

Maintenance of semantic information in capacity-limited item short-term memory

HENK HAARMANN

University of Maryland, College Park, Maryland

and

MARIUS USHER

Birkbeck College, University of London, London, England

We report a semantic effect in immediate free recall, which is localized at recency and is preserved under articulatory suppression but is highly reduced when recall is delayed after an intervening distractor task. These results are explained by a neurocomputational model based on a limited-capacity short-term memory (STM) store, consisting of *activated* long-term memory representations. The model makes additional predictions about serial position functions in semantically cued recall, indicating capacity limitations caused by a displacement type mechanism, which are confirmed in a second experiment. This suggests that in addition to the phonological component in verbal STM, there is an activation/item-limited component with semantically sensitive representations.

Theoretical views on memory organization are divided about the existence and the nature of short-term memory (STM). Although some theorists have proposed that a single unified system can account for the data associated with both STM and long-term memory (LTM; Crowder, 1993; Howard & Kahana, 1999; Lewandowsky & Murdock, 1989), most of the models developed to explain the data on immediate memory rely on distinctive STM (or primary memory) and LTM (or secondary memory) processes (Atkinson & Shiffrin, 1968; Baddeley, 1986; Burgess & Hitch, 1999; Cowan, 1999; Hulme, Roodenrys, Schweickert, Brown, & Gordon, 1997; Nairne, 1990; Raaijmakers & Shiffrin, 1981; Schweickert, 1993). Within this framework, two aspects of verbal STM have generated a strong debate. The first one involves the code for information in verbal STM (phonological or semantic), and the second involves the nature of its capacity limitation (decay vs. displacement from a limited item store). These two aspects of STM are the main target of the present work. In particular, we present data and a computational model that support the existence of a limited-capacity STM system (Cowan, 2001) with lexical/semantic representations.

Robust phonological effects in STM (e.g., Baddeley, 1966; Conrad, 1965) and weaker semantic effects that are also present and of equal size (or larger) in delayed recall tasks (Baddeley & Ecob, 1970; Craik & Levy, 1970; Glanzer & Schwartz, 1971; Levy & Baddeley, 1971) have motivated the view that the information in STM is exclusively phonological and that all semantic effects in STM are mediated by LTM contributions. This view is best summarized in a review by Alan Baddeley: "all cases in which semantic factors have been shown to affect primary memory are attributable to the subject's utilization of retrieval rules, that is, techniques and strategies stored in secondary memory but that facilitate the retrieval of phonemically coded information from primary memory" (Baddeley, 1972, p. 379; a similar view is presented by Crowder, 1979; see however, Shulman, 1970, 1972, and Raser, 1972, for a strong challenge to it and Baddeley, 2000, for a more recent and broader view). The assumption of exclusive phonological STM traces was then included in many influential STM models, such as Baddeley's (1986) model of working memory, the network model of the phonological loop (Burgess & Hitch, 1999), and the reintegration model (Hulme et al., 1997; Schweickert, 1993). Other models, such as TODAM (Murdock, 1982) and the feature model (Nairne, 1990), use representations that, in principle, may include semantic features. However, since in these models, semantic similarity leads to a deterioration in performance (owing to interference), they cannot account for semantic effects in STM, which are facilitatory with regards to item (but not order) information (Poirier & Saint-Aubin, 1995; Saint-Aubin & Poirier, 1999).

In the recent literature, two types of explanations for semantic effects in STM seem to be available. Among

This work was reported at the 40th Annual Meeting of the Psychonomic Society, Los Angeles, November 1999. We thank Sandra Carter, Kimberly Russo, David Vernon, and Eddy Davelaar for running experiments, Jonathan Cohen, Todd Braver, Randy O'Reilly, Mark Chappell, and Eddy Davelaar for helpful and stimulating discussions, and Jean Saint-Aubin and two anonymous referees for their sharp criticism. Correspondence concerning this article should be addressed either to H. Haarmann, Department of Hearing and Speech Sciences, University of Maryland, College Park, MD 20742 (e-mail: hhaarmann@hesp.umd.edu) or to M. Usher, School of Psychology, Birkbeck College, University of London, London WC1E 7HX, England (e-mail: m.usher@bbk.ac.uk).

the models discussed above, the reintegration model was proposed to provide an explanation for the enhancement of item information by semantic relations in STM (Saint-Aubin & Poirier, 1999; Stuart & Hulme, 2000). According to this, semantic effects are due solely to the use of retrieval cues that act on the reconstruction of decaying phonological traces in STM. Alternatively, it has been proposed by Martin and collaborators (R. C. Martin & Romani, 1994; R. C. Martin, Shelton, & Yafee, 1994; Romani & Martin, 1999), on the basis of cognitive and neuropsychological dissociations, that semantic traces are actually being maintained within a component of the STM system associated with the prefrontal cortex. This system may rely on activated representations, as has been indicated by recent neuroimaging studies (Crosson et al., 1999), and is consistent with models such as SAM (Raaijmakers & Shiffrin, 1981) and spreading activation models, which view STM as the activated components of LTM (Anderson, 1983; see also N. Martin, Saffran, & Dell, 1996) but which have not yet accounted for semantic effects in STM. Here, we report a semantic effect in STM, that is strongly reduced when recall is delayed after a distractor task, therefore leading us to reject an explanation on the basis of an episodic LTM contribution to the STM (Baddeley, 1972). The data are explained, however, with a computational model that is similar to the STM store in SAM (in showing displacement-type forgetting) but which views STM as *activated LTM representations* in the prefrontal cortex (Usher & Cohen, 1999). To distinguish between this model and models based on reintegration theories (which assume decay-based forgetting), we also examine data related to STM capacity limitations (by contrasting decay vs. displace-

ment). These two types of data provide strong constraints for models that address the retention of item information in STM.¹

In the next section, the model is briefly introduced, focusing on two predictions: (1) semantic effects in STM that are not due to episodic LTM contributions and (2) a capacity limitation that results in an interaction between serial position and memory load. These predictions are then tested in Experiments 1 and 2.

Computational Model

In a previous article, we presented a neurocomputational model based on a limited-capacity STM store (thought to be related to the prefrontal cortex) consisting of *activated LTM representations* (Usher & Cohen, 1999). The model was shown to explain various data patterns in immediate recall of item information, such as effects of load in the Brown–Peterson task and effects of presentation rate in cued recall (Waugh & Norman, 1965). The model, as will be described below, is minimal in its complexity,² trying to target specifically the capacity-limited STM component that can be observed when the effects of phonological rehearsal and higher level chunking are prevented (Cowan, 2001) and when episodic LTM learning is factored out (Waugh & Norman, 1965).

