Semantic similarity dissociates short-from long-term recency effects: Testing a neurocomputational model of list memory

EDDY J. DAVELAAR and HENK J. HAARMANN
University of Maryland, College Park, Maryland

YONATAN GOSHEN-GOTTSTEIN
Tel Aviv University, Tel Aviv, Israel

and

MARIUS USHER
Birkbeck College, University of London, London, England

The finding that recency effects can occur not only in immediate free recall (i.e., short-term recency) but also in the continuous-distractor task (i.e., long-term recency) has led many theorists to reject the distinction between short- and long-term memory stores. Recently, we have argued that long-term recency effects do not undermine the concept of a short-term store, and we have presented a neurocomputational model that accounts for both short- and long-term recency and for a series of dissociations between these two effects. Here, we present a new dissociation between short- and long-term recency based on semantic similarity, which is predicted by our model. This dissociation is due to the mutual support between associated items in the short-term store, which takes place in immediate free recall and delayed free recall but not in continuous-distractor free recall.

The concept of a limited-capacity short-term memory (STM) buffer has played a central role in theories of memory. Indeed, such a buffer is an essential component of computational models that account for quantitative data in tasks of immediate and delayed free recall, including the search of associative memory model (SAM; Raaijmakers & Shiffrin, 1981; see also Atkinson & Shiffrin, 1968; Kahana, 1996; Waugh & Norman, 1965) and the adaptive control of thought–rational model (ACT–R; Anderson, Bothell, Lebiere, & Matessa, 1998). However, the need to postulate the existence of such a buffer to account for free recall data seems to have fallen into disrepute (Bjork, 2001; Bjork & Whitten, 1974; Crowder, 1982; Glenberg & Swanson, 1986; Greene, 1986, 1992; Howard & Kahana, 1999; Nairne, Neath, Serra, & Byun, 1997; Neath, 1993; Tan & Ward, 2000; but see Healy & McNamara, 1996, and Raaijmakers, 1993, for a defense). Here, we present new data supporting the original notion that a short-term buffer is involved in producing the recency effect in im-

mediate free recall. We will not address serial order recall, for which the evidence for a short-term buffer is thought to be weaker (Nairne, 2002).

One of the findings that initially suggested a role for a short-term buffer was the existence of a *recency effect* (the enhanced recall of the last items of a list) in immediate free recall (Atkinson & Shiffrin, 1968; Glanzer, 1972; Murdock, 1967). This effect was interpreted as reflecting the unloading of the most recent items from a limited-capacity store, with earlier items having been displaced from the buffer by the more recent items. Because distractor tasks are thought to fill the capacity of the buffer, this helped explain why, when a distractor interval is inserted between list presentation and recall—a task known as *delayed free recall*—the recency effect is eliminated (Glanzer & Cunitz, 1966; Postman & Phillips, 1965).

A central reason for questioning the need for such a buffer was the discovery of recency effects in a paradigm that could not be interpreted as being mediated by the short-term buffer. The long-term memory (LTM) paradigm involved free recall of a list of words in which a distractor activity was inserted before and after every word in the list, including the last word (i.e., *continuous-distractor free recall*; see, e.g., Bjork & Whitten, 1974; Glenberg, Bradley, Kraus, & Renzaglia, 1983; Glenberg & Swanson, 1986; Tzeng, 1973). Because the absence of recency effects in delayed free recall had been perceived as one signature for the existence of a short-term buffer, the appearance of recency effects in continuous-distractor free

recall, despite the distractor activity after the last item, led many investigators to question the existence of a buffer. Instead of postulating a short-term buffer, they argued that the recency effect reflects a more general property of memory and that a single memory store provides a more parsimonious account of recency effects (Crowder, 1982; Greene, 1986).

To bolster this interpretation, Greene (1986) argued that many experimental manipulations have similar effects in immediate and continuous-distractor free recall (i.e., associations, rather than dissociations), indicating that short- and long-term recency effects have a common basis. More recently, Howard and Kahana (1999) reported an additional association between the two tasks in the conditional probability of reporting successive items as a function of their lag in the list.

In a recent article, we have argued that a careful consideration of the data and of the competing theories demonstrates that the existence of long-term recency effects does not undermine the notion of a short-term buffer that contributes to short-term recency effects (Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, & Usher, 2005; see also Talmi, Grady, Goshen-Gottstein, & Moscovitch, 2005). First, the previous associations found between short- and long-term recency (i.e., those from Greene, 1986) were obtained across studies that employed different methodologies and different materials. Second, we described a dual-store neurocomputational model that was successful in accounting for some associations between short- and long-term recency. This established that an association between short- and long-term recency cannot, in and of itself, contradict the dual-store interpretation. Third, we highlighted earlier studies that revealed dissociations be-

tween short- and long-term recency effects based on amnesia (Carlesimo, Marfia, Loasses, & Caltagirone, 1996) and order of output (Dalezman, 1976; Whitten, 1978).[1] Finally, based on the predictions of our neurocomputational model, we reported new data showing a dissociation between short- and long-term recency. This dissociation was revealed using identical methodology and materials for the immediate and the continuous-distractor free recall tasks. We found that proactive interference hurt recall of recency items in the continuous-distractor task but not in immediate free recall (Davelaar et al., 2005).

The demonstration of a dissociation between short- and long-term recency is a key finding for establishing the psychological reality of a short-term buffer. Indeed, in his original article, Crowder (1993) raised the challenge that ". . . the burden of evidence [for the existence of a buffer] should be with those who say these two, similar recency effects are caused by different mechanisms" (p. 143). Because only a single dissociation between short- and long-term recency (with identical materials and procedures used across tasks) has been reported to date (Davelaar et al., 2005), we investigated a novel dissociation due to a semantic manipulation, which was predicted by our neurocomputational model (Davelaar et al., 2005).

Our investigation was motivated by a careful examination of a semantic effect that was reported by Glanzer and Schwartz (1971; see Figure 1). In their experiment, participants were shown lists containing four pairs of weak associates[2] and four pairs of unrelated words for an immediate test of free recall. Their results showed better recall for the associated pairs than for the unrelated words. This effect could be mediated by semantic contributions at retrieval (Baddeley, 1972; but see Haarmann & Usher,
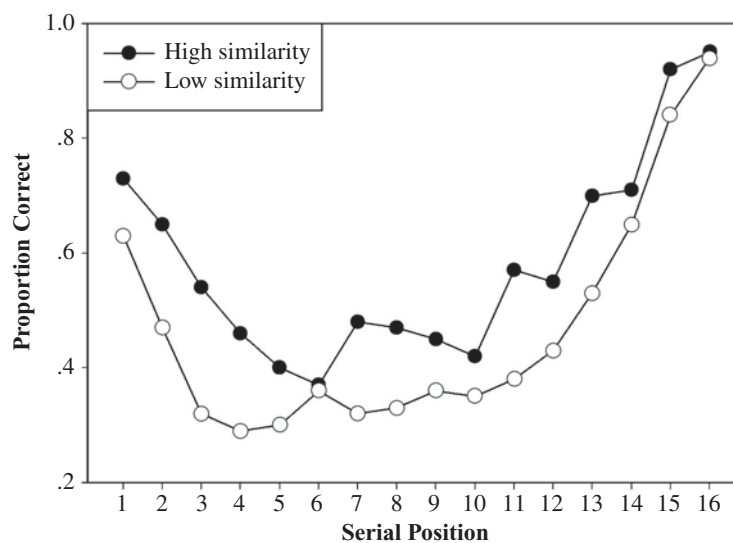

