random conditions, compared to the blocked conditions (Dallet, 1964). Cofer et al. (1966) also compared blocked and random presentation of categorized lists, and found that clustering was consistently enhanced for blocked presentation, while the effect of list organization on recall performance was less consistent. In his comparison of random and blocked lists, D'Agostino (1969) did not examine category clustering, but he did report increased overall recall, as well as increased items recalled per category. In general, researchers have consistently observed increased category clustering with blocked presentation, while recall is only increased about half the time (Puff, 1974).

Borges & Mandler (1972) went beyond the relatively coarse contrast of blocked vs. random presentation, manipulating the spacing between items in a category. They varied the within-category spacing from 0 (i.e. blocked presentation) to 12 (i.e. 12 items in between each item in a given category), with differently spaced categories placed together in the same list. There were 13 "test" categories in the list of 60 words, each with a different within-category spacing. In order to control for influences of serial position effects, they used items from 7 "filler" categories that were placed in the primacy positions (serial positions 1–6) and recency positions (the last 11 serial positions). Participants were informed of the categorical nature of the list, but not of the different spacing conditions. At the end of the list, partici-

pants performed written free recall, followed by a cued recall period where they were given the category names and asked to write down the items in each category. To examine recall performance, Borges & Mandler (1972) examined the number of categories recalled, as well as the number of items recalled within each category, given that at least one word had been recalled from that category. They found that the probability of recalling at least one item from a category increased monotonically with within-category spacing. Borges & Mandler (1972) suggested that this effect is related to the spacing effect observed when items are repeated within a list. Madigan (1969) found that the probability of recalling a repeated item increases with the spacing between the repetitions. The spacing effect may result from contextual variability (Lohnas et al., 2011). If items are associated with a context representation, and context changes slowly over time, then repetitions presented far apart in the list will be associated with highly distinct contexts. This means that the same repeated item can be accessed using a more diverse set of of context cues, making it more likely to be recalled (Lohnas et al., 2011). If items from different categories can be viewed as repetitions of some category representation, then that category will similarly be accessible through more contextual cues if there is a large within-category spacing, compared to if the items in the category are adjacent to one another (Borges & Mandler, 1972).

Although the relation between category recall and within-category spacing is fairly straightforward, Borges & Mandler (1972) found a more complex result when they examined within-category recall. They found that the items per category recall (IPC; the number of items recalled in a category, given that there was at least one recall from that category) varied non-montonically with within-category spacing. In both free and cued recall, IPC was greatest for small (0–3) and large (8–12) within-category spacing, and lowest for intermediate values (4–7). The decrease in IPC from small to intermediate within-category spacing is to be expected, given prior results. Glanzer (1969) found that recall performance decreased as the spacing between related pairs increased from 0 to 7. Furthermore, Borges & Mandler (1972) pointed out that experiments examining blocked vs. random presentation were actually ex-

amining a similar range of within-category spacing: blocked presentation corresponds to a spacing of 0, while "random" spacing, given the constraints on list design, has generally only been varied from 1 to 5. Therefore, the decrease in IPC observed between small and intermediate within-category spacing observed by Borges & Mandler (1972) may correspond to the decrease in recall often observed in random presentation compared to blocked presentation (e.g. D'Agostino 1969). In contrast, the finding that IPC is just as high for within-category spacing of 8–12 as it is for within-category spacing of 0–3 is surprising. Borges & Mandler (1972) suggested a possible explanation. When items in a category are presented nearby to one another, they are present in the short-term store at the same time, allowing participants to detect their common category, and this enhances encoding. When there is a large spacing between items in a category, they do not occupy the short-term store at the same time; therefore, presentation of the second (or third) item in the category triggers retrieval of the earlier presented items in that category, so that they may be rehearsed together. At intermediate spacings, participants may remember recently seeing another item from that category, and not bother to retrieve the other item so that it may be rehearsed with the current item. Borges & Mandler (1972), however, admit that this explanation is rather speculative.

Batchelder & Riefer (1980) proposed a somewhat different explanation for the non-monotonic relation between IPC and within-category spacing. They proposed that two factors influence later recall: (1) items presented nearby in time are likely to be grouped together in memory as a cluster, and (2) items presented further apart are less likely to be placed in a cluster, but any cluster that is formed is more likely to be recalled. However, they did not present simulations of the data described by Borges & Mandler (1972), so this explanation also remains somewhat speculative.

Although a number of researchers have manipulated within-category spacing in categorized free recall, few have controlled for a possible confound of temporal clustering. If items in a category are presented near to one another, then transitions between contiguous items will often be category repetitions. Therefore, the presence of temporal clustering in

recall may artificially inflate the estimate of category clustering on blocked lists compared to random lists (cf. Polyn et al. 2009; Morton et al. in press). Puff (1966) manipulated the number of category repetitions in the input order of a categorized list, with 10 words from each of 3 categories. The number of category repetitions was 0, 9, 18, or 27 (i.e. blocked presentation). He found evidence of monotonic increases with category repetitions, in both recall performance and category clustering. Puff (1966) noted that, if participants tended to follow recall of an item by recalling the item presented immediately after it, this might cause an artificial increase in category clustering scores in conditions with more category repetitions. To control for this, he calculated the number of category repetitions in recall that were also serial repetitions, and found that they accounted for less than 15% of category repetitions in each condition, suggesting that category clustering was much greater than that expected based on serial clustering. However, Puff (1966) only accounted for adjacent forward transitions; this accounts for the most substantial part of the contiguity effect, but fails to account for the graded drop-off of conditional response probability as a function of lag, in both forward and backward directions (Kahana, 1996). Future work should re-examine the effect of input list organization on category clustering, taking temporal organization into account using recently developed methods for measuring the contiguity effect (cf. Howard & Kahana 2002b; Polyn et al. 2009; Morton et al. in press).

Howard & Kahana (2002b) examined the relation between temporal and semantic organization during free recall of random words. As discussed in *Semantic clustering*, they examined conditional response probability at a function of semantic similarity (i.e., the LSA-CRP, based on cosine distance in LSA space), and found that response probability is greatest for related pairs of words. They also examined how this relation between semantic similarity and response probability varied as a function of the temporal distance between items on the list. They found that semantic similarity has a stronger effect on recall order when items were presented near to each other in the list. Howard & Kahana (2002b) also examined whether semantic organization is affected by the temporal spacing (as opposed to

item spacing) of the items in the list. They compared two *continual distraction* conditions, where participants performed a distracting math task in between presentation of items on the list. Each continual distraction list had either 2, 8, or 16 s of math distraction before and after each item. Howard & Kahana (2002b) found that the slope of the LSA-CRP declined as the duration of the distraction periods increased. In other words, semantic organization was greatest when items were presented near in time to one another. In contrast, they found that the effect of temporal contiguity on output order was not affected by the duration of the distraction periods (Howard & Kahana, 2002b). This suggests that temporal and semantic organization do not simply trade off between one another, as might be expected if the two forms of organization were supported by competing processes. Rather, there appears to be a more complex relationship between temporal and semantic organization (Howard et al., 2007).

2.2.6 False memory

While the presence of strong semantic associations between items on a list can lead to increased correct recall (relative to lists where the items are not strongly related; Cohen 1963b), these associations can also cause intrusions of incorrect items. Deese (1959b) examined whether the particular words intruded in a free-recall experiment can be predicted based on pre-existing associations. Participants performed free recall of lists composed of items associated with a critical stimulus. The associated items were taken from a separate norming study where participants were given the critical stimulus and asked to respond with the first word that came to mind. Deese (1959b) used the most frequently given responses to each of 36 different critical stimuli. For example, one list consisted of associates to the critical word "chair"; the items included "sofa", "wood", "cushion", "stool", "sit", etc. (the critical word was never included in the list). He then recorded the frequency with which each critical item was incorrectly recalled. In a separate study, Deese (1959b) had different participants free-associate to the items on each list, to determine how frequently each item

elicited the critical stimulus. He found that the frequency of critical intrusions was highly correlated with the average frequency with which the items on a list elicited the critical stimulus during free association. Furthermore, on the lists with high mean association to the critical word, the rate of intrusions was quite high, occurring up to 44% of the time. In contrast, lists with low associations to the critical word had very few intrusions; for example, there were no critical intrusions to the "butterfly" list.

Roediger & McDermott (1995) replicated the results of Deese (1959b) using the 6 lists that caused the greatest amount of false recall; there was a false recall rate of 40%, averaged over the lists. They also examined the output order of the words, and found that the critical word was usually recalled late in the recall period. Roediger & McDermott (1995) speculated that the previous recalls might make critical items more likely to be recalled. In a second experiment, they tested recognition of presented items, related items not presented on the list (related lures), and critical items (critical lures), using a remember-know procedure. Participants were presented with a target or lure, and responded whether the item was *old* (previously presented) or *new* (not presented on any of the lists). For items rated as old, they then rated whether they remembered the item (i.e., could recall the circumstances of the original presentation), or just knew that the item was on the list, without remembering the exact circumstances. For falsely recognized critical words, participants were almost as likely to make a remember judgment as for words that were actually studied. Therefore, participants did not just have a vague sense that the critical words were actually presented, but rather claimed to remember them being presented (Roediger & McDermott, 1995). One possible explanation is that participants thought of the critical word during encoding, and later misremembered the source of the word (i.e., they thought it was actually presented, when really they just thought about it; Roediger & McDermott 1995).