In the model, conceptual item representations (e.g., words in the lexicon) correspond to cell assemblies connected by loops of recurrent self-excitation and competing via lateral inhibition (Figure 1A), whereas semantic associations correspond to weak excitatory links between the representations. The system is activated by bottom-up input from distributed representations in the posterior system, and its STM property is due to the self-

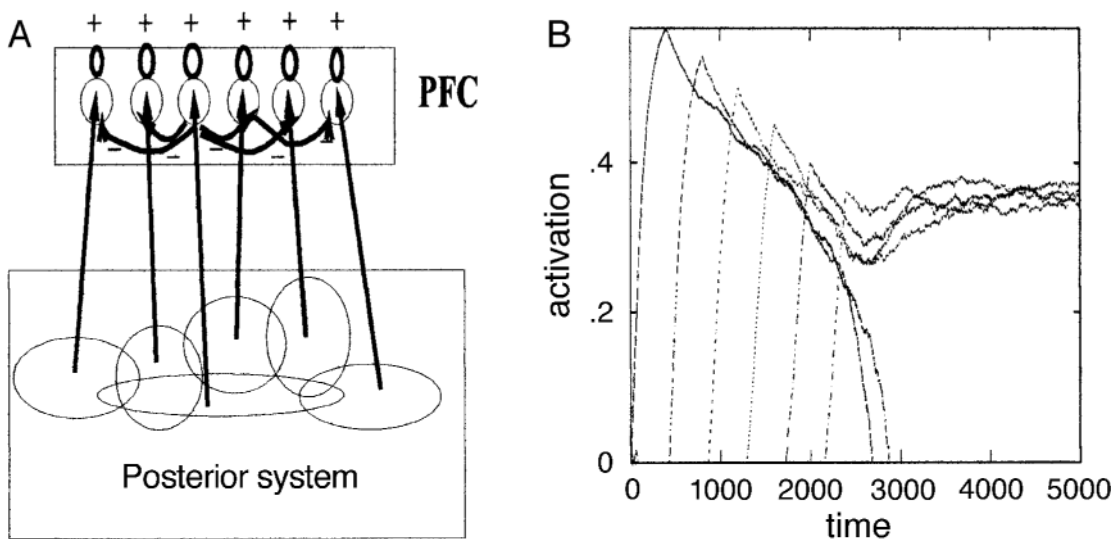


Figure 1. (A) Model description. Items in the capacity-limited short-term memory system, associated with the prefrontal cortex (PFC), are represented by cell assemblies, shown as circles. Horizontal arrows with minus signs indicate lateral inhibition among cell assemblies. The circles with plus signs on top of the cell assemblies indicate recurrent self-excitation. Each cell assembly receives bottom-up input from distributed representations in the posterior system, shown as partially overlapping circles. (B) Activation trajectories for a sequential presentation of six items, showing retention of four items and displacement of two items.

excitation loops that enable the item representations to remain in a state of activation after the offset of the sensory stimulation. This is observed in neurophysiological recordings from the prefrontal cortex (Funahashi, Chafee, & Goldman-Rakic, 1993; E. K. Miller, Erickson, & Desimone, 1996). Note that although the cell assemblies are part of the LTM structure, their activation mediates STM behavior. The competition mediated by lateral inhibition prevents an uncontrolled spread of activation in the system and leads to capacity limitations; when the capacity is exceeded one of the representations is deactivated. This is illustrated in Figure 1B, which shows the activation of the item representations in response to sequential input corresponding to a sequence of six words in a memory test. Although each representation is activated for only 400 iteration steps, its activity persists beyond the input termination. As more representations get activated, however, the activations decrease, owing to the lateral competition, until some of the representations are deactivated (the first two in this illustration). Also, as is shown in Figure 1B, when first activated, each representation reaches a level of activation higher than that of the previously activated representations. This is due to the fact that it benefits from sensory input (in addition to recurrent feedback, as the other representations do). Thus, when the capacity of the system is exceeded, earlier items are more likely to be deactivated (Figure 1B) than later items, whose activations are slightly higher. Moreover, since early items have more opportunity to be displaced (each time that a new item is activated), one obtains a recency effect.

Serial position curves are obtained by checking which units remain active at the end of a list trial and by averaging over 500 simulation trials. It is assumed that activated units in the system are subject to conscious awareness and can be directly reported. Examples of serial position curves are shown in Figures 3–4, demonstrating a recency effect. The model parameters and technical details are presented in the Appendix. We should note again that the model does not predict order information, which we assume to be mediated by another STM system (R. C. Martin & Romani, 1994; Romani & Martin, 1999; Warrington & Shallice, 1969).

Semantic Effects

The effects of semantic relations between retention items on recall can be tested in relation to an experiment performed by Glanzer (1969). In that experiment, lists of 14 words, consisting of seven pairs of semantically weak associates, were presented for immediate recall. The separation (i.e., number of intervening words) between the associates in the list was manipulated, and a continuous improvement in recall with increased proximity between the associates was obtained. The model's prediction is shown in Figure 2, replicating the semantic proximity effect: The number of words retained increases with proximity (in a range of 3–4 items). In the model, the effect takes place because, when associated items are proximal, they are more likely to coexist in an active state.

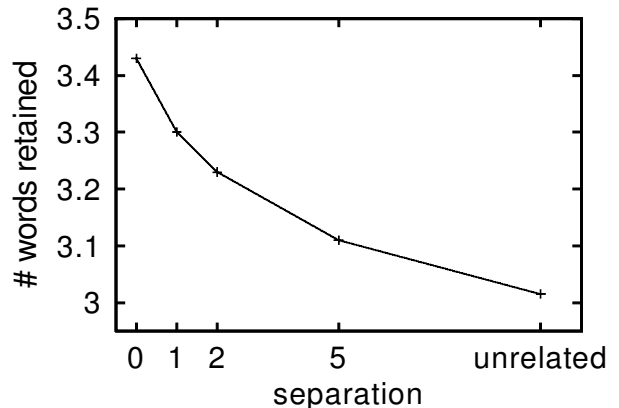


Figure 2. Semantic separation effects in the model. The number of words recalled decreases with the separation (i.e., number of intervening words) between associate pairs, toward a performance value that corresponds to unrelated words.

Coactive associated items support each other via the semantic links that cancel some of the mutual competition, increasing the effective capacity. Serial position effects from this simulation are reported in the next section, together with experimental results.

Capacity Limitation Effects

The model's displacement-type capacity limitation generates an interaction between memory load (i.e., number of items in the retention list) and serial position. In particular, the model predicts that increasing the memory load (e.g., from four to six items) will result in steeper recency functions, affecting item retention more negatively at the early than at the late serial positions (see Figure 3). In the model, this is due to the displacement type of capacity limitation. With higher load, the earlier items have a higher chance to be displaced from the active memory system, relative to recent items that are less affected (as is illustrated in Figure 1B). This prediction distinguishes competition-based capacity-limited models (where items are probabilistically replaced) from decay-based models (Burgess & Hitch, 1999; Hulme et al., 1997; Schweickert, 1993) or retroactive interference (RI) models (Nairne, 1990), in which recall performance is predicted to depend mainly on the distance (temporal for decay theory, or as number of intervening items in RI theories) between the probed item and the recall cue (but not on items before the probed item). This prediction is tested in Experiment 2, using category cued recall.