Figure 1. Serial position curves showing effects of semantic similarity on recall probability in immediate free recall for semantically related and unrelated word pairs. From "Mnemonic Structure in Free Recall: Differential Effects on STS and LTS," by M. Glanzer and A. Schwartz, 1971, *Journal of Verbal Learning & Verbal Behavior*, *10*, pp. 194-198. Copyright 1971 by Elsevier. Reprinted with permission.

2001) and may therefore not be relevant to the debate regarding the existence of a short-term store.

However, upon closer inspection of Glanzer and Schwartz's (1971) data, we noticed that the serial position function for related pairs shows an up-and-down pattern, which—because it resembles the zigzag pattern of the teeth of a saw—we label the *zigzag effect*. This pattern is difficult to account for by a semantic contribution at retrieval and, as we shall argue in this article, likely reflects the operation of a short-term store.

Even though our model had initially been developed to account for a particular set of data in free and cued recall (Davelaar, 2003; Davelaar et al., 2005; Davelaar & Usher, 2002; Haarmann & Usher, 2001; Usher & Cohen, 1999), it turned out to also predict such a zigzag effect in immediate but not in continuous-distractor free recall. The zigzag effect was predicted to emerge only in immediate (and in delayed), but not in continuous-distractor, free recall, because in the former it is due to the temporal activation dynamics of coactivated items in the short-term buffer. To give an idea about why this dissociation is predicted, we begin with a rough sketch of the model. We then present simulations that predict a zigzag effect in immediate and in delayed but not in continuous-distractor free recall. Finally, we present experimental results that confirm this prediction.

Model

The model consists of two interconnected layers, corresponding to a lexical-semantic and an episodic-context representation (for computational details, see the Appendix; Davelaar et al., 2005). The lexical-semantic layer contains localistic units (representing words) with self-excitation, which enables the units to remain active after stimulus offset. It also includes global inhibition, which prevents an unbounded spread of activation, allowing only a few items to be active simultaneously (for analyses, see Davelaar, 2003; Usher & Cohen, 1999).

Both properties, self-excitation and inhibition, give rise to a storage system that behaves like an activation buffer. The model is similar to SAM (Raaijmakers & Shiffrin, 1981) and takes a Hebbian view of short-term maintenance, where short-term retention maps onto activation-based processes. By contrast, long-term retention maps onto weight-based processes (Hebb, 1949) between the lexical-semantic and a changing contextual representation (see also Dennis & Humphreys, 2001; Glenberg & Swanson, 1986; Howard & Kahana, 2002a; Mensink & Raaijmakers, 1988).

Critical to our concerns, the model represents semantic relations between items by assuming weak excitatory connections between the corresponding lexical-semantic units; when two semantically associated units are coactive, the excitatory connection between them makes them support each other, partially offsetting the global inhibition. According to our model, items enter a short-term buffer where an interplay between activation and inhibition determines their duration in the buffer before they decay (see the Appendix for details). An additional factor that can prolong this duration is the extent to which neighboring items can mutually support each other. Such support can be obtained when adjacent items are semantically related, even if only weakly so. Therefore, weakly related items should remain in the buffer for longer durations than do unrelated items.

Unlike other models of STM that are exclusively phonological, our buffer corresponds to the activated part of a lexical-semantic representation and is thus sensitive to semantic information. As in SAM, later items tend to replace earlier items in the buffer. Therefore, the support between associated items is more pronounced for the first item of each pair. Whereas for two successive unrelated words, the first item tends to be displaced before its successor, and it has a lower probability of remaining in the buffer than does its successor (see Figure A1, Appendix), for related word pairs, the two items support each other and tend to be displaced together. As a result, their probability of remaining in the buffer tends to be the same (see the Appendix). This gives an advantage to the first member of each semantic pair relative to the second member, leading to the zigzag pattern for the items that are reported from the buffer (i.e., the recency items in immediate free recall). Thus, the first component that contributes to the zigzag effect is the unloading from the short-term buffer.

An additional component in our model that contributes to the zigzag effect is the long-term episodic-memory component. In our model, the weights between items and their context are determined as a function of (the integral across time of) the activation of items in the buffer. Therefore, items that are activated longer also have stronger episodic memory traces. Because the first item in each associated pair is active longer, an LTM-based zigzag effect should also be observed whenever items had resided together in the buffer. Therefore, in immediate free recall, in addition to an STM contribution to the zigzag effect at recency positions, a smaller, episodically based zigzag effect is also predicted to be found for prerecency positions.

A zigzag effect is also predicted for delayed free recall, where during encoding, items reside in the buffer together but are then displaced by the distractor activity and can therefore be retrieved only from LTM. Because one of the components that contributes to the zigzag effect is the LTM component, a zigzag effect is predicted for all items in delayed recall. However, because the STM component does not contribute to the zigzag effect in this task, the zigzag effect is predicted to be smaller in size than it is for immediate free recall and to show no interaction with position in the list.

Finally, in continuous-distractor free recall, items are separated during encoding by distractor activity, preventing mutual support between related items. A zigzag effect is not predicted to emerge at any serial position because both the advantage in the persistence of the first item (of a pair) in the buffer and the advantage in the episodic weights, both of which mediate the zigzag effect, depend on the mutual support between associates during encoding.

In addition to our specific prediction of different zigzag profiles in the three free recall tasks, we expect that recall

performance will be better in the related than in the unrelated condition, consistent with previous literature (Glanzer & Schwartz, 1971; Greene & Crowder, 1984). To account for this general relatedness effect, the model assumes that, during retrieval, associated items tend to "prime" each other and be reported in clusters (see, e.g., Howard & Kahana, 2002b), thereby increasing the total recall in the related condition compared with the unrelated condition.

To illustrate the model's predictions, we ran a simulation for lists consisting of six related or unrelated pairs in immediate, delayed, and continuous-distractor free recall (see the Appendix for details), mirroring the experimental procedure described in the Method section. Figure 2 presents the simulated serial position functions.