These results may be related to the finding in categorized free recall that items from the same category that are presented near to each other in the list are more likely to be clustered than if they are randomly placed in the list (Puff, 1974). Similarly, in the case of

the Deese-Roediger-McDermott (DRM) false-memory paradigm, related items are presented in a temporally contiguous fashion; that is, they are presented in the same list (Roediger & McDermott, 1995). There is evidence that words are sometimes intruded due to an association with a single word on a list (Zaromb et al., 2006); however, in the absence of common associates on the presented list, intrusions are rare (Deese, 1959b). This suggests that the false-memory effects observed in the DRM paradigm are enhanced by the temporal proximity of the items; if this is the case, then the false-memory phenomenon may be related to effects of within-category spacing on the strength of category clustering.

2.3 Models of semantic organization

Researchers have proposed a number of distinct mechanisms to explain how semantic associations influence behavior in free recall tasks. Some theorists have argued that recall performance in categorized free recall reflects recoding, where items in a category are recoded in terms of a higher-order category representation, such as a category name (e.g., Cohen 1963a; Rotondo 1977). Others have argued that categorized recall is best described in terms of associations, either between items and category representations (e.g., Bousfield & Cohen 1955), or between the items themselves (e.g., Sirotin et al. 2005). Finally, some models of semantic influences on recall do not assume activation of explicit category representations during encoding, but rather assume a blending of information related to individual items (Kimball et al., 2007; Socher et al., 2009); the resulting representations may serve a similar role to category representations in influencing recall behavior.

2.3.1 Recoding in categorized free recall

As discussed in *Memory availability versus accessibility*, some early influential accounts of behavior in categorized free recall focused on Miller's (1956) notion of recoding. Cohen (1963a) hypothesized that participants can effectively expand their memory of a categorized list by recoding the list in terms of chunks of items from the same category. Consistent with

this hypothesis, Cohen (1963b) showed participants could recall more items from categorized lists than from lists composed of unrelated items. There is also evidence that recall of categories (that is, recall of at least one item from a category) is independent from recall of individual items within a category. These measures of recall are differentially affected by manipulations of list length (Tulving & Pearlstone, 1966; Cohen, 1966) and within-category spacing (Borges & Mandler, 1972). Furthermore, researchers have demonstrated that within-category IRTs and between-category IRTs evolve in different ways over the course of recall (Pollio et al., 1969; Patterson et al., 1971), and are differentially affected by the provision of category cues during recall (Patterson et al., 1971). For recoding to occur, presumably participants must notice the categorized structure of the list, so that items in the same category can be grouped together into a chunk. Researchers have proposed that, for a participant to take advantage of an association between a pair of words (including noticing that they are members of the same category), the words must be in short-term memory at the same time (Puff, 1974). Consistent with this hypothesis, both recall (Glanzer, 1969) and category clustering (D'Agostino, 1969) are increased when related items are presented adjacent to one another on the list.

The recoding process might involve encoding a category name, rather than each item in the category (Cohen, 1963a; Cofer et al., 1966). However, even if one recalls a category name successfully, there is no guarantee that all of the items in that category can be recalled on the basis of this category cue. Cohen (1963a) illustrated the later issue by examining a distinction between exhaustive (E) and non-exhaustive (NE) categories. For example, an E category might include the cardinal directions (*EAST*, *WEST*, *NORTH*, *SOUTH*), while a NE category might be four-footed animals (e.g., *BEAR*, *DOG*, *DEER*, *ELEPHANT*). Participants generally recalled all items from E categories if they remembered any items from that category. Reliably fewer items were recalled from each recalled NE category (Cohen, 1963a). Later experiments with NE categories confirmed that recall from categorized lists follows a "some-or-none" pattern, where either no items are recalled from a given category, or

a substantial proportion of items are recalled from that category (Cohen, 1966). Evidently, chunking is most effective when the items in a category can be completely replaced by a category label without loss of information (as is the case for E categories, but not NE categories). Although simply free-associating to the category name may be an effective strategy in the case of E categories, it is clear that participants do not generally adopt this strategy in categorized free recall of items from subsets of larger categories: Intrusions of extralist items that belong to one of the categories on the list are very infrequent relative to correct recalls (Cofer et al., 1966). Therefore, recall of items within an NE category must rely on some information in addition to the category name. The use of associations between a representation of list context and items on the list has been proposed as a critical source of information to focus recall on the list items (Raaijmakers & Shiffrin, 1980, 1981; Howard & Kahana, 2002a). List context may influence recall either by serving as a memory cue to only retrieve items studied in that context (Raaijmakers & Shiffrin, 1980; Howard & Kahana, 2002a), or by facilitating correct recognition of candidate items that are generated by free-associating to the category name (Bahrick, 1971).

Rotondo (1977) developed a quantitative framework to describe chunking in free recall. The framework uses discrete structural equation systems to characterize recall in terms of chunks and the items within each chunk. Unlike measures of category clustering, which focus on properties of recall order, the framework ignores output order completely, focusing instead on the co-occurrence of items in recall. For example, recall of two items in the same chunk should be somewhat dependent, since recalling either item depends on the chunk first being recalled; as a result, items in the same chunk should tend to be recalled or forgotten together. In contrast, recall of items in different chunks should, according to the assumptions of the framework, be independent of one another. In order to apply the framework, the experimenter must first specify which items may be chunked together (such as items in the same taxonomic category). The framework assumes that there are parameters indicating *chunk strength*, which determines the probability of a given chunk being recalled.

For each item in a chunk, there is a *membership strength* parameter, which determines the probability of recalling each item given that the chunk is recalled. A specific model (with a given chunk structure) can then be fit to recall performance data to estimate the chunk strength and membership parameters. Furthermore, different models with different types of chunks can be evaluated for their ability to fit the data (Rotondo, 1977). For example, one can contrast a model where there are several chunks for items in different categories with a model where each item stored in its own chunk. Rotondo (1977) examined co-occurrence data from a free-recall experiment with 5 items from each of 10 categories, and showed that a model with different independent chunks for each category fit better than a model with separate chunks for each item. Furthermore, Rotondo (1977) found that the category-chunk model provided a good fit to most categories, consistent with the assumptions that recall of different categories is independent and that recall of items within a category is independent.

The model of Rotondo (1977) allows estimation of the tendency for individual categories and items within those categories to be recalled; however, it is silent on the mechanisms underlying recoding and subsequent recall. Raaijmakers & Shiffrin (1980) used the SAM framework, which has been successfully applied to many results in free recall, to implement a simple model that proposes an account of retrieval dynamics in categorized free recall. Previous simulations using SAM relied on two types of associations: interitem and context-to-item associations. Context-item associations are important for focusing search of the long-term store (LTS) on the current list. Interitem associations further guide search, giving rise to the contiguity effect (Kahana, 1996). In simulating categorized free recall, Raaijmakers & Shiffrin (1980) included category-item associations. They assumed that, during encoding, each item becomes associated with a representation of its category. The initial category-item associations are set to 0; the associations are assumed to be learned entirely during the study period. Additionally, there are residual associations between each item and every other category; this is important to explain output interference in cued recall (Smith, 1971). To investigate results in categorized free recall experiments, which have focused on category

organization and recall rather than serial position effects, Raaijmakers & Shiffrin (1980) used a simplified version of the model where no interitem associations are learned, only context-item and category-item associations. During free recall, first an item is retrieved using the context cue. The recalled item then retrieves its category representation; then both context and the category are used to cue LTS. After a fixed number of failures to retrieve a new item, the category cue is dropped, and context alone is used to search until a new category is found. Cued recall procedes in a similar fashion, except that category cues do not need to be searched for since they are supplied by the experimenter. When an item is recalled, the associations between it and the cues used to retrieve it are incremented by a fixed amount.

Using this simple model, Raaijmakers & Shiffrin (1980) were able to account for the several manipulations examined by Tulving & Pearlstone (1966). In both the model and the data, recall is greater in cued recall; in uncued recall, some categories will not be retrieved and used as cues, while in cued recall every category will always be cued. The advantage for cued recall decreases as the number of items per category increases (from 1 item per category to 4), since larger categories are more likely to be recalled in uncued recall. In contrast, the advantage for cued recall increases with list length, since the probability of accessing a given category decreases in uncued recall as more categories are added to the list (Raaijmakers & Shiffrin, 1980). Furthermore, the model accounts for the fact that percent recall within each category declines as category size increases; this is because search within a category procedes with replacement, so that earlier recalled items interfere with recall of new items (Raaijmakers & Shiffrin, 1980).

Although there is evidence that recall within a category is mostly independent of items in other categories (e.g., within-category recall is not affected by changes in list length; Cohen 1966), Raaijmakers & Shiffrin (1980) found that the residual category-item associations are important for explaining output interference effects. The residual associations influence recall, so that a given category cue will mostly support items in that category, but it will also somewhat support items in other categories. In other words, the model assumes that

a given category cue does not perfectly target only items from that category. When, as in Smith (1971), successive categories are probed with category names, incrementing of associations causes recalled items to become more strongly associated to the list context. As a result, search attempts during recall of later categories will experience more interference from words that have already been recalled. This causes recall to decrease as a function of category output position; using best-fitting parameters based on the Tulving & Pearlstone (1966) data, SAM provides a good fit to the rate of decrease in recall performance observed by Smith (1971) (Raaijmakers & Shiffrin, 1980).