EXPERIMENT 1

Glanzer (1969) has proposed two nonexclusive hypotheses to explain semantic-separation effects in his experiment. The first one is consistent with our model: Associates support each other in STM. The second hypothesis is that the effect is due to more effective transfer to LTM. In this case, at retrieval the improved retention is due to a contribution from episodic LTM, consistent with the theories discussed earlier (Baddeley, 1972;

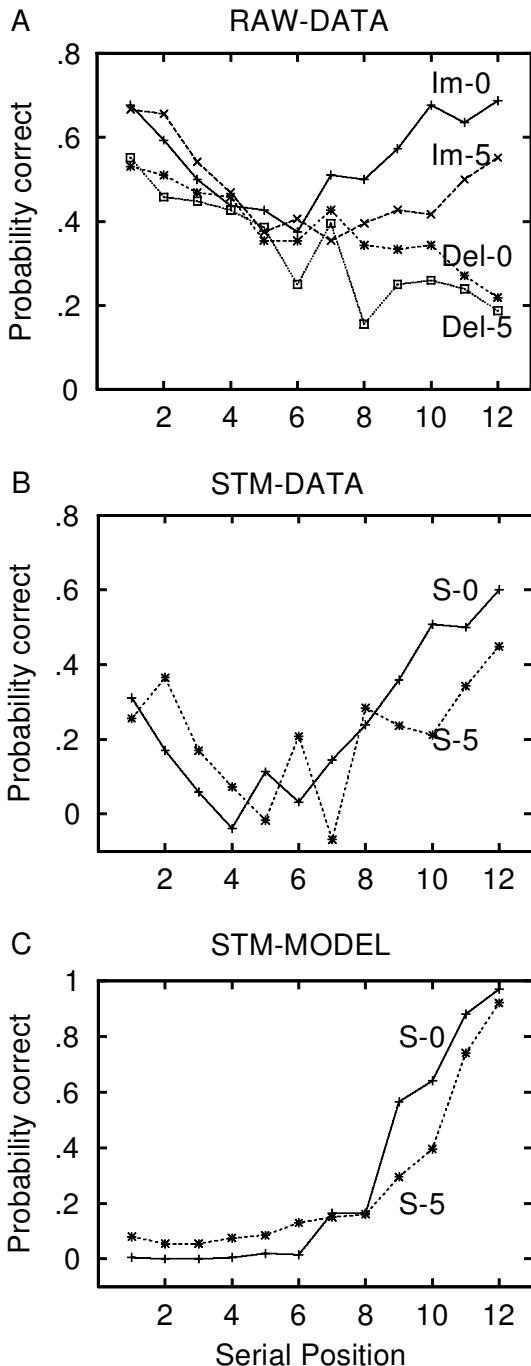


Figure 3. Recency curves in free recall. (A) Raw data in Experiment 1. Im, immediate; Del, delayed; symbols 0 and 5 correspond to the separation condition. (B) Short-term memory (STM) component estimated from the data. (C) STM maintenance in the model (S-0 and S-5 correspond to the separation conditions).

Crowder, 1979). The idea is that some information that was not retained in STM may have been transferred to episodic LTM and retrieved from there. At recall, a subject may rely on two distinct processes, which can suc-

ceed or fail independently (Baddeley, 1968, 1970; Waugh & Norman, 1965) in the retrieving of information: (1) a direct read-out from STM and (2) a retrieval attempt from episodic LTM.³

To demonstrate a direct contribution of coactive retention of semantic associates in STM, we ran a free recall experiment similar to that of Glanzer (1969), in which the associates can be either adjacent (separation-0 condition) or separated by five intervening words (separation-5 condition). In addition, the recall task was varied orthogonally—namely, immediate recall versus delayed recall with an interpolated distractor task (using sentence comprehension). Since distractor tasks are expected to empty the contents of verbal STM (Brown, 1958; Glanzer & Cunitz, 1967; Peterson & Peterson, 1959), we can estimate the LTM component to the recall on the basis of the performance in the delay task and use this to estimate the STM component (Baddeley, 1970; Glanzer & Schwartz, 1971; Levy & Baddeley, 1971; Waugh & Norman, 1965). If a semantic STM system exists, there should be a larger STM component in the separation-0 condition than in the separation-5 condition after factoring out the LTM component to recall, estimated on the basis of the delayed performance. Previous free recall studies, using this decomposition method, have not been successful in obtaining a semantic STM component. Those studies have typically reported equal or larger semantic effects in delayed recall than in immediate recall (or equal STM components in semantically related vs. unrelated conditions), leading to the conclusion that semantic effects reside solely in LTM (Baddeley, 1972; Craik & Levy, 1970; Glanzer & Schwartz, 1971; Levy & Baddeley, 1971). The estimation of the STM and LTM components has not been performed for the semantic-separation effect reported in the study of Glanzer. Neither has it been reported whether the semantic-separation effect can prevail across a distractor interval that is effective enough to wipe out the recency effect.

Method

Subjects and Design. Sixteen undergraduate students, who were native speakers of English, participated either for extra course credit or for a \$7 payment. The experiment conformed to a 2 (separation, separation-0 or separation-5) \times 2 (delay, immediate recall or delayed recall) within-subjects design.

Materials. We created a pool of 144 word pairs of semantic associates, using the same criteria as Glanzer (1969). In order to reduce guessing strategies, semantic associates were weak. In particular, each second word was given as a response to the first word by fewer than 3% of the subjects in the association norm group (e.g., *scissors-steel*; Palermo & Jenkins, 1964). From the pool of 144 word pairs, we created four lists of six trial sequences of 6 word pairs each. Assignment of a word pair to a list was without replacement, and across lists first words and second words were matched for word frequency, number of letters, number of syllables, and abstractness. There were two versions per list, corresponding to the two separation conditions. In the separation-0 condition, the two words in each of the 6 word pairs were immediately adjacent, whereas in the separation-5 condition, they were separated by five unrelated words (e.g., *broader horizon king leader comfort uneasy kittens fluffy get purchase eating lunch vs. broader*

king comfort kittens get eating horizon leader uneasy fluffy purchase lunch). We also created 12 sentence pairs for the interpolated distractor task in the delayed recall condition. Each pair consisted of a stimulus sentence broken down into six separately presented phrases and a verification statement (e.g., *The lawyer I that attended I the hearing I brought I the map I to the courtroom. An attorney was present at an investigation. True or False?*). Half of the verification statements were true, and half were false.

Procedure. Stimulus presentation was visual and computer controlled. The events on a trial were as follows. The subjects silently read a sequence of 12 words, which appeared at the center of the display (1 word/sec). A recall cue ("??") prompted the subjects to report as many words as they could. In the immediate condition, the recall cue appeared immediately after the offset of the last word, whereas in delayed recall, a sentence verification task lasting about 10 sec was interspersed between the offset of the last word and the onset of the recall cue. The subjects did not know at the start of a trial whether the task would be immediate recall or delayed recall.

In the sentence verification task, the subjects had to silently read a relative clause sentence, presented one phrase at a time at the center of the display, followed by a verification statement, and to make a true/false forced-choice response. The instructions stressed that the subjects should aim for perfect accuracy on the sentence verification task even if it interfered with their recall performance. Each subject performed 6 trials per condition, for a total of 24 trials. Assignment of the four lists of trial sequences to conditions was counterbalanced, so that each unique word pair occurred in all four conditions but was seen only once per subject.

Analysis: STM/LTM components. LTM and the STM contributions to the recall were estimated by using the same method as that in previous recall studies (Baddeley, 1970; Craik & Levy, 1970; Glanzer & Schwartz, 1971; Levy & Baddeley, 1971). At every serial position, it is assumed that $P(\text{IM}) = P(\text{STM}) + [1 - P(\text{STM})]P(\text{LTM})$, where $P(\text{IM})$ is the recall probability in the immediate condition, $P(\text{STM})$ and $P(\text{LTM})$ are the STM and LTM independent contributions to it, and $P(\text{LTM})$ is estimated on the basis of the delayed recall $P(\text{LTM}) = P(\text{delay})$. The STM contributions are obtained from the raw recency data [$P(\text{IM})$ and $P(\text{delay})$] and from the formulas above. Since the model described does not include an LTM learning component, the estimation of the STM component from the data is helpful for the purpose of comparison with the prediction of the model.