As can be seen in Figure 2, the model predicts a zigzag effect similar to the pattern observed in the immediate free recall data of Glanzer and Schwartz (1971). More importantly, the model predicts a dissociation between the profiles of the zigzag effect across the three free recall tasks. To obtain a more precise measure of the zigzag effect, we computed a score of the presence and magnitude of the zigzag effect (see Figure 3). To obtain this zigzag score, a difference score was calculated for each item by subtracting the item's expected recall from the item's actual recall. The expected recall for an item was determined by a linear interpolation of the recall of the two neighboring items (i.e., their average).[3] Next, we calculated a *grouped zigzag index* by summing the difference scores at odd serial positions (corresponding to the first item in each pair) and at the even serial positions (corresponding to the second item in each pair) and subtracting the even difference scores from the odd difference scores. This grouped zigzag index is positive when a zigzag pattern exists.[4] We did this separately for the first half (corresponding to prerecency items) and the second half (corresponding to recency items) of the list, because items from the second list-half have a higher probability of being reported directly from the buffer.[5]

Figure 3 presents the grouped difference scores for the three tasks (left panel) and the grouped scores for the related condition, broken down by list-half (right panel). Examination of Figure 3 reveals that, for immediate free recall, the model predicts a zigzag effect in the related, but

not in the unrelated, condition, with the zigzag effect in the related condition larger for the second half of the list than for the first half. For delayed free recall, the model predicts a zigzag effect in the related (but not in the unrelated) condition, which is of similar magnitude for both list-halves (and halfway between the magnitudes in the first and second halves of the immediate free recall list). Finally, for continuous-distractor free recall, the model predicts no zigzag effects at any serial position.

The Present Experiment

A critical aspect of the model's ability to predict the zigzag effect is that it relies on the mutual support of co-active items in an activation-based buffer, which has its effect both on the maintenance in the buffer and on the transfer to episodic memory. Therefore, the zigzag pattern forms an ideal focus point for testing the assertion of the need for a short-term buffer (in addition to a contextual retrieval mechanism) in accounting for data in free recall paradigms. In order to test the model predictions, we ran an experiment using lists consisting of either six unrelated or six weakly associated word pairs. These lists were used in all three free recall tasks (immediate, delayed, and continuous distractor), with each of three groups of participants doing only one of the three tasks.

METHOD

Participants

A total of 119 undergraduate students from the University of Maryland and the University of London, all native speakers of English, participated in the experiment for course credit or a monetary reward (of about $7).

Design

The experiment conformed to a 2 × 3 mixed-factor design, crossing the within-subjects factor relatedness (related or unrelated) with the between-subjects factor task (immediate, $n = 28$; delayed, $n = 40$; continuous distractor, $n = 51$). We added the within-subjects factor list-half in the analysis of the zigzag scores.

Materials

The word pairs for the memory lists were adopted from Experiment 1 in Haarmann and Usher (2001) and included a random subselection of 72 of their 144 word pairs of weak semantic associates.
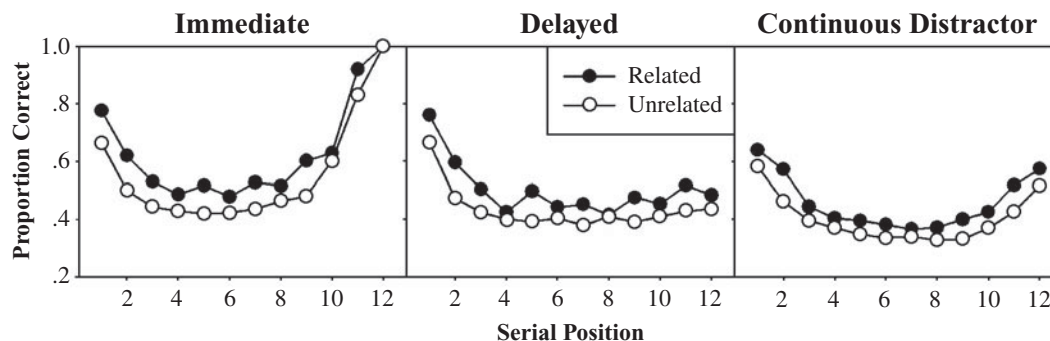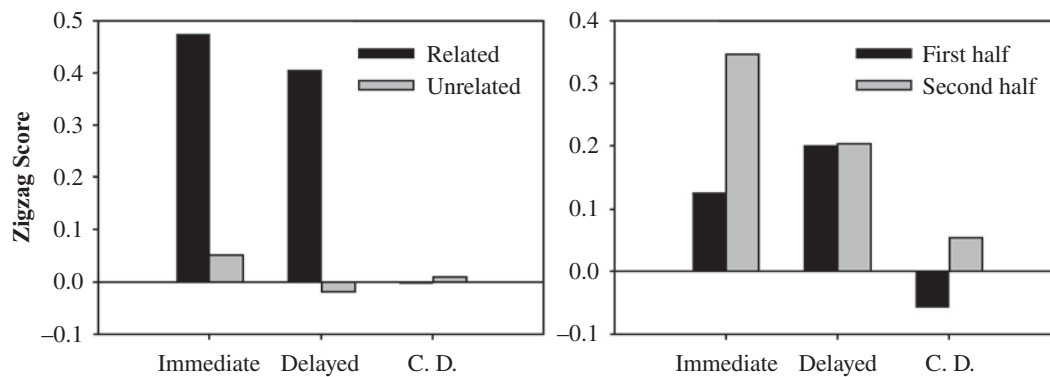

Figure 2. Model predictions for the mean recall probability in immediate (left), delayed (middle), and continuous-distractor (right) free recall as a function of serial position.

Figure 3. (Left) Model predictions for the total grouped zigzag scores for the three free recall tasks. The larger the value, the larger the overall zigzag effect. (Right) Grouped zigzag scores for the related condition only, broken down by list-half. C. D., continuous distractor.

From these, we created four lists of 12-trial sequences, with each list containing 6 trials in each relatedness condition. Trials in the related condition consisted of a sequence of 6 pairs of semantic associates (e.g., *light*, *candle*; *sleep*, *blanket*; *spider*, *snake*; *heavy*, *stone*; *street*, *curb*; *hotter*, *boil*). Trials in the unrelated condition consisted of a sequence of 6 different pairs of unrelated words (e.g., *bald*, *wonder*; *liver*, *prayer*; *distance*, *bench*; *cry*, *eagle*; *faster*, *bitter*; *table*, *working*), which were created by regrouping pairs of semantic associates with the constraint that words from the same pair did not occur in the same trial and were separated from each other by at least 1 other trial. Assignment of word pairs to lists and list-halves and of lists to participants was counterbalanced in such a way that, across participants, all pairs were presented in all unique conditions. The same lists were used in all three tasks.

Procedure

Presentation of all stimuli was visual and computer controlled. In immediate free recall, the participants silently read a sequence of 12 words, presented at a rate of 1 word per sec, with 75 msec between words. Immediately after the offset of the last word, a recall cue ("?") prompted the participants to report as many words as they could, out loud, in any order, within 45 sec.