Although the model of Raaijmakers & Shiffrin (1980) provides a detailed theory of retrieval dynamics in categorized free recall, it provides no predictions for why a given item in a category is more likely to be recalled than others. This is because it relies purely on associations learned during the experiment, without taking pre-existing associations into account. In order to explain why some items in a given category are recalled, but not others, several researchers have explained recall behavior in categorized free recall in terms of pre-existing associations between individual items and higher-order category representations. Using this framework, some aspects of recall performance can be predicted from free-association data. The frequency with which participants generate a given exemplar when given a category cue is called taxonomic frequency; this variable is an important determinant of performance in categorized free recall. Recall is better for words that have higher taxonomic frequency, both when the words occur in lists with multiple categories (Bousfield & Cohen, 1955; Bousfield et al., 1958; Cofer et al., 1966) and when they occur in lists with only one category (Deese, 1959a). These results suggest that variations in the strength of pre-experimental associations have an important role in influencing free recall.

Models of categorized free recall typically embed assumptions about recall strategy. For example, Raaijmakers & Shiffrin (1980) assume that a category cue is used to attempt recall of individual words until a fixed number of retrieval failures occur. The threshold for giving up on a category cue and searching for a new category is a free parameter of the model,

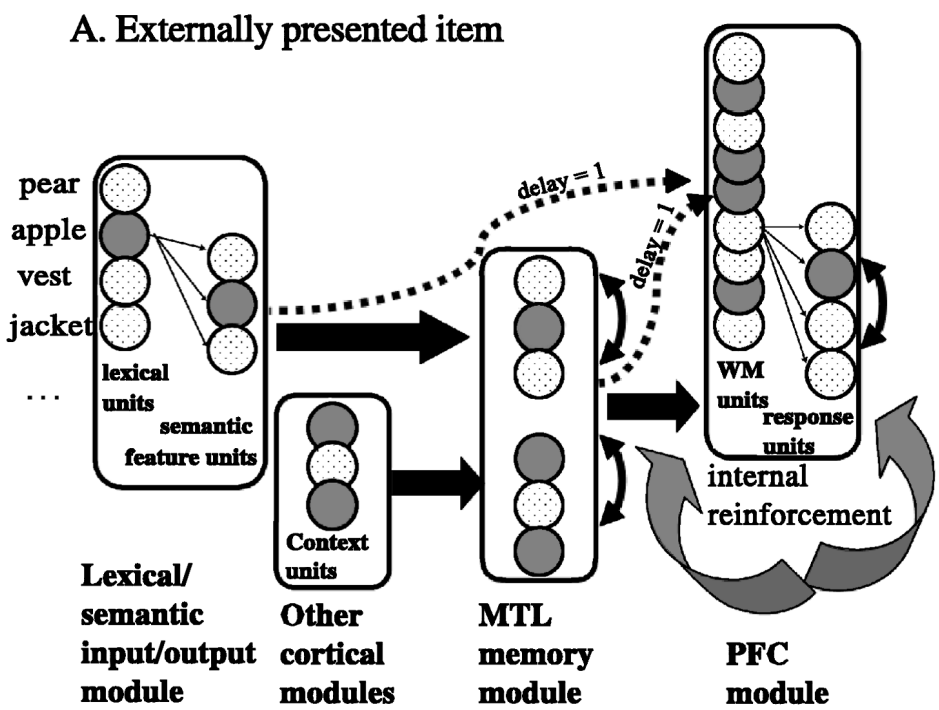


Figure 3. A model of strategic organization in free recall. Adapted from Becker & Lim (2003).

which the experimenters can change to make the model's recall performance approximately the same as that of actual participants. This approach avoids the question of how participants determine what the threshold should be set to, or more generally, how they determine what rule they should use to decide when to switch categories in order to maximize their overall recall and minimize repeats and intrusions.

Becker & Lim (2003) began to address this issue by developing a neural network model that uses rapid reinforcement learning to adaptively guide memory search (Fig. 3). The model assumes that items within a category are associated with similar patterns in a semantic feature layer. Another layer represents the current context; as in a number of models of episodic memory, context changes slowly over time (Estes, 1955; Mensink & Raaijmakers, 1989; Howard & Kahana, 2002a). During encoding, the context and semantic feature layers project to a medial temporal lobe (MTL) memory module, which stores memories of "episodes" including item and context information. Information about the current item and

the current context are represented in a working memory layer, which projects to a response layer. In the response layer, only one unit is active at a time; this unit becomes associated with the MTL layer, allowing the response layer to guide later retrieval. In simulations of categorized free recall, the units in the response layer can take on a similar role to the category name cues used in SAM (Raaijmakers & Shiffrin, 1980); although there is no explicit coding of categories in the model, the response units are sensitive to the semantic similarity structure of presented items, so that items in the same category tend to activate the same response unit.

At the start of recall, the last item on the list and its associated activity in MTL projects to the working memory layer, which then activates a response unit. This response layer activity projects back to the MTL, where an associated episode is retrieved, reactivating context and semantic information. The semantic information drives a competition among units in a lexical layer representing individual items, causing one to be retrieved. Response suppression in the lexical layer is used to prevent recently recalled items from being repeated. After an item has been retrieved, its associated semantic features and the current context are fed into MTL; in the MTL module, a measure of network activity called *harmony* is calculated to determine how well the item and context pattern matches existing memory traces. If the harmony is too high, the retrieved item is assumed to be a repeat; if the harmony is too low, the retrieved item is assumed to be an intrusion of an item not on the list (Becker & Lim, 2003).

When the model detects a repeat or intrusion, negative reinforcement learning occurs in the PFC module, making the current response unit less likely to be used in the immediate future. If the just-recalled word is not recognized as a repeat or intrusion, it is assumed to be a correct recall, which triggers positive reinforcement learning in the PFC. In this way, response units that tend to produce correct recalls will be used more often, and response units that tend to cue for repeats or intrusions will be used less often. As a result, the model can learn that it is optimal to keep using certain response units until they start triggering

repeats or intrusions; at that point, it is best to shift to a different response unit (Becker & Lim, 2003). Since response units tend to correspond to different categories, learning can manipulate the amount of category clustering. Category clustering tends to increase with learning, since it provides for a relatively efficient search of items on the list. Therefore, the model suggests that increases in clustering may be driven by a tendency to continue cuing with a category for a longer period; this is consistent with the finding that increases in clustering with practice correlate with increases in between-category IRTs (Thompson, 1978). Becker & Lim (2003) also examined a version of the model with reduced connections in the PFC module, in order to simulate the effects of PFC damage in patients; they found that these virtual lesions cause reduced category clustering, similar to that observed in frontal patients (Stuss et al., 1994).

2.3.2 Interitem associations

Although the use of higher-order associations allowed a version of the SAM model to account for a number of findings in categorized free recall (Raaijmakers & Shiffrin, 1980), it is unclear whether this model could account for the influence of subtle semantic associations that is observed in recall of "unrelated" words, which are not clearly divided into separate taxonomic categories. Sirotin et al. (2005) attempted to characterize both subtle semantic clustering and category clustering solely in terms of existing interitem associations. The original version of SAM (Raaijmakers & Shiffrin, 1981) only accounts for associations between items that are learned during the experiment. Sirotin et al. (2005) added another set of semantic associations between items, which do not change during the experiment; the resulting modified version of SAM is referred to as eSAM. To obtain a realistic model of semantic associations, they used interitem similarity estimates based on LSA (Landauer & Dumais, 1997) in one version of the model, and WAS (Steyvers et al., 2004) in another version. They set the associative strength between each pair of items based on the cosine similarity of the high-dimensional representations of the items.

During recall, contextual, episodic, and semantic cues are combined to determine which item is recalled next. Critically, the model uses a multiplicative rule to combine the cues in order to determine the total cue strength for each item in memory (Sirotin et al., 2005). To see the importance of this multiplicative rule, consider a case where an item has just been correctly recalled from the list. Both the list context and the item will then be used as cues to attempt retrieval of another item from the list. The item cue influences recall both through episodic and semantic associations. The item might be strongly semantically related to another item that was not presented on the list, which would therefore be an intrusion if recalled. If the context, episodic, and semantic cue strengths were combined additively to determine the total cue strength, then a very strong semantic association could overwhelm the context and episodic strengths (which should be low, given that the item was not presented on the list), causing the item to be intruded. By using a multiplicative rule for combining the different types of associations, the model effectively filters out any words that are not strongly cued by all three cue types. This allows recall to be strongly influenced by semantic associations (causing category clustering), without leading to an excessive amount of intrusions (Sirotin et al., 2005). The influence of each of the context, episodic, and semantic cues is scaled by separate free parameters of the model, allowing temporal and semantic clustering to be scaled separately.

To test the ability of eSAM to account for semantic organization in lists of random words, Sirotin et al. (2005) simulated an experiment by Howard & Kahana (2002b). They found that eSAM provides an excellent quantitative fit of conditional response probability as a function of semantic similarity (measured either by LSA or WAS). Using the same parameters, the model simultaneously provides a good fit of recall probability as a function of serial position, conditional response probability as a function of lag, and the probability of starting recall with each serial position (Sirotin et al., 2005).