Results and Discussion

The serial position curves for the four conditions in the experiment are reported in Figure 2A, showing that the distractor interval was successful in totally eliminating recency effects. The serial position curves for the estimated STM component (obtained after factoring out the episodic LTM contribution) are as predicted by the model. There is a larger STM component (over the last four serial positions) when the associates are successive (54%) than when they are distant [31%; $F(1, 15) = 19.7$, $MS_e = 0.016$, $p < .001$; Figure 3].⁴ The model suggests a mechanism that generates the effect. Semantically related items that are adjacent are more likely to remain coactive in the STM store and, thus, support one another via excitatory links than are nonadjacent ones.

A comparison of the raw recall data in the immediate and delayed recall conditions also indicates that the semantic-separation effect resides in STM and does not prevail in episodic LTM. In immediate recall, a clear semantic effect (17%) is obtained for the recency items [last 4 serial positions; $F(1, 15) = 26.95$, $MS_e = 0.085$, $p < .001$], but not at earlier serial positions ($F < 1$), indicat-

ing an STM component (Waugh & Norman, 1965) and explaining why the main effect of separation [$F(1, 15) = 16.95$, $MS_e = 0.027$, $p < .001$] was qualified by an interaction between serial position and separation [$F(1, 15) = 2.31$, $MS_e = 0.036$, $p = .012$]. In the delayed condition, the semantic-separation effect over the same recency serial positions is much smaller (6%) and does not reach significance [$F(1, 15) = 1.7$, $MS_e = 0.015$, $p = .21$]. This pattern of results is supported by main effects of delay [$F(1, 15) = 58.17$, $MS_e = 0.024$, $p < .001$] and separation [$F(1, 15) = 12.03$, $MS_e = 0.017$, $p < .01$] and by an interaction between delay and separation at recency [$F(1, 15) = 7.36$, $MS_e = 0.068$, $p = .016$; Figure 2A]. No interaction between semantic separation and delay is found for the middle ($F < 1$) or primacy items ($F < 1$) at serial positions 4–8 and 1–3, confirming the model's prediction of a semantic-separation effect at recency.

Our immediate recall data replicate the effect reported by Glanzer (1969), with one difference. Whereas we obtained a separation effect only at recency, Glanzer found a semantic-separation effect that did not interact with serial position. Our findings also contrast with previous studies, which obtained equal or larger semantic effects in delayed recall and which did not obtain a semantic STM component after factoring out LTM contributions to retrieval (Baddeley, 1972; Craik & Levy, 1970; Glanzer & Schwartz, 1971; Levy & Baddeley, 1971). Nevertheless, in the study of Glanzer and Schwartz, larger values of STM components were reported in a semantically related versus unrelated condition for five out of six recency positions; however, this difference did not reach statistical significance. There may be a few reasons for the different results. First, in our study, the words were presented at a faster rate (of 1 sec/word),⁵ allowing less time for elaborative processing that results in LTM storage and, therefore, engaging the STM system to a larger extent.⁶ Second, in our study, the delayed and immediate conditions were randomized (and not blocked) so that subjects could not prepare different strategies in advance of the list presentation. Third, our distractor task was more difficult (requiring difficult syntactic parsing, plus semantic analysis and comparison of a stimulus sentence and a verification statement), preventing a shallow/automatic processing of the distractor material (as is indicated by the full abolishment of recency effects).

The results are challenging for theories that assume that information is retained in STM as an exclusively phonological decaying trace (Baddeley, 1972, 1986). According to such theories, one should predict equal STM components in the adjacent and the separated conditions (since they are phonologically equivalent), as well as semantic effects that are preserved in magnitude in delayed recall (since it is the encoding of information in episodic LTM that is assumed to explain the semantic effects). Neither of these predictions was confirmed by our data. A more recent computational model of the phonological loop (Burgess & Hitch, 1999) assumes that STM information is encoded in decaying weights that link item units with context and with phonological rep-

resentations. That model can explain the reduction in performance that is an effect of phonological similarity. However, since item units (in the model) do not encode associations, the model cannot explain the effects of semantic similarity (which go in the opposite direction). Moreover, adding links between semantically related item units (as in our model) would not help, since a major property of the model is the “winner take all” inhibition between item units (which, therefore, cannot support each other).

One possible way by which a reconciliation of these results with the basic statement that STM traces are exclusively phonological may be attempted is by relying on a recent modification of the redintegration theory (Stuart & Hulme, 2000). According to this, “interitem associations may create a mutually supportive network of item nodes that makes a long-term memory representation of each candidate item more accessible at the point of item retrieval” (Stuart & Hulme, 2000, p. 801). To understand this mechanism and the difference between it and the model we propose (which also assumes a network of supportive associations), it is important to make some of its assumptions explicit. What distinguishes the redintegration explanation from ours is that the associative support takes place only at retrieval and that the contributor is the word previously retrieved (this assumption is used by Stuart & Hulme, 2000, to account for the difference in recall between pure and alternating sequences in their experiment). By contrast, we assume that *in addition* to retrieval support, associates also support each other during retention. A number of considerations indicate that this is a more likely explanation. First, whereas our explanation predicts a gradual effect of proximity between associates on recall (as was reported by Glanzer, 1969 and is illustrated in Figure 2), the redintegration theory seems to predict a discontinuity, where only adjacent associates are facilitated relative to baseline.

Second, it is not clear why the advantage in retrieval for semantic associates should depend at all on degraded phonological traces. Presumably, the phonological information combines nonlinearly with the associative information to amplify the associative effect. However, it is unclear (in the absence of a computational model) why episodic contextual information (which exists for items at primacy or in delayed recall) could not obtain a similar amplification of the semantic effect. We thought to test the assumption that phonological traces mediate the associative effect directly by replicating our free recall experiment for the immediate condition under articulatory suppression (AS). Previous AS experiments have demonstrated that AS interferes with the encoding of visually presented words into a phonological code. Again, we found a proximity effect, an effect of serial position showing recency, and an interaction between serial position and proximity owing to an effect of proximity between associates at recency, but not at primacy,⁷ suggesting that the effect is not wholly mediated by phonological traces at retrieval. Third, the retrieval-only

assumption is contradicted by converging evidence from probed recognition tasks (Raser, 1972; Shulman, 1970, 1972) where the retrieval is controlled. Of particular interest is an experiment performed by Raser in which subjects are presented (in some trials) with lists of 10 words, followed by a probe that is either a synonym of a word in the list (e.g., probe=NUDE for list word BARE) or a synonym of a homophone of a word in the list (e.g., probe=NUDE for list word BEAR) and the subjects are required to decide whether the probe has the same meaning as one of the list words. The subjects are able to make such a memory discrimination at a level much higher than chance, showing also a clear pattern of recency. This, then, indicates that a semantic STM trace is being maintained, since a purely phonological one could not distinguish between the meanings of the homophones (BEAR and BARE), whereas subjects do. Nevertheless, a more complete exploration of semantic associative effects, in which retrieval processes are controlled, may be needed to fully confirm the mediation of semantic STM effects by retention processes.

Finally, a strong assumption of the redintegration theories (Schweickert, 1993) is that phonological traces in STM decay in time. We now turn to a test of serial position and load in item STM that can distinguish easily between theories based on decay versus displacement.