In delayed free recall, the participants had to perform a distractor task, reading and solving a sequence of arithmetic problems out loud, before the recall cue was given. These problems involved addition or subtraction of two randomly generated integers, ranging from 1 to 9, and always resulted in a positive sum. If a buttonpress did not follow after 2 sec, the next problem was automatically presented, but the participants were typically able to respond within the allotted time. The experimental program stopped the presentation of further problems when 12 sec of total problem-solving time had passed. To encourage the participants to pay attention to the arithmetic, error feedback was given, and the participants were instructed to have an error rate smaller than 20%. In continuous-distractor free recall, the events in a trial were the same as in immediate free recall, with the exception that before each word on a list and after the last word on a list, the distractor task was given.

RESULTS

The data of 13 participants in the delayed free recall task and of 10 participants in the continuous-distractor free recall task were excluded from the analysis because of an error rate of more than 20% on the distractor task. However, the results, including the critical triple interaction and subsequent simple effects, were not compromised when all of the data were included.

The serial position functions for all three groups are in agreement with the general literature, showing primacy effects for all three tasks and recency effects for only immediate and continuous-distractor free recall (see Figure 4, for related pairs). In addition, in all three tasks, performance was better in the related than in the unrelated condition, replicating previous findings (Glanzer & Schwartz, 1971; Greene & Crowder, 1984; Haarmann & Usher, 2001). As predicted by the model (see Figure 2), a zigzag pattern can be observed in the related condition for immediate and delayed free recall (see Figure 5, left). More importantly, as predicted by our model, in immediate free recall, the zigzag effect is stronger in the second than in the first half of the list, whereas in delayed free recall, no clear difference between list-halves is evident (see Figure 5, right). No zigzag effect can be discerned in continuous-distractor free recall.

Statistical analyses verified the above description of the experimental results. An ANOVA on the raw recall data revealed the predicted main effects of relatedness [$F(1,93) = 126.038$, $MS_e = 0.054$, $p < .001$], serial position [$F(5.396,501.851) = 24.016$, $MS_e = 0.104$, $p < .001$], and task [$F(2,93) = 18.936$, $MS_e = 0.333$, $p < .001$]. These reflect better recall in the related than in the unrelated condition, the curvilinear serial position effect, and better recall in immediate than in delayed or continuous-distractor free recall, respectively. Of the interactions, only serial position × task [$F(10.792,501.851) = 7.229$, $MS_e = 0.104$, $p < .001$] and relatedness × serial position [$F(7.2,669.624) = 2.971$, $MS_e = 0.058$, $p < .005$] reached significance. The former reflects the expected differences in the serial position profile, with stronger recency in the immediate than in the continuous-distractor free recall, and no recency in delayed free recall. The latter interaction indicates that the relatedness effect changes with serial position, which is further analyzed using the zigzag scores.

An overall ANOVA on the grouped zigzag scores for the first list-half (Serial Positions 2–6) and second list-
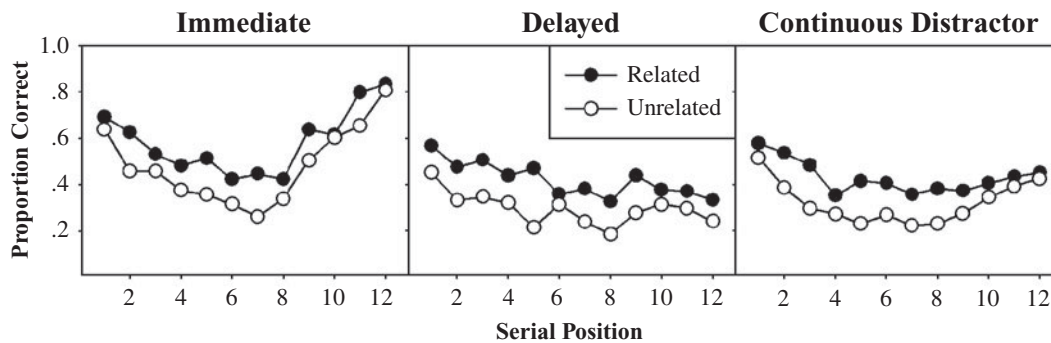
Figure 4. Experimental results for the average proportion correct in immediate (left), delayed (middle), and continuous-distractor (right) free recall as a function of serial position. Compare with Figure 2.

half (Serial Positions 7–11) revealed main effects of re-latedness [$F(1,93) = 10.301$, $MS_e = 0.390$, $p < .005$] and task [$F(2,93) = 3.109$, $MS_e = 0.396$, $p < .05$], cor-responding to a zigzag effect in the related, but not in the unrelated, condition and in immediate and delayed, but not in continuous-distractor, free recall. Most importantly, a significant triple interaction between relatedness, list-half, and task [$F(2,93) = 5.234$, $MS_e = 0.326$, $p < .01$] was obtained, which was analyzed further. As predicted by the model, only in immediate free recall was there an interaction between relatedness and list-half [$F(1,27) = 6.01$, $MS_e = 0.349$, $p < .05$], which was due to a stronger zigzag effect in the related condition than in the unrelated condition for the second list-half [$t(27) = 3.037$, $p < .01$] and a marginally larger zigzag effect for the second compared with the first list-half in the related condition [$t(27) = 2.027$, $p = .053$]. In delayed free recall, only a main effect of relatedness reached significance [$F(1,26) = 8.474$, $MS_e = 0.298$, $p < .01$] because of a zigzag effect in the related, but not in the unrelated, condition, whereas in continuous-distractor free recall, the zigzag effect was not affected by relatedness or list-half (both $ps > .10$). Compared with a zero baseline (no zigzag effect), only the

second list-half in the related condition in immediate free recall [$t(27) = 4.308$, $p < .001$] and both list-halves in the related condition in delayed free recall [first list-half, $t(26) = 3.198$, $p < .005$; second list-half, $t(26) = 3.075$, $p < .005$] showed a zigzag effect that deviated signifi-cantly from zero.

DISCUSSION

The major empirical aim of this study was to demon-strate a dissociation between short- and long-term recency with semantic similarity, which was predicted by our dual-store model of recency effects (Davelaar et al., 2005). We replicated a pattern discerned (but not analyzed) in the data of Glanzer and Schwartz (1971), who used lists consisting of pairs of weak semantic associates. Our model predicted this pattern, which we referred to as a zigzag effect, to be present in immediate, but not in continuous-distractor, free recall, suggesting a new dissociation between the two recall tasks. In addition, the model predicted that in im-mediate free recall, the zigzag effect would be larger in the second than in the first half of the list, but that in delayed free recall, the zigzag effect would not vary with list-half.
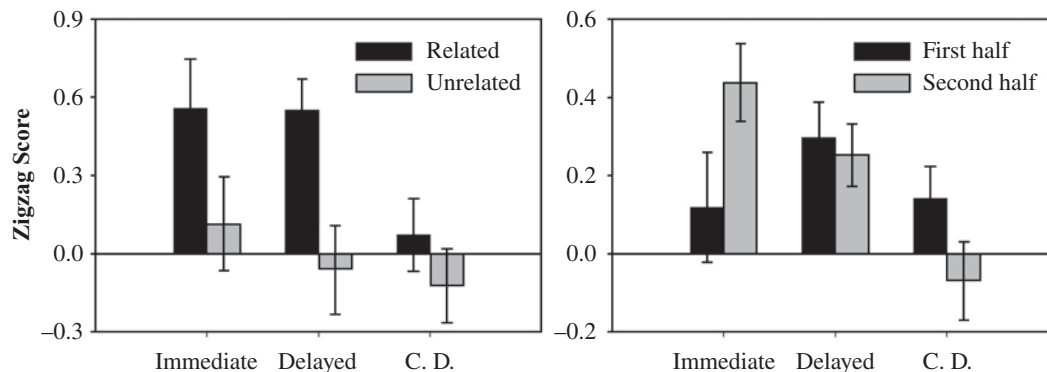