In addition to examining the effect of semantic similarity between pairs of random items, Sirotin et al. (2005) also examined category clustering in a categorized free recall experiment

(Kahana & Wingfield, 2000). They found that a version of the model with semantic associations based on LSA greatly underpredicted the amount of category clustering, because LSA does not capture differences between taxonomic categories with high fidelity. The version of the model with semantic associations based on WAS provided a better fit, though the amount of category clustering was still somewhat lower than that observed in the data (Sirotin et al., 2005).

It remains to be investigated whether eSAM can account for effects of manipulating list length (Cohen, 1966), category size (Tulving & Pearlstone, 1966), and within-category spacing (Puff, 1974). In particular, it will be important for future work to address whether a model such as eSAM can account for findings that have been interpreted as evidence for recoding (e.g., Cohen 1963a). Whereas recoding theories of categorized free recall assume that items become associated to higher-order representations, eSAM has no higher-order representations and instead relies on interitem associations. In contrast to recoding theories, which assume that retrieval procedes by first retrieving a category representation, then using it as a cue to retrieve items, eSAM assumes that there is only one cuing step for each retrieval attempt. Therefore, in the eSAM framework, clustering by taxonomic category (e.g., Bousfield 1953) is not treated as a distinct phenomenon from clustering by more subtle variations in semantic associations (e.g., Howard & Kahana 2002b), but rather both types of clustering are treated as occurring on a continuum of associative strength in semantic memory. The approach of eSAM contrasts with models that assume that items in a category are grouped together into a discrete chunks (Rotondo, 1977; Raaijmakers & Shiffrin, 1980). Further investigation is required to determine which of these framework provides a better account of the critical data from categorized free recall experiments.

The framework of retrieved-context models has also been extended to account for semantic organization (Polyn et al., 2009). The context maintenance and retrieval (CMR) model builds on the temporal context model (TCM; Howard & Kahana 2002a; Howard 2004; Howard et al. 2005; Sederberg et al. 2008) to simultaneously account for temporal, source,

and semantic influences on recall (Polyn et al., 2009). To account for semantic organization in free recall of random words, Polyn et al. (2009) assumed that connections between the context layer and the feature layer reflect longstanding semantic associations between items and their associated states of context. When an item is recalled, it retrieves its associated context; this state of context then projects through the context-to-item connections, causing items that are semantically related to the just-recalled item to become activated on the feature layer. Thus, items that are semantically related to the just-recalled item have greater activation and are more likely to win the competition to be recalled (driven by the leaky, competing accumulator model; Usher & McClelland 2001) than items that are unrelated. CMR successfully accounts for semantic clustering, while simultaneously providing a good fit to primacy, recency, and contiguity effects, as well as clustering by source context (Polyn et al., 2009). In contrast to eSAM, which relies on separate context, temporal, and semantic cues (Sirotin et al., 2005), CMR assumes that recall is driven by a single context cue that simultaneously constrains memory search to be influenced by temporal and semantic information[2]. Although CMR provides a good account of basic semantic clustering, it remains to be seen whether the behavior of the model is consistent with the pattern of results obtained in categorized free recall, or with findings suggesting that temporal and semantic organization interact with one another (e.g., Borges & Mandler 1972; Howard & Kahana 2002b)

2.3.3 Hybrid models

While some models of semantic clustering in free recall rely purely on higher-order category representations to guide recall (Rotondo, 1977; Raaijmakers & Shiffrin, 1980), and others assume that recall is guided by interitem associations (Sirotin et al., 2005; Polyn et al., 2009), there are also models of free recall that focus on item-related memory cues, but allow them

[2]The source context cue, which is represented in a separate subregion of context that is allowed to evolve somewhat differently than the temporal subregion, may be best thought of as a separate cue that may vary independently of the temporal and semantic cues (but see Polyn et al. 2012).

to combine to effectively create higher-order representations on the fly. The central idea in each of these "hybrid" models is that processing during encoding or retrieval of an item is dependent on the surrounding context. For example, several theories of within-category spacing effects in categorized free recall assume that sustained item-related activity during encoding sums over multiple items, so that category-specific information is emphasized when items from the same category are presented in succession (Puff, 1974).

Kimball et al. (2008) used the SAM framework to examine behavior in the false-memory paradigm, which critically depends on context-sensitive processing of items. First, they examined how the eSAM model performs when applied to a false-memory task. They found that, in order to account for the number of critical-word intrusions, they had to increase the weighting of semantic cues to the point that non-critical intrusions were also common. Kimball et al. (2008) then explored several variants of eSAM to determine what additional mechanisms are necessary to account for data showing high rates of critical-word intrusions and low rates of other intrusions (Deese, 1959a). They examined two proposed mechanisms: (1) spreading activation during encoding, and (2) compound cuing during recall.

They examined a spreading-activation mechanism operating during encoding that was proposed by Puff (1966) and others. In this mechanism, a presented item causes activation of related items in a semantic network (cf. Collins & Loftus 1975). In addition to learning associations between the presented item and list context, the model assumes that related items (partially activated due to spreading activation) also become associated to context (Kimball et al., 2008). They explored 3 variants of this mechanism: (1) the associates of each individual item are learned; (2) activation spreads from each of the items currently in the short-term store (STS), and each associated item is learned proportional to its total activation; and (3) activation is related to each of the items in the STS, and the activation of each related item is combined multiplicatively. When the spreading-activation mechanism operates over the multiple items held in the buffer, items that are related to each of the items will tend to be activated the most. When the items all correspond to a similar theme

or category (e.g. different types of furniture), then high-frequency associates to those items will tend to be activated (e.g. *CHAIR*).

Kimball et al. (2008) also examined a potential mechanism operating during retrieval. In eSAM, each recalled item is used as a cue to attempt retrieval of more items from the list. In fSAM, Kimball et al. (2008) examined the possibility of compound cuing, where multiple items are simultaneously used as cues. They examined 3 variants of semantically guided retrieval: (1) only the just-recalled word is used as a cue, and it cues for semantically related items (this is the same as the cuing used in eSAM); (2) all recalled items in the STS are used as cues, and their cue strengths are combined additively; and (3) multiple item cues are used, and their cue strengths are combined multiplicatively.

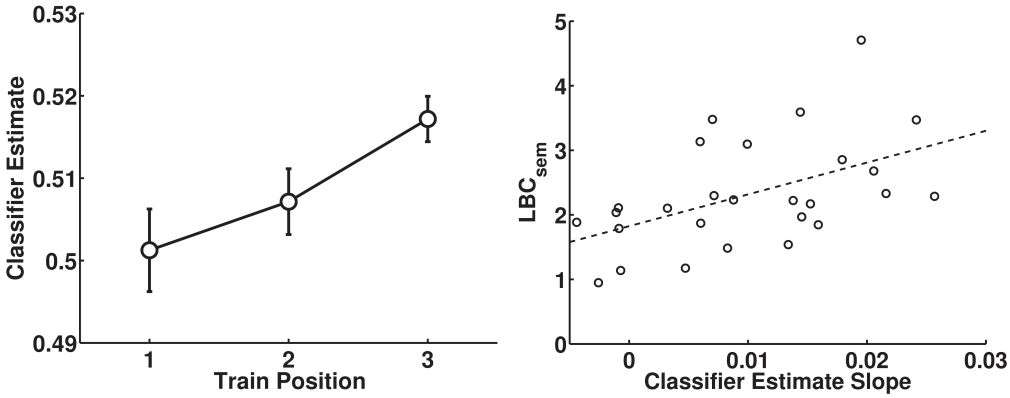
They found that both spreading activation during encoding and compound cuing during retrieval was necessary to focus the most cue strength on items that were related to several of the items in the list, rather than just one; this increased recall of the critical item without making other intrusions too frequent (Kimball et al., 2008). They found, for both mechanisms, that multiplicative combining across items (of spreading activation during encoding, or cue strength during retrieval) provided a better fit than additive combination. Given that fSAM can use these mechanisms to increase the number of critical intrusions during recall, the same mechanisms might allow SAM to make better predictions regarding the degree of category clustering observed in categorized free recall, which is underpredicted somewhat by eSAM (Sirotin et al., 2005).

Although CMR (Polyn et al., 2009) assumes that semantic similarity between items only has an effect during retrieval, recent work suggests that temporal context active during encoding should be sensitive to semantic information abut the studied items. Rao & Howard (2008) examined semantic learning over time in TCM; as discussed in more detail in Section 3, they exposed a version of TCM to a large text corpus. They found that, over time, TCM learns about co-occurrence structure of material it has been exposed to. Items that are semantically similar come to elicit similar states of context. This contrasts with the

assumptions typically used to simulate free recall using TCM or CMR, where the states of context elicited by items are assumed to be orthogonal. If, instead, semantically similar items retrieve similar states of context, then CMR predicts that semantic information about items will be integrated into context during encoding (Morton et al., in press). Therefore, a version of CMR incorporating this assumption will exhibit sustained activity in context reflecting the semantic similarity of studied words (Morton & Polyn, 2012); this may have a similar effect to the spreading-activation mechanism used in fSAM.