EXPERIMENT 2

The aim of Experiment 2 was to test the model’s prediction of an interaction between memory load and serial position (in particular, that load has its largest effect on the items at the earlier serial positions). Since we were interested in item retention (and wanted to avoid potentially confounding variables, such as output interference during recall), we used a cued-recall procedure requiring the report of only one item per trial. For cued recall, neither decay (Burgess & Hitch, 1999; Hulme et al., 1997; Schweickert, 1993) nor RI (Lewandowsky & Murdock, 1989; Nairne, 1990) models of STM predict an interaction between memory load and serial position, because the number of items preceding a word at a particular position (e.g., -4) affects neither the length of the time interval nor the number of items between that word and its cue.⁸ Moreover, since in the model that we match to the data, units correspond to conceptual (lexical/semantic) representations, and in order to minimize reliance on phonological strategies (which could rely on a different memory system than the one addressed), we used semantic category cues. A secondary aim of the experiment was to test whether word length (WL, a variable often associated with the temporal decay hypothesis) affects performance in this task. In a previous study, Avons, Wright, and Pammer (1994) demonstrated that, indeed, WL effects are diminished in cued-recall tests (as compared with serial order recall tests). In order to minimize the contribution of episodic LTM retrieval, the word material was presented with replacement from a

fixed word pool. Repetition of words across trials generates proactive inhibition, which has been shown to affect retrieval from LTM more than that from STM (Craik & Birtwistle, 1971; Halford, Maybery, & Bain, 1988).

Method

Subjects and Design. We recruited 24 undergraduate students with the same subject characteristics as those in Experiment 1. The subjects were randomly assigned to one of two groups. One group of subjects ($n = 12$) performed cued recall with sequences of four words, and the second group ($n = 12$) with sequences of six words. Each group received an equal number of trials with short words and with long words.

Materials. The word material consisted of six semantic categories, with four words in each category. The short words were one syllable long (e.g., *grape*, *pear*, *lime*, and *peach* for *fruit*), and the long words were two, three, and four syllables long (e.g., *banana*, *orange*, *apple*, *strawberry*). Short and long words were matched for frequency, concreteness, and imaginability. A pretest of these materials, using serial recall of four-word sequences, showed a highly significant effect of word length (35% advantage for short words, $p < .001$). From these materials, we created a list of 20 short- and 20 long-word trials for each subject group. Word length was blocked, and the order of blocks was counterbalanced. Words were selected with replacement, but the same word was not presented in consecutive trials or in the same serial positions. Category cues probed serial positions -4 , -3 , -2 , -1 (from the end of the list) in four-word lists and serial positions -5 , -4 , -2 , and -1 , in six-word lists. There were 5 trials per serial position. The order of the serial positions tested was randomized.

Procedure. Sequences of four or six words (each one from a different semantic category) were shown, using a computer-controlled visual presentation, with a presentation rate of 1 word/sec. One second after the last item in the sequence, a category cue appeared, and the subjects were required to say the corresponding item in that trial. An example of a four-word trial sequence with short words is *fan*, *car*, *lime*, *rat*, *FRUIT*, and a corresponding example with long words is *radio*, *locomotive*, *strawberry*, *monkey*, *FRUIT*. A category cue appeared 1 sec after the last item in the sequence, and the subjects were required to say the corresponding item in that trial.

Analysis. The percentage correctly recalled was analyzed in a mixed factor by-subjects analysis of variance crossing the factors memory load (four or six words), WL (short or long), and serial position (-4 , -2 , -1). Memory load was between-subjects, whereas both WL and serial position were within-subjects. Serial positions -4 , -2 , and -1 constituted identical probed positions (from the end of the list) in four- and six-word sequences.

Results and Discussion

All reported results were collapsed across word-length, which was not involved in any effects ($F < 1$). The results were as predicted by the model (Figure 4A). As for the model, the recency functions show that increasing the load from four to six words affected the early serial positions more than it did the late ones (Figure 4B, lines labeled L4 and L6), as is reflected by main effects of memory load [$F(1,22) = 22.04$, $MS_e = 0.021$, $p < .001$] and serial position [$F(2,44) = 34.47$, $MS_e = 0.018$, $p < .001$] and an interaction between memory load and serial po-

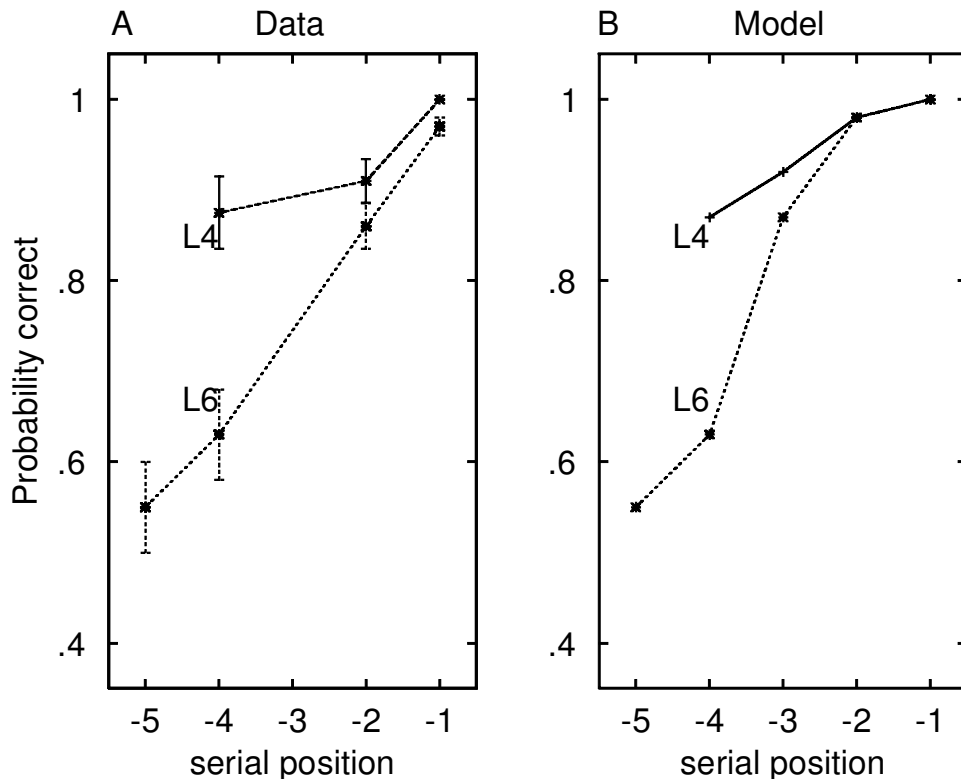


Figure 4. (A) Serial position curves in category cued recall for lists of four and six items (L4 and L6). Error bars show standard errors. (B) Model predictions with the same parameters as for Figure 3C (see the Appendix). An independent contribution of .47, corresponding to correct responses owing to guessing and episodic long-term memory recall, is assumed in all the conditions.

sition. [$F(2,44) = 8.74$, $MS_e = 0.018$, $p = .001$]. The statistical test of the interaction was restricted to identical probed serial positions in both memory load conditions (i.e., -1 , -2 , -4 from the end of the list). The interaction between memory load and serial position remained highly significant when the easiest (i.e., final) serial position was removed from the analysis [$F(1,22) = 8.13$, $MS_e = 0.027$, $p = .01$].