Figure 5. (Left) Experimental results for the total grouped zigzag scores for the three free recall tasks. The larger the value, the larger the overall zigzag pattern seen in the raw recall data in Figure 4. (Right) Grouped zigzag scores for the related condition only, broken down by list-half. Lines are standard error bars. C. D., continuous distractor.

Our empirical results supported these model predictions, lending further support for the view that short- and long-term recency are due to separate mechanisms.

Relations With Other Computational Theories

We have implemented these mechanisms within an activation-based dual-store model with changing context. However, we do not see the results as restricted to our specific model. On the contrary, we believe that they can be obtained in any generic framework that includes a short-term buffer, such as SAM (Raaijmakers & Shiffrin, 1981), eSAM (extended search of associative memory model; Sirotin, Kimball, & Kahana, 2005), and ACT–R (Anderson et al., 1998). Our aim, rather, is to highlight that a short-term buffer is an important component in accounting for data in free recall paradigms, whatever the specifics of its implementation. In such a generic framework, short-term recency is due to the most recently presented items still residing in a short-term buffer, from which they are reported without error, whereas long-term recency is due to a retrieval mechanism that operates on a changing context representation; its accuracy depends on the similarity between encoding and retrieval context (Dennis & Humphreys, 2001; Howard & Kahana, 2002a; Mensink & Raaijmakers, 1988).

Recently, Kahana and colleagues (Sirotin et al., 2005) have used a SAM framework to account for semantic effects in free recall, which are difficult to address in a single-store model. As in our model, eSAM includes a short-term buffer in which items have semantic links; this buffer helps account for category clustering and semantically based recall transitions. We have shown that in our model, the zigzag effect is due to associates that are simultaneously active in the buffer supporting each other (via excitatory semantic links), leading the first member of a pair to stay in the buffer longer. As in immediate and delayed free recall, consecutively presented items reside in the buffer together, and zigzag effects are found in these tasks, but not in continuous-distractor free recall, in which the distractor activity prevents coactivation of consecutively presented (semantically related) words. Moreover, the larger zigzag effect at recency in immediate free recall is consistent with the view that recency items in this task are retrieved from an STM store. The finding that a list-half effect is absent in delayed free recall further supports the notion that all items in this task are reported from episodic memory and that the zigzag effect is due to transfer from the activation-based component (STM) during encoding.

Implications for the Single–Dual Store Debate

Although our results undermine the arguments against a short-term buffer based on associations between these tasks (Greene, 1986), we still need to address one concern. Because the recency effect is larger in size in immediate than in continuous-distractor free recall, one could object that this factor confounds the dissociation (the single-store theorist would then need to explain the difference in magnitude, although the ratio between the interpresentation

and retention intervals was identical in the two tasks) (e.g., Bjork, 2001; Nairne et al., 1997). To rule out this concern, we conducted a supplementary analysis that was restricted to a subset of those participants ($n = 19$) in the continuous-distractor group whose average long-term recency effect was identical to that of the short-term recency effect in the immediate free recall group (30% drop-off in recall over the last four serial positions). Although the effect of relatedness was significant, no zigzag effect was obtained in continuous-distractor free recall for the second list-half in the related condition [$t(18) = 0.42, p = .68$], indicating that the dissociation with the zigzag effect is not an automatic function of the recency magnitude. This was further indicated by the finding of a significant zigzag effect [$t(16) = 2.17, p = .016$] for a subset of participants ($n = 17$) from the immediate free recall group whose short-term recency was identical to that of the continuous-distractor group (15% drop-off).

In principle, one cannot rule out that a single-store contextual retrieval model that accounts for the dissociations reported here may be developed, but this remains a challenge for the future. However, we believe that this will require the introduction of additional, new processing assumptions that implicitly dissociate continuous-distractor, immediate, and delayed free recall (beyond standard assumptions that distractors replace items from the focus of attention and allow context to change).[6] Moreover, such a model will need to explain an increasing list of dissociations between short- and long-term recency effects (medial temporal lobe amnesia, proactive interference, directed output order, lag recency effects) that were mentioned in the introduction (see Davelaar et al., 2005, for a detailed discussion).

Implications for Semantic Processing in List Memory

We have used a set of simplifying assumptions—which were motivated by our experimental manipulation—about the nature of semantic processing. These simplifications involve the presence of only two levels of associative strength, unrelated ($\alpha_2 = 0$) or related ($\alpha_2 > 0$), and of a single associative connection per word. Despite these simplifications, the model has been able to make a strong, testable prediction. Future computational work could incorporate graded associative-strength levels by using, for example, *latent semantic analysis* (LSA; Landauer & Dumais, 1997) or published association norms (Nelson, McEvoy, & Schreiber, 1998). This would allow not only for quantitative modeling, but also for a detailed consideration of semantic intrusions in free recall (e.g., Sirotin et al., 2005).

In this article, we have focused on a dissociation between short- and long-term recency with semantic similarity, showing a zigzag pattern in short-term but not in long-term recency. Also by measuring the effect of semantic similarity (as quantified by LSA) on *conditional response probabilities* (CRPs) for retrieval of successive words, Howard and Kahana (2002b) reported another dissociation between immediate (and delayed) and continuous-distractor free recall. These researchers reanalyzed their previous ex-

periment (Howard & Kahana, 1999) and found that the conditional response probability of retrieving word i after retrieving word j depends not only on the proximity of words i and j in the list, but also on the semantic similarity between the two words. Critically, the dissociation is due to the impact of associative strength on recall transitions between successive words in the list, which was larger in delayed free recall than in the continuous-distractor task.