Socher et al. (2009) examined a version of TCM that incorporates semantic information in context. Rather than assuming a TCM-like semantic learning mechanism (as suggested by Rao & Howard 2008), however, they incorporated representations derived from the topic model (Steyvers et al., 2004). Socher et al. (2009) assumed that presentation of an item causes retrieval of the distribution of that item over topics (see *Non-metric models* in Section 1); this topic information is then integrated into context. Over the course of the list, context changes slowly to reflect the gist of the set of presented words. As in standard TCM (Howard & Kahana, 2002a), context continues to evolve during the recall period, as context associated with recalled items is retrieved. Socher et al. (2009) examined the distinct contributions of episodic context (that is, blends of context developed during encoding) and semantic context (distributions over topics retrieved by recalled items during recall). They found that the episodic influence was necessary to focus recall on the current list, and keep intrusions low. The semantic influence also served a clear role; when only episodic cuing was used, the model made no intrusions. Although Socher et al. (2009) demonstrated that their model, LDA-TCM, produces reasonable recall behavior in free recall of random words, they did not examine whether it can account for semantic clustering.

Morton et al. (in press) found neural evidence for the type of sustained activation that theorists have proposed as being vital for explaining within-category similarity effects and false-memory effects (Puff, 1974; Kimball et al., 2008). Using scalp EEG, they measured oscillatory power at a range of frequencies, while participants performed a categorized free



(a) (b)

Figure 4. Sustained patterns of oscillatory power during encoding predict subsequent category clustering. (**a**) The persistence of category-related neural patterns is seen in the increased fidelity of category patterns (as measured by multivariate pattern analysis) when multiple same-category items are presented in succession. The classifier's estimate of the strength of the current category is plotted as a function of position within a train of same-category item presentations. On average, classifier estimates rose with successive same-category stimuli. Error bars represent 95% confidence intervals based on within-subject error (Loftus & Masson, 1994). (**b**) The slope of the regression of classifier estimate on train position was correlated with individual differences in category clustering as measured by $\mathrm{LBC}_{\mathrm{sem}}$ ($r = 0.500$, $p < 0.01$). Two outliers have been removed from the plot; with them included, the correlation is still significant ($r = 0.421$, $p < 0.05$). Figures reproduced from Morton et al. (in press).

recall task with categories that have distinct neural representations (celebrities, landmarks, and objects). They examined oscillatory power during presentation of each item, and used multivariate pattern analysis (Norman et al., 2006) to examine category-specific activity. This analysis provides an estimate of how active each category is during each item, based on the measured oscillatory power. They found evidence that category-specific activity is integrated over time: As different items from the same category are presented in succession, the classifier evidence for that category increases (Fig. 4a; Morton et al. in press). Furthermore, this activity is related to subsequent category clustering. The rate of increase in classifier estimates predicts individual differences in category clustering (Fig. 4b), as measured by LBC_{sem} (Stricker et al., 2002). These findings are consistent with theories proposing that sustained activation during encoding is important for supporting subsequent clustering (Puff, 1974).

2.4 Conclusions

Many researchers have argued that behavior in categorized free recall reflects a two-stage process: recall of category names, followed by recall of items from within that category (Cohen, 1963a; Tulving & Pearlstone, 1966; Raaijmakers & Shiffrin, 1980). However, unlike interitem association models, models based purely on category-item associations do not seem well-equipped to account for clustering based on non-obvious variation in semantic associations (Howard & Kahana, 2002b). Few researchers have attempted to account for category clustering with solely interitem associations; there is some evidence that interitem associations can lead to both category clustering (Sirotin et al., 2005) and semantic clustering of material with no obvious semantic structure (Polyn et al., 2009). However, models that focus on interitem associations have not been applied to many of the important findings in categorized free recall experiments; future work should examine whether these models can account for effects of list length, category size, and cued vs. uncued recall manipulations. If models relying on interitem associations are sufficient to explain category clustering, then

they should be favored in terms of parsimony, relative to models that assume that different processes or representations are engaged in categorized free recall and standard free recall.

While there is a great deal of evidence from categorized free recall and false memory paradigms that the proximity of associated items is an important determinant of recall performance, few formal models of free recall have accounted for these findings. Simulations of false memory paradigms using the fSAM model suggest that spreading activation during encoding is an important mechanism to consider in free recall (Kimball et al., 2008). Similar elaborations of the retrieved-context model framework are beginning to consider the impact of integrative processing of stimuli during encoding and retrieval (Socher et al., 2009; Morton & Polyn, 2012). Proposed mechanisms determining the role of semantic associations within the SAM and retrieved-context model frameworks can be highly constrained by data in both categorized free recall and false memory paradigms. By investigating formal models of these distinct tasks, where recall is either facilitated or hurt by semantic organization, it will be possible to gain increased insight as to how prior semantic memory influences new episodic learning. Understanding of how existing knowledge and biases influence the formation and retrieval of new memories will have important consequences for legal, educational, and clinical applications.

3 Models of declarative memory

Although models of memory have typically focused on explaining either semantic or episodic memory, some researchers have begun to develop models that attempt to describe the entire declarative memory system that subserves both episodic and semantic memory. The influential complementary learning systems (CLS) framework (McClelland et al., 1995; O'Reilly & Norman, 2002; Norman & O'Reilly, 2003) proposes that semantic and episodic memory have different functions that require distinct neural architectures and therefore involve different brain regions. However, recent work suggests that the putative episodic memory system

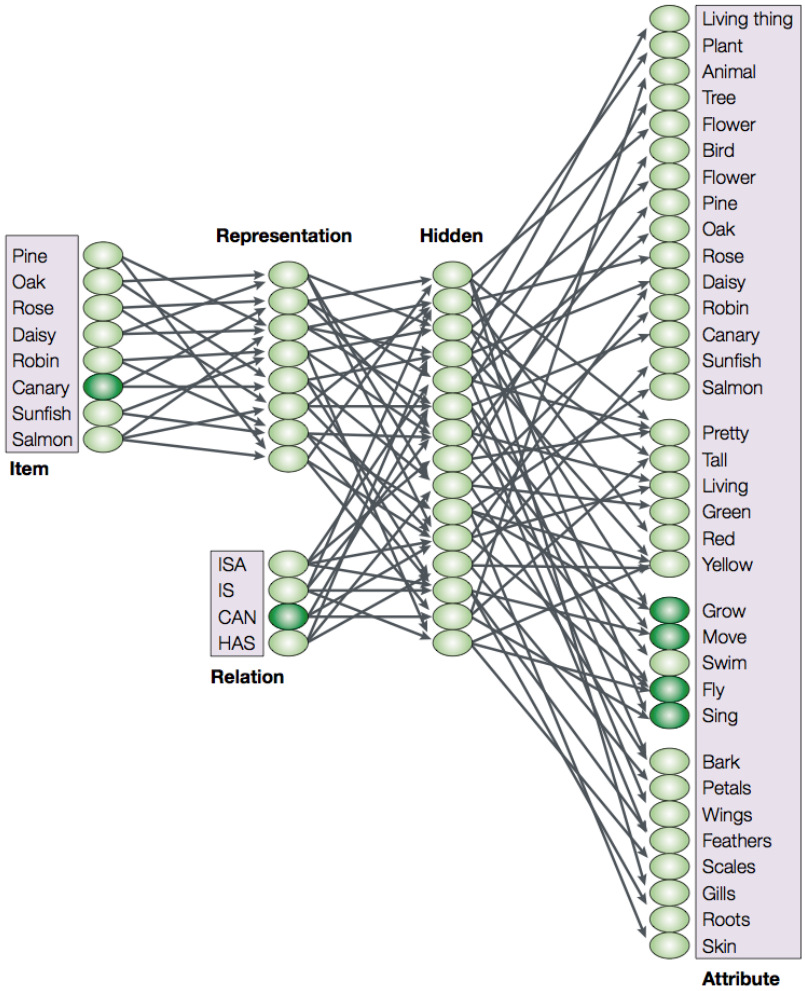


Figure 5. The Rumelhart connectionist model of semantic learning. Reproduced from McClelland & Rogers (2003).

in the medial temporal lobe is also involved in performing inference, a critical component of semantic learning (Kumaran, 2012). In contrast to the CLS framework, the temporal context model (TCM) proposes that both episodic and semantic learning involves similar mechanisms (Howard et al., 2011), raising the possibility of developing a unified model of declarative memory. The following sections discuss each of these approaches to understanding the relation between episodic and semantic memory.

3.1 Complementary learning systems

McClelland et al. (1995) proposed that episodic and semantic memory have opposing goals: semantic memory requires slow learning in order to extract statistical regularities in the environment, while episodic memory requires fast learning and storage of events as distinct episodes that are separate from one another even if they are very similar. They proposed that this functional difference between episodic and semantic memory requires distinct networks that are optimized for these distinct tasks: Episodic memory is subserved by the hippocampus, while semantic memory is accumulated slowly in the cortex (McClelland et al., 1995).