The cued-recall results are also consistent with other models that postulate displacement type forgetting in a limited-capacity store such as SAM (Raaijmakers & Shiffrin, 1981; but see also Atkinson & Shiffrin, 1968; Baddeley, 2000; Cowan, 1999; Waugh & Norman, 1965). The interaction between load and serial position is less easily explained in STM models based on temporal decay (Burgess & Hitch, 1999; Hulme et al., 1997; N. Martin et al., 1996; Schweickert, 1993) or on RI (Nairne, 1990), because the load manipulation did not affect the duration of or the number of words in the interval between the tested word and the cue.⁹

A mechanism based on proactive interference (PI) could predict a load effect but, by itself, not be sufficient to explain the main effect of recency. Nevertheless, it is possible that a more complex model combining PI and RI could generate a load effect, a recency effect, and the steeper recency effect with increased load, as obtained in our experiment. However, there are several considerations that argue against interference models of capacity limitations in item STM (as opposed to displacement models, as used here and in the SAM model; Raaijmakers & Shiffrin, 1981). First, in Experiment 1, we obtained improved recall for semantically similar adjacent words, whereas interference models (both PI and RI) predict that similarity of representations diminishes performance. A related finding, the fact that the elimination of recency by distractor activity during delay intervals depends little on the similarity between the retention material and the distractors (Brown, 1958; Glanzer, Gianutsos, & Dubin, 1969), was used by Glanzer and colleagues to support displacement, rather than interference, models for recency in free recall. Second, with regards to PI, the manipulation is typically thought to generate interference between items presented in different trials and to accumulate across trials (Wickens, Moody, & Dow, 1981). Clearly, only a within-trial PI could be relevant for load \times serial position interactions. PI caused by interfering items occurring in the same trial as the target has never been reported for different-category items such as we presented in our experiment. Moreover, to obtain PI with same-category items (Tehan & Humphreys, 1998), an additional set of conditions, not satisfied in our experiment, need to be met.¹⁰ Third, in immediate recall, short, but not long, lists of items are immune from the build-up of PI across lists of same-category items (Cowan, 2001; Halford et al., 1988; Wickens et al., 1981). When long lists are used, recency, but not prerecency, items are immune from PI (Craik & Birtwistle, 1971; Goshen-Gottstein, Ashkenazi, & Usher, 2001). This suggests that items in capacity-limited STM are protected from PI,

whereas items in LTM are not (Cowan, 2001; see also Goshen-Gottstein et al., 2001, for a dissociation between effects of PI on short-term and long-term recency).

Although some features in our cued-recall experiments are similar to results reported in order cued recall (Avons et al., 1994; Murdock, 1968), other features differ. A large WL effect was found in serial recall in the pretest in our experiment and also by Avons et al. However, the WL effect was completely abolished in cued recall in our experiment, whereas it was small but still present in cued recall in the study by Avons et al. Furthermore, the increase in the slope of the recency effect with increased list length was obtained in both our experiment and Murdock's. However, a primacy effect was present only in the latter study. These differences are expected if category cued recall engages phonological rehearsal strategies (contributing to WL and primacy effects) to a much lesser extent than does order cued recall, suggesting that our task may be more suitable for studying capacity limitations in item STM.

GENERAL DISCUSSION

The results of both Experiments 1 and 2 confirmed the predictions of the model (Usher & Cohen, 1999). In Experiment 1, there was a larger STM component at recency items for adjacent word associates than for separated associates. In Experiment 2, increasing the retention load had a larger effect on early than on late serial positions. The model we presented interprets these results in terms of a capacity-limited STM system in which a small number of conceptual representations can be simultaneously active. This type of item-limited capacity (sensitive to semantic structure) was one of the best-known properties of earlier STM theory (Broadbent, 1975; G. A. Miller, 1956) but was somehow neglected later on. Recently, this approach has been brought back into the spotlight (Baddeley, 2000; Cowan, 2001; Potter, 1993). For example, Cowan (2001) provides a review of literature demonstrating that when additional strategies (such as rehearsal) are prevented, a capacity of about 3–4 items is a robust finding for many studies of STM, consistent with our results (see Figure 2). The only alternative explanation of the semantic proximity effect is the revised redintegration theory (Stuart & Hulme, 2000). We have discussed a number of objections to this explanation. However, since the redintegration theory was never used within a computational model to account for the semantic effect, it is difficult to fully rule it out. Nevertheless, the results of our second experiment conflict with one of the basic assumptions of the redintegration theory: the temporal decay of phonological traces.

The neurocomputational model we have presented is consistent with recent neuropsychological evidence, which suggests the existence of separate phonological and semantic components in STM (R. C. Martin & Romani, 1994; R. C. Martin et al., 1994). Our model suggests a mechanism for how item information is maintained in the capacity-limited part of the semantic STM component.

Preservation of this component explains why patients with a phonological STM deficit, such as P.V. (Baddeley, Papagno, & Vallar, 1988) and E.A. (R. C. Martin & Romani, 1994), have a problem with the formation of new phonological, but not semantic, long-term memories (Baddeley et al., 1988; Romani & Martin, 1999). This supports the proposal of R. C. Martin and colleagues that these dissociations, originally interpreted as rejecting the dual-memory model (Atkinson & Shiffrin, 1968) for lack of transfer from STM to LTM (Warrington & Shallice, 1969), may be better interpreted in terms of a dissociation between phonological and semantic processing.¹¹

Our model is also consistent with traditional memory models, such as SAM (Raaijmakers & Shiffrin, 1981), which postulate an item-limited STM that maintains activated conceptual LTM representations. This needs to be distinguished from episodic LTM representations that require specific encoding of links between the retention events and the relevant context or list tags and that are expected to rely on structural (weight) changes that should be unaffected by short periods of distractor activity. Interestingly, the activation-based STM provides a very simple type of episodic information that may be enough to distinguish items from the current trial (or list) from items in previous trials, but not between items that were encoded with regards to arbitrary lists. Especially when the same items occur repeatedly in highly similar contexts, as in our cued-recall experiment, activation-based STM may provide a more reliable system for maintaining a focus on recent and relevant information than does episodic LTM, where similarity with less recent items may produce interference. The system we describe may thus achieve the function of an episodic buffer, as was recently proposed by Baddeley (2000).

The model also shares many properties with a series of activation-based models that have been used in the context of language comprehension (Haarmann, Just, & Carpenter, 1997; Just & Carpenter, 1992), in mental arithmetic (Anderson, Reder, & Lebiere, 1996), and in contextual disambiguation and control (Braver, Barch, & Cohen, 1999; Cohen, Braver, & O'Reilly, 1996; Cohen & Servan-Schreiber, 1992). For example, an important characteristic of the models developed by Cohen and collaborators is that context information needs to be maintained in an active state (and across intervening items) in order to be used in the control (or biasing) of information processed later on. It is important to note that in order for the context to facilitate information processing, it has to be retained in a central type of representation (conceptual, and not phonological or orthographic).

The model we presented has a number of limitations. First, it targets only a very specific STM component and does not address various interactions between semantic and phonological representations in memory (Davelaar & Usher, 2001). Second, it does not address the relation between STM retention and learning in episodic memory. Future work is needed to address these important characteristics of STM and to further

test the model's predictions regarding the brain mechanisms mediating immediate memory performance.