This dissociation, involving a larger semantic clustering in delayed than in continuous-distractor free recall, can easily be accommodated in a dual-store framework in which associative interconnections are formed between items while they reside in a short-term buffer. This approach has been used successfully in eSAM (Sirotin et al., 2005). This version of SAM does not include a changing context representation necessary for accounting for long-term recency effects, but includes a matrix of LSA-based associative connections (separate from the episodic item–context connections) that enables the model to account for clustering effects. We have not added such a component (i.e., SAM-like strengthening of associations between items in the buffer), because, although it would account for clustering effects, it would produce a sharper lag recency function for delayed as compared with continuous-distractor free recall; such a sharper lag recency function would be inconsistent with data that indicate temporal scale invariance of lag recency (Howard & Kahana, 1999). The sharpening of lag recency in the delayed free recall condition should be expected (within this approach) because both contextual overlap and buffer-induced associations would contribute to the episodic retrieval of items in delayed free recall, but only contextual overlap would contribute to episodic retrieval in continuous-distractor free recall.

Our model can nevertheless be extended so that it would account for both dissociations in clustering (i.e., the larger clustering in delayed as opposed to continuous-distractor free recall) and the scale invariance of lag recency for lists of unrelated words. The approach we explored makes a weaker assumption about the nature of associative enhancement—namely, that the connections between co-active items that *already* have some associative links are strengthened during encoding.[7] This is consistent with theories of the role of the prefrontal cortex in memory encoding. Some of these theories assume that the prefrontal cortex is involved in the maintenance of semantic information in working memory (Gabrieli, Poldrack, & Desmond, 1998) and that it plays a role in the strengthening of associations between weakly related words in a memory list presented for free or cued recall (Fletcher, Shallice, & Dolan, 2000; Moscovitch, 1994).[8]

To explore such a mechanism in our model with minimal modifications, we can assume a larger value of the semantic-retrieval parameter, g, between words of an associated pair in the immediate and delayed free recall tasks (where the associated words were coactive in the buffer) than in the continuous-distractor free recall task (where the associated words were not coactive in the buffer).[9] We ran a supplementary simulation study in which we used a g value of 1.75 in delayed free recall and a g value of 1.4 in continuous-distractor free recall, with lists of six pairs of related words (as in our experiment).

We computed a measure of semantic clustering, which is analogous to the interaction between relatedness (as measured by LSA) and lag recency (Howard & Kahana, 2002b). In our simulation, the clustering measure was computed as the difference between the sum of the transitions within associative pairs and the sum of the transitions outside associative pairs. Specifically, the CRP for a $+1$ transition for odd items was summed with the CRP of a -1 transition for even items (within-pair transitions). From this we subtracted the sum of the CRP for a -1 transition for odd items and the CRP of a $+1$ transition for even items (outside-pair transitions). This resulted in a semantic clustering effect of .58 in delayed free recall and of .40 in continuous-distractor free recall. These results are consistent with the clustering data reported by Howard and Kahana (2002b). Importantly, the zigzag dissociation was not affected by these changes. This account is tentative, and future studies recording the order of the memory recall will be important in order to further examine the various effects (zigzag, contiguity, clustering, etc.) of semantic relations between words in list memory.

Conclusions

In summary, we have presented a novel dissociation between short- and long-term recency with semantic similarity, which was predicted by our dual-store model. The model attributes the overall main effect of relatedness to semantic facilitation during encoding (coactive associates support each other in STM) and retrieval (retrieval of an item facilitates retrieval of its associate). However, the zigzag effect was due entirely to semantic facilitation between associates that resided in an activation-based short-term buffer during encoding. Whereas the semantic facilitation at retrieval was independent of the processes that underlie short- and long-term recency (active maintenance and similarity in a study–test context), the semantic facilitation during encoding operated in the same system that leads to recency in immediate free recall, namely, the activation-based short-term buffer. These empirical results lend further support to the view that a dual-store theory, combining a short-term buffer with a contextually based LTM encoding and retrieval mechanism, is a viable and fruitful theory of dissociations and associations in short- and long-term recency.

REFERENCES

ANDERSON, J. R., BOTHELL, D., LEBIERE, C., & MATESSA, M. (1998). An integrated theory of list memory. *Journal of Memory & Language*, **38**, 341-380.

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, pp. 89-195). New York: Academic Press.

BADDELEY, A. D. (1972). Retrieval rules and semantic coding in short-term memory. *Psychological Bulletin*, **78**, 379-385.

BADDELEY, A. D., & HITCH, G. J. (1974). Working memory. In G. H.

Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89). New York: Academic Press.

Baddeley, A. D., & Hitch, G. J. (1977). Recency re-examined. In S. Dornic (Ed.), *Attention and performance VI* (pp. 647-667). Hillsdale, NJ: Erlbaum.

Bjork, R. A. (2001). Recency and recovery in human memory. In H. L. Roediger III, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 211-232). Washington, DC: American Psychological Association.

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

Carlesimo, G. A., Marfia, G. A., Loasses, A., & Caltagirone, C. (1996). Recency effect in anterograde amnesia: Evidence for distinct memory stores underlying enhanced retrieval of terminal items in immediate and delayed recall paradigms. *Neuropsychologia*, **34**, 177-184.

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

Crowder, R. G. (1982). The demise of short-term memory. *Acta Psychologica*, **50**, 291-323.

Crowder, R. G. (1993). Short-term memory: Where do we stand? *Memory & Cognition*, **21**, 142-145.

Dalezman, J. J. (1976). Effects of output order on immediate, delayed, and final recall performance. *Journal of Experimental Psychology: Human Learning & Memory*, **2**, 597-608.

Davelaar, E. J. (2003). *Active memory: Its nature and its role in mnemonic and linguistic behaviour.* Unpublished doctoral thesis, University of London.

Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The demise of short-term memory revisited: Empirical and computational investigations of recency effects. *Psychological Review*, **112**, 3-42.

Davelaar, E. J., & Usher, M. (2002). An activation-based theory of immediate item memory. In J. A. Bullinaria & W. Lowe (Eds.), *Connectionist models of cognition and perception: Proceedings of the Seventh Neural Computation and Psychology Workshop* (pp. 118-130). Singapore: World Scientific.

Dennis, S., & Humphreys, M. S. (2001). A context noise model of episodic word recognition. *Psychological Review*, **108**, 452-478.

Fletcher, P. C., Shallice, T., & Dolan, R. J. (2000). "Sculpting the response space"—An account of left prefrontal activation at encoding. *NeuroImage*, **12**, 404-417.

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

Glanzer, M. (1972). Storage mechanisms in recall. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 5, pp. 129-193). New York: Academic Press.

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

Glanzer, M., & Schwartz, A. (1971). Mnemonic structure in free recall: Differential effects on STS and LTS. *Journal of Verbal Learning & Verbal Behavior*, **10**, 194-198.

Glenberg, A. M., Bradley, M. M., Kraus, T. A., & Renzaglia, G. J. (1983). Studies of the long-term recency effect: Support for a contextually guided retrieval hypothesis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **9**, 231-255.

Glenberg, A. M., & Swanson, N. G. (1986). A temporal distinctiveness theory of recency and modality effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 3-15.

Greene, R. L. (1986). A common basis for recency effects in immediate and delayed recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 413-418.

Greene, R. L. (1992). *Human memory: Paradigms and paradoxes.* Hillsdale, NJ: Erlbaum.