The Rumelhart model of semantic learning is consistent with the idea that semantic memory must be built up slowly using interleaved learning. The model uses error-driven backpropagation learning to learn facts about objects such as "robins have wings" (Rumelhart, 1990). The model learns through supervised learning that, for example, when the "canary" unit is active, and the "CAN" relation unit is active, then the attribute units "grow", "move", "fly", and "sing" should be active (Fig. 5). The challenge for the model is to not only efficiently learn a set of relations for a number of different items, but also to learn general relations (e.g., that all birds have wings). With slow, interleaved learning, the model learns to associate similar items (i.e., items that mostly have the same properties) with similar activity on the representation layer; for example, birds elicit similar patterns, while trees elicit a distinct set of patterns. The Rumelhart model is able to generalize well to new learning, such as learning about properties of "sparrow" after learning about "robin" and "canary" (McClelland et al., 1995). Because "sparrow" shares properties with other birds, it will come to elicit a similar pattern to other birds on the representation layer. Given this representation, the model will be able to infer some facts about "robin" even if it has not been trained on them specifically; for example, after learning other properties of "robin", activating "robin" and "HAS" will cause activation of the "wings" unit on the attribute layer. However, extracting statistical regularities such as sets of items with shared proper-

ties requires learning to procede very slowly. If new, conflicting information is learned too quickly, it results in catastrophic interference that disrupts previous learning (McClelland et al., 1995).

In contrast, episodic learning involves encoding episodes that share many features, and yet are stored separately; for example, one can (often) remember where one's car is parked, without interference from previous days that were similar in many respects (Kumaran & McClelland, 2012). O'Reilly & McClelland (1994) demonstrated that the properties of the hippocampus make it well-suited for the task of storing episodes while minimizing interference.

McClelland et al. (1995) proposed that learning in the cortex is accomplished slowly over time through a process of consolidation. Consolidation involves the hippocampus training the cortex slowly over time. McClelland et al. (1995) used this proposed mechanism to explain the finding of retrograde amnesia in patients with medial temporal lobe (MTL) damage. Memories from up a few years before the damage may be affected; this suggests that memories are initially hippocampally dependent, but gradually are stored in the cortical network, making them hippocampally independent (McClelland et al., 1995).

3.2 REMERGE

Although the Rumelhart model is able to account for a number of findings in semantic learning, such as the development of concepts in children and the effects of semantic dementia (McClelland & Rogers, 2003), there is evidence that people can perform inference on a much shorter timescale than is typically assumed by models of semantic memory (Preston et al., 2004). This faster form of inference has been demonstrated in transitive inference, paired-associate inference, and acquired equivalence paradigms (Kumaran, 2012). This review will focus on the paired associate inference (PAI) paradigm, because data from this paradigm have also been simulated using TCM, as discussed in the next section. In the PAI paradigm, participants are first presented with a number of stimulus pairs. In the first stage, they are

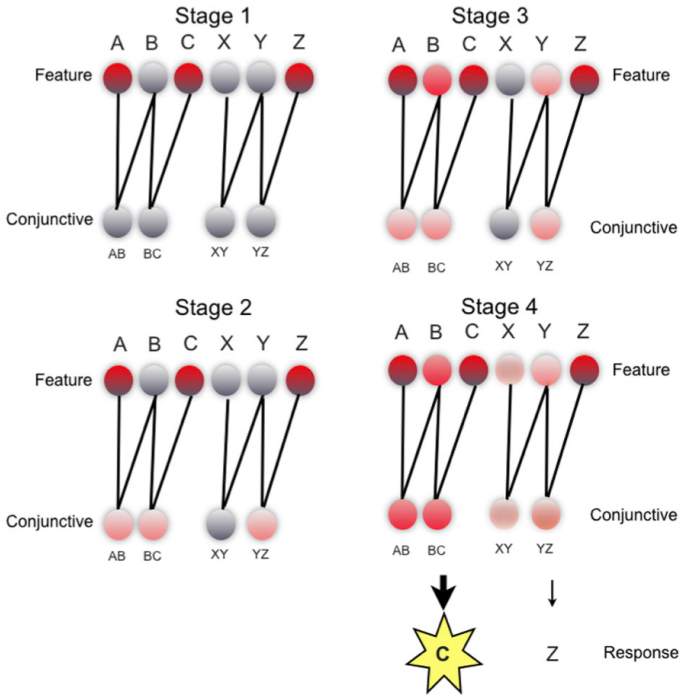


Figure 6. Illustration of recurrent activation in REMERGE during the test phase of a paired-associate inference task. Reproduced from Kumaran (2012).

presented with pairs of items A-B and X-Y. In the second stage, they learn overlapping pairs B-C and Y-Z. At test, participants are shown an item (e.g. A), and must choose between two responses (e.g. B or X). After the study periods, participants perform two types of test: premise pair testing of pairs that were previously studied (e.g. A is presented, and the participant chooses between B and X), and pairs that were not studied, but were studied in overlapping pairs. For example, a participant might be presented with A and asked to choose between C and Z. If participants are using inference, then they will choose C (which is related to A through the mediator B). Both rats (Bunsey & Eichenbaum, 1996) and humans (Preston et al., 2004) demonstrate inference in this paradigm.

Interestingly, there is evidence that behavior in the PAI task is dependent on the hippocampus: Hippocampally lesioned rats are able to learn the premise pairs, but not the inference pairs (Bunsey & Eichenbaum, 1996). Kumaran & McClelland (2012) suggested that recent findings of hippocampal anatomy may shed light on its role in fast inference. Previous models of hippocampal anatomy have assumed that activation spreads through the

hippocampal subregions in one direction, starting in the superficial layers of the entorhinal cortex (EC), propagating through the dentate gyrus (DG), CA3, and CA1, with retrieved memories activated in the deep layers of the EC (O'Reilly & McClelland, 1994). However, new evidence suggests that the deep and superficial layers of the EC are connected, allowing activity in deep EC to propagate again through the hippocampal system (Kumaran & McClelland, 2012). Kumaran & McClelland (2012) propose that this recurrence is integral to the role of the hippocampus in fast inference. They developed the REMERGE (recurrence with episodic memory results in generalization) model, which presents a simplified model of recurrent activity in the hippocampus. The model is based on a class of interactive activation and competition (IAC) models that use multiple distributed representations that are bidirectionally connected in order to allow satisfaction of multiple constraints simultaneously (Kumaran & McClelland, 2012).

REMERGE primarily involves two distributed representations that have bidirectional connections between them. There is a feature layer, which can be used to represent items (e.g. A or B in the PAI task). A separate conjunction layer represents pattern-separated events stored in the hippocampus (e.g. the studied AB pair). In addition, a response layer contains the possible responses; associations between the conjunction layer and the response layer allow the network to learn the correct response to a probe item. For example, if A-B is learned, where A is the probe and B is the correct response, then the AB unit in the conjunctive layer will become associated with the B unit on the response layer. On transitive test trials, the bidirectional connections allow activity to spread such that features from related episodes become activated. For example, consider a situation where the pairs A-B, B-C, X-Y, and Y-Z have been learned. At this point, each of the items in the feature layer is associated to units in the conjunctive layer corresponding to the pairs they have been presented in (for example, B is connected to both AB and BC units). Then, if A is presented, and either C or Z must be chosen as a response, then A, C, and Z are activated on the feature layer. These active units in the feature layer then activate the AB, BC, and YZ units in

the conjunctive layer (see Fig. 6). At this point, the conjunctive layer influences the activity on the feature layer; this recurrent activation in the network allows it to generalize and perform inference. Since both AB and BC are connected to the B unit in the feature layer, B becomes more active than X or Y. The active A, B, and C units then ensure that the AB and BC conjunctive units are most active; those units then activate B and C response units, and C is chosen since B is not an available option on this trial (Kumaran & McClelland, 2012). Through the interaction of features and pattern-separated conjunctions of features, the model is able to perform inference during retrieval, even though each event is stored in a distinct code (Kumaran, 2012).

3.3 Temporal context model

The temporal context model (TCM) has also been applied to explain performance in the PAI task, using an encoding-based mechanism, in contrast to the retrieval-based mechanism used in REMERGE (Kumaran, 2012). TCM was originally developed to explain findings in the free-recall paradigm (Howard & Kahana, 2002a). By assuming that items are associated with a slowly changing representation of temporal context, and that retrieved context can be used as a cue, TCM provides a successful account of recency and contiguity effects in free recall (see *Free recall* in Section 2; Howard & Kahana 2002a; Howard 2004; Howard et al. 2005; Sederberg et al. 2008; Polyn et al. 2009). Importantly, TCM assumes that context evolution is driven by the presentation of items; this assumption contrasts with other memory models which assume that context changes randomly (Estes, 1955; Mensink & Raaijmakers, 1988, 1989; Sirotin et al., 2005). TCM assumes that, even before an experiment, stimuli (which are typically common words) are associated with the states of context in which they were previously encountered (Howard & Kahana, 2002a). The pre-experimental contexts associated with different items are assumed to be orthogonal. When an item is presented, its pre-experimental context is retrieved. Over time, the current state of context integrates activity related to studied items. Items become associated with the current state of context

through Hebbian learning. Therefore, each item is associated with a state of context, which includes a recency-weighted combination of the pre-experimental contexts associated with recently studied items. Critically, when an item is recalled, it again triggers recall of its pre-experimental context (as well as the experimental contextual state that was active when the item was studied). This pre-experimental context then provides a purely forward-going cue, since that context was only active during items that were presented later in the list. This mechanism allows TCM to account for the tendency of participants to make forward transitions in a list (Howard & Kahana, 2002a).