REFERENCES

- ANDERSON, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning & Verbal Behavior*, **22**, 261-295.
- ANDERSON, J. R., REDER, L. M., & LEBIERE, C. (1996). Working memory: Activation limitations on retrieval. *Cognitive Psychology*, **30**, 221-256.
- ATKINSON R. C., & SHIFFRIN R. M. (1968). Human memory: A proposed system and its control processes. In W. K. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2., pp. 89-105), New York: Academic Press.
- AVONS, S. E., WRIGHT, K. L., & PAMMER, K. (1994). The word-length effect in probed and serial recall. *Quarterly Journal of Experimental Psychology*, **47A**, 207-231.
- BADDELEY, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. *Quarterly Journal of Experimental Psychology*, **18**, 312-365.
- BADDELEY A. D. (1968). Prior recall of newly learned words and the recency effect in free recall. *Canadian Journal of Psychology*, **22**, 157-163.
- BADDELEY A. D. (1970). Estimating the short term component in free recall. *British Journal of Psychology*, **61**, 13-15.
- BADDELEY, A. D. (1972). Retrieval rules and semantic coding in short-term memory. *Psychological Bulletin*, **78**, 379-385.
- BADDELEY, A. D. (1986). *Working memory*. Oxford: Oxford University Press, Clarendon Press.
- BADDELEY, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, **4**, 417-423.
- BADDELEY, A. D., & COB, J. R. (1970). Simultaneous acoustic and semantic coding in STM. *Nature*, **227**, 288-289.
- BADDELEY, A. D., PAPAGNO, C., & VALLAR, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory & Language*, **27**, 586-595.
- BRAVER, T. S., BARCH, D. M., & COHEN, J. D. (1999). *Mechanisms of cognitive control: Active memory, inhibition, and the prefrontal cortex* (Tech. Rep. No. PDP.CNS.99.1). Pittsburgh: Carnegie Mellon University, PDP Research Group.
- BROADBENT, D. E. (1975). The magical number after fifteen years. In R. A. Kennedy & A. Wilkes (Eds.), *Studies in long-term memory* (pp. 3-18). New York: Wiley.
- BROWN, J. (1958). Some tests of the decay theory of immediate memory. *Journal of Experimental Psychology*, **10**, 2-21.
- BURGESS, N., & HITCH, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, **106**, 551-581.
- COHEN, J. D., BRAVER, T. S., & O'REILLY, R. C. (1996). A computational approach to prefrontal cortex, cognitive control and schizophrenia: Recent developments and current challenges. *Philosophical Transactions of the Royal Society of London: Series B*, **351**, 1515-1527.
- COHEN, J. D., & SERVAN-SCHREIBER, D. (1992). Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review*, **99**, 45-77.
- CONRAD, R. (1965). Order error in immediate recall of sequences. *Journal of Verbal Learning & Verbal Behavior*, **4**, 161-169.
- COWAN, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62-101). Cambridge: Cambridge University Press.
- COWAN, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral & Brain Sciences*, **24**, 87-185.
- CRAIK, F. I. M., & BIRTWISTLE, J. (1971). Proactive inhibition in free recall. *Journal of Experimental Psychology*, **91**, 120-123.
- CRAIK, F. I. M., & LEVY, B. A. (1970). Semantic and acoustic information in primary memory. *Journal of Experimental Psychology*, **86**, 77-82.
- CROSSON, B., RAO, S. M., WOODLEY, S. J., ROSEN, A. C., BOBHOLZ, J. A., MAYER, A., CUNNINGHAM, J. M., HAMMEKE, T. A., FULLER, S. A.,

- BINDER, J. R., COX, R. W., & STEIN, E. A. (1999). Mapping of semantic, phonological, and orthographic verbal working memory in normal adults with functional magnetic resonance imaging. *Neuropsychology*, **13**, 171-87.
- CROWDER, R. G. (1979). Similarity and serial order in memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 13, pp. 319-353). New York: Academic Press.
- CROWDER, R. G. (1993). Short-term memory: Where do we stand? *Memory & Cognition*, **21**, 142-145.
- DAVELAAR, E. J., & USHER, M. (2001). *Exploring a computational account of item short-term memory*. Manuscript in preparation.
- FUNAHASHI, S., CHAFFEE, M. V., & GOLDMAN-RAKIC, P. S. (1993). Prefrontal neuronal activity in rhesus monkeys performing a delayed anti-saccade task. *Nature*, **365**, 753-756.
- GLANZER, M. (1969). Distance between related words in free recall: Trace of the STS. *Journal of Verbal Learning & Verbal Behavior*, **8**, 105-111.
- GLANZER, M., & CUNITZ, A. R. (1967). Two storage mechanisms in free recall. *Journal of Verbal Learning & Verbal Behavior*, **6**, 351-360.
- GLANZER, M., GIANUTSOS, R., & DUBIN, S. (1969). The removal of items from short-term storage. *Journal of Verbal Learning & Verbal Behavior*, **8**, 435-447.
- 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.
- GOSHEN-GOTTSTEIN, Y., ASHKENAZI, A., & USHER, M. (2001). *Proactive interference attenuates recency in the continuous distractor task but not in free recall*. Manuscript submitted for publication.
- HAARMANN, H. J., JUST, M. A., & CARPENTER, P. A. (1997). Aphasic sentence comprehension as a resource deficit: A computational approach. *Brain & Language*, **59**, 76-120.
- HALFORD, G. S., MAYBERY, M. T., & BAIN, J. D. (1988). Set-size effects in primary memory: An age-related capacity limitation? *Memory & Cognition*, **16**, 480-487.
- 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.
- HULME, C., ROODENRYS, S., SCHWEICKERT, R., BROWN, A., & GORDON, D. A. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 1217-1232.
- JUST, M. A., & CARPENTER, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, **99**, 122-149.
- LEVY, B. A., & BADDELEY, A. (1971). Recall of semantic clusters in primary memory. *Quarterly Journal of Experimental Psychology*, **23**, 8-13.
- LEWANDOWSKY, S., & MURDOCK, B. B. (1989). Memory for serial order. *Psychological Review*, **96**, 25-57.
- MARTIN, N., SAFFRAN, E. M., & DELL, G. S. (1996). Recovery in deep dysphasia: Evidence for a relation between auditory-verbal STM capacity and lexical errors in repetition. *Brain & Language*, **52**, 83-113.
- MARTIN, R. C., & ROMANI, C. (1994). Verbal working memory and sentence comprehension: A multiple-components view. *Neuropsychology*, **8**, 506-523.
- MARTIN, R. C., SHELTON, J. R., & YAFEE, L. S. (1994). Language processing and working memory: Evidence for separate phonological and semantic capacities. *Journal of Memory & Language*, **33**, 83-111.
- MILLER, E. K., ERICKSON, C. A., & DESIMONE, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience*, **16**, 5154-5167.
- MILLER, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, **63**, 81-97.
- MURDOCK, B. B. (1968). Serial order effect in short-term memory. *Journal of Experimental Psychology*, **76**, 1-15.
- MURDOCK, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, **89**, 609-626.
- NAIRNE, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, **18**, 251-269.
- PALERMO, D., & JENKINS, J. (1964). *Word association norms: Grade school through college*. Minneapolis: University of Minnesota Press.
- PETERSON, L., & PETERSON, M. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, **58**, 193-198.
- POIRIER, M., & SAINT-AUBIN, J. (1995). Memory for related and unrelated words: Further evidence on the influence of semantic factors in immediate serial recall. *Quarterly Journal of Experimental Psychology*, **48A**, 384-404.
- POTTER, M. C. (1993). Very short-term conceptual memory. *Memory & Cognition*, **21**, 156-161.
- RAAJMAKERS, J. G., & SHIFFRIN, R. M. (1981). Search of associative memory. *Psychological Review*, **88**, 93-134.
- RASER, G. (1972). Recoding of semantic and acoustic information in short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **11**, 692-697.
- ROMANI, C., & MARTIN, R. (1999). A deficit in the short-term retention of lexical-semantic information: Forgetting words but remembering a story. *Journal of Experimental Psychology: General*, **128**, 56-77.
- SAINT-AUBIN, J., & POIRIER, M. (1999). Semantic similarity and immediate serial recall: Is there a detrimental effect on order information? *Quarterly Journal of Experimental Psychology*, **52A**, 367-394.
- SCHWEICKERT, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory & Cognition*, **21**, 168-175.
- SHULMAN, H. G. (1970). Encoding and retention of semantic and phonemic information in short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **9**, 499-508.
- SHULMAN, H. G. (1972). Semantic confusion errors in short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **11**, 221-227.
- STUART, G., & HULME, C. (2000). The effects of word co-occurrence on short-term memory: Associative links in long-term memory affect short-term memory performance. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 796-802.
- TEHAN, G., & HUMPHREYS, M. S. (1998). Creating proactive interference in immediate recall: Building a DOG from a DART, a MOP, and a FIG. *Memory & Cognition*, **26**, 477-489.
- TSODYKS, M., PAWELZIK, K., & MARKRAM, H. (1998). Neural networks with dynamic synapses. *Neural Computation*, **10**, 821-835.
- USHER, M., & COHEN, J. D. (1999). Short-term memory and selection processes in a frontal-lobe model. In D. Heinke, G. W. Humphries, & A. Olsen (Eds.), *Connectionist models in cognitive neuroscience* (pp. 78-91). London: Springer-Verlag.
- WARRINGTON, E., & SHALLICE, T. (1969). The selective impairment of auditory verbal short-term memory. *Brain*, **92**, 885-896.
- WAUGH, N. C., & NORMAN, D. A. (1965). Primary memory. *Psychological Review*, **72**, 89-104.
- WICKENS, D. D., MOODY, M. J., & DOW, R. (1981). The nature and timing of the retrieval process and of interference effects. *Journal of Experimental Psychology: General*, **110**, 1-20.