Greene, R. L., & Crowder, R. G. (1984). The effect of semantic similarity on long-term recency. *American Journal of Psychology*, **97**, 441-449.

Haarmann, H. [J.], & Usher, M. (2001). Maintenance of semantic information in capacity-limited item short-term memory. *Psychonomic Bulletin & Review*, **8**, 568-578.

Healy, A. F., & McNamara, D. S. (1996). Verbal learning and memory: Does the modal model still work? *Annual Review of Psychology*, **47**, 143-172.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. New York: Wiley.

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**, 75-116.

Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, & 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 & Language*, **46**, 85-98.

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

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

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

Moscovitch, M. (1994). Cognitive resources and dual-task interference effects at retrieval in normal people: The role of the frontal lobes and medial temporal cortex. *Neuropsychology*, **8**, 524-534.

Murdock, B. B., Jr. (1967). Recent developments in short-term memory. *British Journal of Psychology*, **58**, 421-433.

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

Nairne, J. S., Neath, I., Serra, M., & Byun, E. (1997). Positional distinctiveness and the ratio rule in free recall. *Journal of Memory & Language*, **37**, 155-166.

Neath, I. (1993). Distinctiveness and serial position effects in recognition. *Memory & Cognition*, **21**, 689-698.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). *The University of South Florida word association, rhyme, and word fragment norms*. Available from w3.usf.edu/FreeAssociation/.

Pecher, D., & Raaijmakers, J. G. W. (1999). Automatic priming effects for new associations in lexical decision and perceptual identification. *Quarterly Journal of Experimental Psychology*, **52A**, 593-614.

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

Raaijmakers, J. G. W. (1993). The story of the two-store model of memory: Past criticisms, current status, and future directions. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 467-488). Cambridge, MA: MIT Press.

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

Schrijnemakers, J. M. C., & Raaijmakers, J. G. W. (1997). Adding new word associations to semantic memory: Evidence for two interactive learning components. *Acta Psychologica*, **96**, 103-132.

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**, 787-805.

Talmi, D., Grady, C. L., Goshen-Gottstein, Y., & Moscovitch, M. (2005). Neuroimaging the serial position curve: A test of single-store versus dual-store models. *Psychological Science*, **16**, 716-723.

Tan, L., & Ward, G. (2000). A recency-based account of the primacy effect in free recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 1589-1625.

Tzeng, O. J. L. (1973). Positive recency effect in delayed free recall. *Journal of Verbal Learning & Verbal Behavior*, **12**, 436-439.

Usher, M., & Cohen, J. D. (1999). Short-term memory and selection processes in a frontal-lobe model. In D. Heinke, G. W. Humphreys, & A. Olson (Eds.), *Connectionist models in cognitive neuroscience* (pp. 78-91). New York: Springer.

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

Whitten, W. B. (1978). Output interference and long-term serial position effects. *Journal of Experimental Psychology: Human Learning & Memory*, **4**, 685-692.

NOTES

1. Howard and Kahana (1999) reported a further dissociation, involving an interaction between lag recency and output order in immediate, but not in continuous-distractor, free recall. Although their article focuses mainly on the association between immediate and continuous-distractor free recall with lag recency, in immediate, but not in continuous-distractor free recall, the lag recency varies for the first few items reported (see note 8 in Howard & Kahana, 1999; for an STM interpretation of this finding, see Davelaar et al., 2005).

2. The rationale of using weak associates is to diminish the contribution of semantic guessing.

3. Difference(n) = $P(n) - [P(n+1) + P(n-1)]/2$, where n, $n+1$, and $n-1$ denote the serial position of the item, its right neighbor, and its left neighbor, respectively, and P denotes the recall probabilities of the corresponding items. The difference score analysis of the zigzag effect did not include Serial Positions 1 and 12, since the difference score is undefined at these two serial positions.

4. A smooth serial position function may still generate a zigzag score at recency (second half) because of its convexity/concavity. However, this effect will be negligible.

5. Note that although the typical estimate for the buffer capacity is about four items (see Cowan, 2001), the presence of semantic associates can increase the effective capacity (see Davelaar et al., 2005). Furthermore, as the displacement from the buffer is probabilistic, items from Positions 7 and 8 still have a chance (albeit smaller than that for items from Positions 9–12) to be reported directly from the buffer.

6. The distinction between a single- and a dual-store model is not always obvious. For example, although the temporal context model (Howard & Kahana, 2002a) is interpreted as a single-store model, one may also interpret it as involving two (activation/weight) components. This is most clear in Howard, Fotedar, Datey, and Hasselmo (2005), where an activation-based memory is localized in the entorhinal cortex and a weight-based memory is localized in the hippocampus.

7. Priming studies have indicated that items without preexperimental associations require multiple paired presentations before facilitatory semantic priming effects arise (Pecher & Raaijmakers, 1999; Schrijnemakers & Raaijmakers, 1997). One adaptive possibility is that the associative enhancement of coactive items is largest for the weak-to-middle range of preexperimental association levels (according to the sensitivity of a sigmoidal function).

8. At present, data are not available about similar processes in continuous-distractor paradigms, when words are separated by distractor intervals.

9. The assumption is that the g parameter, which represents the semantic-retrieval strength, is enhanced during the encoding processes for pairs of words that have a preexisting association and were coactive in the buffer during encoding. Thus, at retrieval, g is larger for associates in delayed than in continuous-distractor free recall.

APPENDIX
Model Equations and Parameters

The model consists of two interconnected systems: a lexical-semantic and an episodic changing context representation. Episodic learning involves a matrix of connections between the context and the item layers. These connections are made during encoding, linking the active units in the lexical layer with the active unit in the context layer. During retrieval, items can be unloaded from the buffer or can be reported, following a slower, competitive retrieval process. This process is driven by the active context via the episodic links (retrieval from LTM) and by the semantic association between the recalled item and its associates (via the semantic links in the lexical layer).

Lexical-Semantic Layer: Activation-Based Buffer

The input activations of the lexical units, labeled as x_i, are updated in parallel as follows:

$$x_i(t+1) = \lambda x_i(t) + (1-\lambda)\left\{\alpha_1 F\left[x_i(t)\right] + \alpha_2 F\left[x_{i*}(t)\right] - \beta\Sigma F\left[x_j(t)\right] + I_i(t) + \xi\right\}, \tag{A1}$$

where t is the iteration time step, λ is the decay time constant, α_1 and α_2 ($0 < \alpha_2 < \alpha_1$) are the parameters for self-excitation and excitation from associated items, $i*$, β is the global inhibition, and I is the sensory input supplemented with zero-mean Gaussian noise, ξ, with standard deviation σ. F is the output activation function: $F(x > 0) = x/(1 + x)$; $F(x < 0) = 0$ (Haarmann & Usher, 2001; Usher & Cohen, 1999).