Retrieved context also allows TCM to account for behavior in PAI tasks (Howard et al., 2005). When the pair A-B is learned, TCM assumes that B becomes associated with the pre-experimental context elicited by A. When the pair B-C is later studied, presentation of B causes retrieval of the context that was previously associated with it, which includes the pre-experimental context associated with A. C then becomes associated with this retrieved context. As a result, presentation of A will provide a good cue for B but also will be a good cue for C, which is associated with A's pre-experimental context (Howard et al. 2005; see also Howard et al. 2009). In this way, TCM can generalize from learning A-B and B-C, to also develop an association between A and C.

Thus, TCM successfully accounts for data in episodic memory tasks (Sederberg et al., 2008; Polyn et al., 2009) and fast inference tasks (Howard et al., 2005). Recent work has extended TCM to also account for the development of semantic knowledge (Rao & Howard, 2008). The application of TCM to semantic learning relies on information about the tendency for pairs of words to co-occur in similar contexts. This use of co-occurrence information is used in other models of semantic knowledge, such as LSA and BEAGLE (Landauer & Dumais, 1997; Jones & Mewhort, 2007). When trained on pairs of successive words, TCM's retrieved context mechanism allows words to become associated even when they were never presented together. If two items were studied in similar contexts (e.g. next to the same word), they will become associated even if they are never presented together (Rao & Howard, 2008).

However, the co-occurrence information extracted by TCM fails to capture other information given by the order of words. For example, in the sentence "The baker reached into the oven and pulled out the FLOOB", information about the meaning of the unknown word FLOOB can be extracted from its position in the sentence, to determine that FLOOB probably refers to some baked good such as bread (Howard et al., 2011). However, a fairly simple modification to TCM allows it to also use order information to determine the similarity structure of a set of words. In the pTCM model, context-to-item associations are used to predict expected items, given the current state of context. This predicted activity determines the semantic representation of the current word; on subsequent presentations of the word, this semantic representation is used to update context. Over time, each item's representation comes to reflect the contexts that it fits into (Howard et al., 2011). The pTCM model bears some resemblance to the order information encoded in each item's representation in the BEAGLE model (Jones & Mewhort, 2007).

3.4 Conclusions

The work of Howard et al. (2011) in developing pTCM illustrates that a form of semantic learning (development of similarity structure) can be driven by context maintenance, retrieval, and prediction operations. This suggests that episodic and semantic learning might rely on similar mechanisms (Howard et al., 2011). However, there is also evidence that slower, semantic learning occurs on a larger timescale than episodic learning, which requires learning after a single presentation (McClelland et al., 1995). The finding that hippocampal damage is associated with retrograde amnesia, sometimes dating back years, suggests that a slow consolidation process is necessary for learning in the cortex (McClelland et al., 1995).

Evidence for fast, hippocampally dependent generalization (Bunsey & Eichenbaum, 1996; Preston et al., 2004) complicates the picture further. It is possible that generalization in the hippocampus may facilitate cortical learning of semantic information through the consolidation process (Kumaran & McClelland, 2012). However, there is evidence that semantic

learning does not depend critically on the hippocampus: Patients whose hippocampi were damaged at an early age have relatively unimpaired performance on tasks designed to assess semantic knowledge (Vargha-Khadem et al., 1997). It will be important for future work with pTCM, which attempts to bridge between episodic and semantic learning by utilizing similar mechanisms for both, to account for neuropsychological findings suggesting that the two forms of memory rely on distinct brain networks.

An important challenge for future work will be to develop a model that can account for performance on list-learning tasks like free recall, fast inference in paradigms such as paired-associate inference, and gradual learning of semantic relations; in addition, theories of declarative memory should account for the effects of focal brain damage on performance in each of these tasks. Although there is much work to be done to account for results from each of these domains, both the complementary learnings systems framework and the temporal context model framework have taken significant steps toward a comprehensive theory of declarative memory.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, p. 89-105). New York: Academic Press.

Bahrick, H. P. (1971). Accessibility and availability of retrieval cues in the retention of a categorized list. *Journal of Experimental Psychology*, *89*(1), 117–125.

Batchelder, W., & Riefer, D. (1980). Separation of storage and retrieval factors in free recall of clusterable pairs. *Psychological Review*, *87*(4), 375–397.

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories:

A replication and extension of the connecticut category norms. *Journal of Experimental Psychology Monograph*, *80*(3 Pt. 2), 1–46.

Becker, S., & Lim, J. (2003). A computational model of prefrontal control in free recall: Strategic memory use in the california verbal learning task. *Journal of Cognitive Neuroscience*, *15*, 821–832.

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, *6*, 173–189.

Borges, M. A., & Mandler, G. (1972). Effect of within-category spacing on free recall. *Journal of Experimental Psychology*, *92*, 207–214.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology*, *49*, 229–240.

Bousfield, W. A., & Cohen, B. H. (1953). The effects of reinforcement on the occurrence of clustering in the recall of randomly arranged associates. *The Journal of Psychology*, *36*, 67–81.

Bousfield, W. A., & Cohen, B. H. (1955). The occurrence of clustering in the recall of randomly arranged words of different frequencies-of-usage. *Journal of General Psychology*, *52*, 83–95.

Bousfield, W. A., Cohen, B. H., & Whitmarsh, G. A. (1958). Associative clustering in the recall of words of different taxonomic frequencies of occurrence. *Psychological Reports*, *4*, 39–44.

Bunsey, M., & Eichenbaum, H. B. (1996). Conservation of hippocampal memory function in rats and humans. *Nature*, *379*(6562), 255–257.

Cofer, C. N., Bruce, D. R., & Reicher, G. M. (1966). Clustering in free recall as a function of certain methodological variations. *Journal of Experimental Psychology*, *71*, 858–866.

Cohen, B. H. (1963a). An investigation of recoding in free recall. *Journal of Experimental Psychology*, *65*(4), 368-376.

Cohen, B. H. (1963b). Recall of categorized word lists. *Journal of Experimental Psychology*, *66*(3), 227-234.

Cohen, B. H. (1966). Some-or-none characteristics of coding behavior. *Journal Of Verbal Learning And Verbal Behavior*, *5*(2), 182–187.

Cohen, B. H., Bousfield, W. A., & Whitmarsh, G. A. (1957). Cultural norms for verbal items in 43 categories. (Technical Report No. 22).

Collins, A. M., & Loftus, E. F. (1975). Spreading activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428.

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

D'Agostino, P. R. (1969). The blocked-random effect in recall and recognition. *Journal of Verbal Learning and Verbal Behavior*, *8*, 815–820.

Dallet, K. M. (1964). Number of categories and category information in free recall. *Journal of Experimental Psychology*, *68*(1), 1–12.

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.

Deese, J. (1959a). Influence of inter-item associative strength upon immediate free recall. *Psychological Reports*, *5*, 305–312.

Deese, J. (1959b). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*, 17–22.

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review*, *62*, 145–154.

Feldbaum, C. (1998). A semantic network of English: The mother of all WordNets. *Computers and the Humanities*, *32*(2-3), 209-220.

Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto Word Pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods and Instrumentation*, *14*, 375–399.

Glanzer, M. (1969). Distance between related words in free recall: Trace of the STS. *Journal of Verbal Learning and Verbal Behavior*, *8*, 105–111.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, *5*, 351–360.

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psycholgical Review*, *114*(2), 211–44.

Howard, M. W. (2004). Scaling behavior in the temporal context model. *Journal of Mathematical Psychology*, *48*, 230–238.

Howard, M. W., Fotedar, M. S., Datey, A. V., & Hasselmo, M. E. (2005). The temporal context model in spatial navigation and relational learning: Toward a common explanation of medial temporal lobe function across domains. *Psychological Review*, *112*(1), 75–116.

Howard, M. W., Jing, B., Rao, V. A., Provyn, J. P., & Datey, A. V. (2009). Bridging the gap: Transitive associations between items presented in similar temporal contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(2), 391-407.

Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 923–941.

Howard, M. W., & Kahana, M. J. (2002a). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*, 269–299.

Howard, M. W., & Kahana, M. J. (2002b). When does semantic similarity help episodic retrieval? *Journal of Memory and Language*, *46*, 85–98.

Howard, M. W., Shankar, K. H., & Jagadisan, U. K. K. (2011). Constructing semantic representations from a gradually changing representation of temporal context. *Topics in Cognitive Science*, *3*(1), 48–73.

Howard, M. W., Venkatadass, V., Norman, K. A., & Kahana, M. J. (2007). Associative processes in immediate recency. *Memory & Cognition*, *35*(7), 1700–1711.

Howard, M. W., Viskontas, I. V., Shankar, K. H., & Fried, I. (2012). Ensembles of human mtl neurons "jump back in time" in response to a repeated stimulus. *Hippocampus*, *22*(9), 1833–1847.

Johns, B. T., & Jones, M. N. (2010). Evaluating the random representation assumption of lexical semantics in cognitive models. *Psychonomic Bulletin & Review*, *17*(5), 662-672.

Jones, M. N., & Mewhort, D. J. K. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review*, *114*(1), 1–37.

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, *24*, 103–109.

Kahana, M. J. (2012). *Foundations of human memory* (1 ed.). New York, NY: Oxford University Press.

Kahana, M. J., & Wingfield, A. (2000). A functional relation between learning and organization in free recall. *Psychonomic Bulletin & Review*, *7*, 516–521.

Kimball, D. R., Bjork, E. L., Bjork, R. A., & Smith, T. A. (2008). Part-list cuing and the dynamics of false recall. *Psychonomic Bulletin & Review*(15), 296–301.