We model the list presentation by sequentially turning on the inputs I_1 to I_L (L = list length) in Equation A1 for 500 iterations. An illustration of the activation in the lexical layer upon presentation of a list of unrelated or related word pairs is shown in Figure A1. When the activation level of an item is active above the threshold (horizontal line), the item is said to be in the buffer. Notice that activated items displace other items from the buffer (top panel) in the unrelated condition, but that associated items tend to be displaced together (bottom panel). This leads to the first member of a pair acquiring a stronger episodic trace compared to the second pair member.

Context Layer

The context layer consists of a linear array of context units, of which, at each time-step, only one is activated. During list presentation and during retrieval, the context changes with probabilities P^+ and P^- ($P^+ > P^-$) to the next or previous context unit, implementing a biased random-walk process along a linear array.
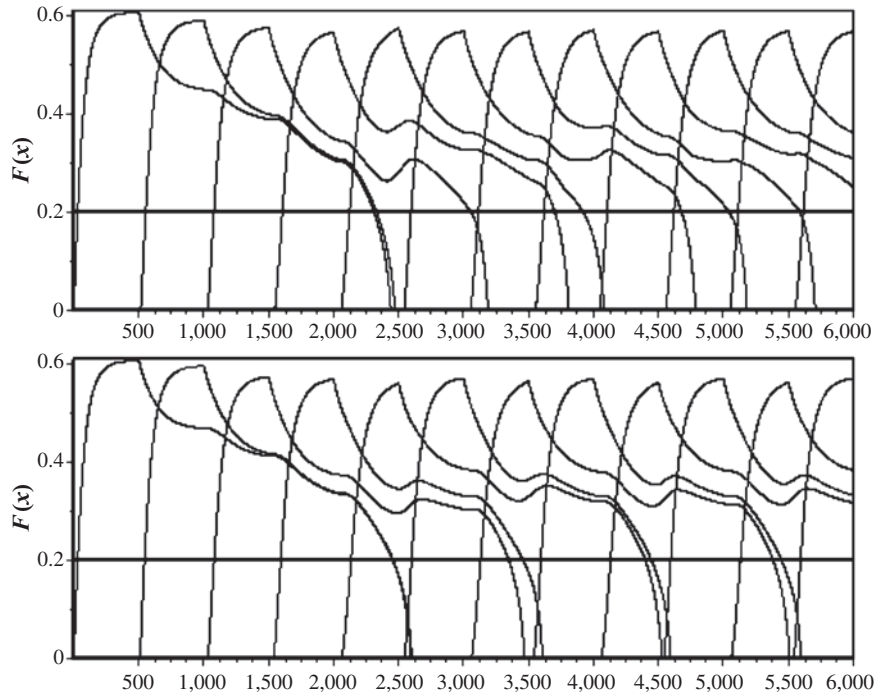
Figure A1. Activation levels of 12 items, presented for immediate free recall, in the unrelated (top) and pairwise related (bottom) conditions. The horizontal and vertical axes show iteration number (time step) and activation level, respectively.

Episodic Encoding

Activated item representations become associated with active context units, forming a matrix of episodic connections with strengths W_i that are set to zero and grow linearly with the time integral of each item's activation above a threshold, φ_1:

$$W_i = \int_0^t \theta\big[F(x_i) - \varphi_1 \big] dt, \tag{A2}$$

where $\theta(x) = \varepsilon x$ for $x > 0$ and $\theta(x) = 0$ for $x < 0$ is the Heaviside θ function, with ε being the learning rate parameter.

Episodic Retrieval

At retrieval, active items are reported first, followed by retrieval of inactive items. Context continues to change during retrieval, with k attempts for the end context and k attempts for the start context. Episodic retrieval is a competitive process, where the selection probability, $P^{sel}(i|c,j)$, is the probability that item i is selected for retrieval given a previously retrieved item j and context c. For $j \neq i^*$ (the item associated with i),

$$P^{sel}(i \mid c, j) = \frac{\exp(W_{c,i})}{\sum \exp(W_{c,j})}. \tag{A3}$$

When $j = i^*$ (i.e., for items associated with the previously retrieved item), we assume an additional input to item i of value g, which corresponds to an associative retrieval process from i^* to i.[A1]

$$P^{sel}(i \mid c, i^*) = \frac{\exp(W_{c,i} + g)}{\sum \exp(W_{c,j} + g)}. \tag{A4}$$

Whether the selected item will be reported depends on a probabilistic process of recovering its phonological form, in which the total input to the selected item is compared with a retrieval threshold, φ_2.

$$P^{rec}(i \mid c, i^*) = \frac{1}{1 + \exp\big[\varphi_2 - (W_{c,i} + g_i) \big]\big/ \tau}. \tag{A5}$$

To limit the number of free parameters, we kept all parameter values in the simulation the same as we have used in previous work (Davelaar et al., 2005): $\lambda = 0.98$, $\alpha_1 = 2.0$, $\beta = 0.20$, $I = 0.33$, $\sigma = 1.0$, $\varepsilon = 0.02$, $\varphi_1 = 0.2$, $\varphi_2 = 4$, $k = 20$, $\tau = 2.0$, $P^+ = .10$, $P^- = .05$. Only $\alpha_2 = 0.08$ and $g = 1.75$ were free to vary and were estimated through a grid search ($\alpha_2 = 0.02$ to 0.10 step 0.02; $g = 1.0$ to 2.0 step 0.5), in order to find those values that closely matched the whole pattern (serial position function and zigzag scores) in our data. In the simulations, each item was presented for 500 iterations, and each simulation consisted of 1,000 trials. The system had 48 units (only 12 were activated).

In the model, short- and long-term recency effects are due to different mechanisms. Short-term recency is due to the last few presented items' still being in an active state (they are still in the short-term buffer) and therefore being reported, whereas long-term recency is due to the last presented items' having a higher likelihood of being associated with context units that are active at the start of the retrieval phase (cf. Bjork & Whitten, 1974; Glenberg et al., 1983; Howard & Kahana, 1999, 2002a). We showed that the model not only accounted for associations between the two types of recency effects (with a manipulation of list length; Greene, 1986), but also for several critical dissociations in the experimental and neuropsychological literature (see Davelaar et al., 2005).

In the model, short- and long-term primacy effects are due to different mechanisms. Primacy effects in immediate and delayed free recall are mainly due to the first few items' residing in the buffer longer than middle list items, leading to greater episodic strengths, and are due, to a smaller degree, to reinstatement of the start context during retrieval. Primacy effects in continuous-distractor free recall are entirely due to the reinstatement of the start context. Therefore, the model produces primacy effects without recourse to a rehearsal mechanism, consistent with data showing that when measures are taken that eliminate the use of a rehearsal strategy, small primacy effects are still found (Baddeley & Hitch, 1974, 1977; Howard & Kahana, 1999).

NOTE

A1. This component is monotonically related to the interitem association parameter, α_2. However, since we are not modeling the detailed dynamics of the retrieval process, we capture its effect at retrieval via the impact of g.