Kimball, D. R., Smith, T. A., & Kahana, M. J. (2007). The fSAM model of false recall. *Psychological Review*, *114*(4), 954–93.

Kumaran, D. (2012). What representations and computations underpin the contribution of the hippocampus to generalization and inference? *Frontiers in Human Neuroscience*, *6*(157), 1-11.

Kumaran, D., & McClelland, J. L. (2012). Generalization through the recurrent interaction of episodic memories: A model of the hippocampal system. *Psychological Review*, *119*(3), 573-616.

Landauer, T. K., & Dumais, S. T. (1997). Solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*, 476–490.

Lohnas, L. J., Polyn, S. M., & Kahana, M. J. (2011). Contextual variability in free recall. *Journal of Memory and Language*, *64*(3), 249-255.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments & Computers*, *28*(2), 203–208.

Madigan, S. A. (1969). Intraserial repetition and coding processes in free recall. *Journal of Verbal Learning and Verbal Behavior*, *8*, 828–835.

Manning, J. R., & Kahana, M. J. (2012). Interpreting semantic clustering effects in free recall. *Memory*, *20*(5), 511-517.

Manning, J. R., Polyn, S. M., Baltuch, G., Litt, B., & Kahana, M. J. (2011). Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. *Proceedings*

of the National Academy of Sciences of the United States of America, *108*(31), 12893–12897.

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), 419–57.

McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, *4*(4), 310–322.

Mensink, G.-J. M., & Raaijmakers, J. G. W. (1988). A model for interference and forgetting. *Psychological Review*, *95*, 434–455.

Mensink, G.-J. M., & Raaijmakers, J. G. W. (1989). A model for contextual fluctuation. *Journal of Mathematical Psychology*, *33*, 172–186.

Miller, G. A. (1956, Mar). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2).

Miller, J. F., Weidemann, C. T., & Kahana, M. J. (2012). Recall termination in free recall. *Memory & Cognition*, *40*(4), 540–550.

Morton, N. W., Kahana, M. J., Rosenberg, E. A., Baltuch, G. H., Litt, B., Sharan, A. D., et al. (in press). Category-specific neural oscillations predict recall organization during memory search. *Cerebral Cortex*.

Morton, N. W., & Polyn, S. M. (2012). *A neurally constrained model of category clustering in free recall.* Psychonomic Society Annual Meeting. Minneapolis, MN.

Murdock, B. B. (1960). The immediate retention of unrelated words. *Journal of Experimental Psychology*, *60*, 222–234.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*, 482–488.

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

Murdock, B. B., & Okada, R. (1970). Interresponse times in single-trial free recall. *Journal of Verbal Learning and Verbal Behavior*, *86*, 263–267.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments and Computers*, *36(3)*, 402–407.

Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary learning systems approach. *Psychological Review*, *110*, 611–646.

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*(9), 424–430.

O'Reilly, R. C., & McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. *Hippocampus*, *4*(6), 661–682.

O'Reilly, R. C., & Norman, K. A. (2002). Hippocampal and neocortical contributions to memory: Advances in the complementary learning systems framework. *Trends in Cognitive Sciences*, *6*(12), 505–510.

Patterson, K. E., Meltzer, R. H., & Mandler, G. (1971). Inter-response times in categorized free recall. *Journal of Verbal Learning and Verbal Behavior*, *10*, 417-426.

Pollio, H. R., Kasschau, R. A., & DeNise, H. E. (1968). Associative structure and the temporal characteristics of free recall. *Journal of Verbal Learning and Verbal Behavior*, *10*, 190-197.

Pollio, H. R., Richards, S., & Lucas, R. (1969). Temporal properties of category recall. *Journal of Verbal Learning and Verbal Behavior*, *8*, 529-536.

Polyn, S. M., & Kahana, M. J. (2008). Memory search and the neural representation of context. *Trends in Cognitive Sciences*, *12*, 24–30.

Polyn, S. M., Kragel, J. E., Morton, N. W., McCluey, J. D., & Cohen, Z. D. (2012). The neural dynamics of task context in free recall. *Neuropsychologia*, *50*(4), 447-457.

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

Postman, L., & Phillips, L. W. (1965). Short-term temporal changes in free recall. *Quarterly Journal of Experimental Psychology*, *17*, 132-138.

Preston, A. R., Shrager, Y., Dudukovic, N. M., & Gabrieli, J. D. E. (2004). Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus*, *14*(2), 148–152.

Puff, C. R. (1966). Clustering as a function of the sequential organization of stimulus word lists. *Journal of Verbal Learning and Verbal Behavior*, *5*, 503-506.

Puff, C. R. (1974). A consolidated theoretical view of stimulus-list organization effects in free recall. *Psychological Reports*, *34*(1), 275-288.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 14, p. 207-262). New York: Academic Press.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, *88*, 93–134.

Rao, V. A., & Howard, M. W. (2008). Retrieved context and the discovery of semantic structure. In J. C. Platt, D. Koller, Y. Singer, & S. Roweis (Eds.), *Advances in neural information processing systems* (p. 1193-1200). Cambridge, MA: MIT Press.

Recchia, G., & Jones, M. N. (2009). More data trumps smarter algorithms: Comparing pointwise mutual information with latent semantic analysis. *Behavior Research Methods*, *41*(3), 647-656.

Roediger, H. L.(1973). Inhibition in recall from cueing with recall targets. *Journal Of Verbal Learning And Verbal Behavior*, *12*(6), 644–657.

Roediger, H. L., & McDermott, K. B.(1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*, 803-814.

Roenker, D. L., Thompson, C. P., & Brown, S. C.(1971). Comparison of measures for the estimation of clustering in free recall. *Psychological Bulletin*, *76*(1), 45–48.

Romney, A. K., Brewer, D. D., & Batchelder, W. H. (1993). Predicting clustering from semantic structure. *Psychological Science*, *4*, 28-34.

Rotondo, J. A.(1977). Discrete structural models of organization in free recall. *Journal of Mathematical Psychology*, *16*(2), 95-120.

Rumelhart, D. E. (1990). An introduction to electronic and neural networks. In S. F. Zornetzer, J. L. Davis, & C. Lau (Eds.), (p. 405-420). Academic Press.

Sederberg, P. B., Howard, M. W., & Kahana, M. J.(2008). A context-based theory of recency and contiguity in free recall. *Psychological Review*, *115*(4), 893–912.

Sederberg, P. B., Miller, J. F., Howard, W. H., & Kahana, M. J. (2010). The temporal contiguity effect predicts episodic memory performance. *Memory & Cognition*, *38*(6), 689–699.

Shepard, R. N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science*, *210*(4468), 390-398.

Shuell, T. J. (1969). Clustering and organization in free recall. *Psychological Bulletin*, *72*, 353-374.

Sirotin, Y. B., Kimball, D. R., & Kahana, M. J. (2005). Going beyond a single list: Modeling the effects of prior experience on episodic free recall. *Psychonomic Bulletin & Review*, *12*(5), 787-805.

Smith, A. D. (1971). Output interference and organized recall from long-term memory. *Journal of Verbal Learning and Verbal Behavior*, *10*(4), 400-408.

Socher, R., Gershman, S. J., Perotte, A. J., Sederberg, P. B., Blei, D. M., & Norman, K. A. (2009). A bayesian analysis of dynamics in free recall. In Y. Bengio, D. Schuurmans, J. Lafferty, C. K. I. Williams, & A. Culotta (Eds.), *Advances in neural information processing systems.* MIT Press.

Solway, A., Geller, A. S., Sederberg, P. B., & Kahana, M. J. (2010). Pyparse: A semiautomated system for scoring spoken recall data. *Behavior Research Methods*, *42*(1), 141–147.

Steyvers, M., Shiffrin, R. M., & Nelson, D. L. (2004). Word association spaces for predicting semantic similarity effects in episodic memory. In A. F. Healy (Ed.), *Cognitive psychology and its applications: Festschrift in honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer.* (pp. 237–249). Washington, DC: American Psychological Association.

Stricker, J. L., Brown, G. G., Wixted, J. T., Baldo, J. V., & Delis, D. C. (2002). New semantic and serial clustering indices for the California Verbal Learning Test–Second Edition: Background, rationale, and formulae. *Journal of the International Neuropsychological Society*, *8*, 425–435.

Stuss, D. T., Alexander, M. P., Palumbo, C. L., Buckle, L., Sayer, L., & Pogue, J. (1994). Organizational Strategies of Patients With Unilateral or Bilateral Frontal Lobe Injury in Word List Learning Tasks. *Neuropsychology*, *8*, 355–355.

Thompson, C. P. (1978). Evidence for learning-to-cluster as a retrieval strategy. *American Journal of Psychology*, *91*(1), 115–124.

Tulving, E. (1962). Subjective organization in free recall of "unrelated" words. *Psychological Review*, *69*(4), 344-354.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory.* (p. 381-403). New York: Adademic Press.

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*, 381-391.

Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*(3), 550-592.

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language*, *50*(3), 289–335.

Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connely, A., Van Paesschen, W., & Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science*, *277*, 376-380.

Zaromb, F. M., Howard, M. W., Dolan, E. D., Sirotin, Y. B., Tully, M., Wingfield, A., et al. (2006). Temporal associations and prior-list intrusions in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(4), 792–804.