NOTES

1. Since semantic manipulations do not affect order information (Poirier & Saint-Aubin, 1995; Saint-Aubin & Poirier, 1999), we focus here only on the STM for item information, using free- and cued-recall tasks.

2. The model contains five parameters, of which only one, the noise level, was varied in order to explain data from the two experiments presented here. All the other parameters were fixed to the same values as those in previous work (Usher & Cohen, 1999).

3. Note that assuming an independence between these two processes does not require that the two systems did not interact prior to memory retrieval. Similar assumptions are used in the multinomial model for STM (Hulme et al., 1997; Schweickert, 1993).

4. The assumption that items can be maintained in STM independently of whether they can be encoded and retrieved from LTM after the delay task was supported by data in Baddeley (1968). However, even if this assumption fails, the corresponding bias obtained in the STM estimation is not likely to affect the results. One alternative and extreme assumption could assume anticorrelation (an item is either in STM or in LTM, but not in both). According to this, the STM measure is simply the subtraction between the immediate and the delay condition (i.e., our measure, but without the denominator). Estimation based on this difference measure still shows an effect of semantic proximity at recency

[11% difference; $F(1,15) = 7.36$, $MS_e = 0.055$, $p = .016$]. The opposite assumption is positive correlation (items maintained in STM at the end of trial are more likely to be also encoded in and retrieved from LTM). In this case, the independence measure would underestimate the STM component, so that the true STM component is even larger than what we report. If we reported a null effect of the STM effect on the STM component (as in articles cited), this assumption would need more careful examination. Since our effect was significant, a positive correlation can only further amplify it.

5. In Glanzer and Schwartz (1971), a presentation rate of 2.3 sec/word was used, whereas in Glanzer (1969), where two presentation rates, of 1 and 3 sec/word, were used, the interaction between separation and serial position (which was statistically insignificant) was not reported for the 1-sec presentation rate separately.

6. Also, the longer response interval in Glanzer's (1969) study (90, instead of 45 sec) may have encouraged his subjects to search LTM more than did our subjects, possibly increasing the contribution of LTM to the semantic-separation effect and explaining why it was equally present at all serial positions.

7. Twelve subjects were tested. The design, materials, and procedure were the same as those for the immediate recall condition in Experiment 1, except that the subjects repeated the string "bla" out loud at a rate of 3/sec during the presentation of the word sequence. The overall, repeated measurements analysis of variance revealed a main effect of proximity [$F(1,11) = 6.39$, $MS_e = 6.6$, $p = .028$], an effect of serial position [$F(11,121) = 5.49$, $MS_e = 4.7$, $p < .001$], showing recency, and an interaction between serial position and proximity [$F(11,121) = 2.99$, $MS_e = 8.9$, $p = .001$]. Post hoc comparisons revealed an effect of proximity between associates at recency [$F(1,11) = 34.91$, $MS_e = 3.5$, $p < .001$], but not at primacy ($F < 1$).

8. The situation is somehow more complex for the Lewandowsky and Murdock (1989) model, where in addition to retroactive interference, other factors, such as an explicit dependency of item encoding strength on serial position, are assumed. This involves a more complex set of assumptions than the model we presented, in which serial position dependency is an emergent property of the model. Nevertheless, a careful inspection of the data in Lewandowsky and Murdock (Figures 3 and 13) reveals that the model could not fully account for load \times serial position interactions.

9. A smaller interaction can be obtained in the decay model of Burgess and Hitch (1999); for example, increasing the load from four to six items results in a decrease in correct recognition from .91 to .83, at the same serial position of four items from probe (N. Burgess, personal communication, December 15, 2000).

10. Using category cued recall, Tehan and Humphreys (1998) found reliable within-trial PI for same-category items only when the interfering item was a much more typical member of the category than was the target and was phonologically related to the target or other recent item(s).

11. Romani and Martin (1999) stressed that P.V.'s learning of new words was, in fact, impaired and suggested that P.V.'s preserved ability to learn associations between existing words could have been mediated by an intact semantic STM, such as the one proposed in our model. Further support for their proposal came from a double dissociation observed in two patients, E.A. (similar to P.V.) and A.B., who showed selective phonological and semantic impairments, respectively, in both immediate recall and long-term learning tasks (Romani & Martin, 1999). The association between a semantic STM deficit and a semantic LTM learning deficit (in particular, a failure to learn associations between existing words) suggests an important cognitive function for semantic STM—namely, the maintenance of word meanings for semantic binding.

APPENDIX

Model Equations and Parameters

In the model, each unit connects to itself by a recurrent excitatory connection of strength α_1 , it sends excitation of strength α_2 to its associates, and it inhibits other units by weights of strength β . The dynamics behavior is determined by the following iterative stochastic equations (which correspond to the numeric solution of a differential equation), for each of the units (denoted by x_i), which are updated in parallel (y_i denotes the semantic associate of x_i):

$$x_i(t+1) = \lambda x_i + (1 - \lambda) \left[\alpha_1 F(x_i) + \alpha_2 F(y_i) - \beta \sum_{j \neq i}^N F(x_j) + I_i + noise \right], \quad (A1)$$

where λ is the decay time constant (which is taken as .99, corresponding to Euler integration steps of .01), $F(x) = \frac{1}{1+e^{-x}}$ is the activation function used,^{A1} I_i is the sensory input to each representation (which is set to zero after the unit is activated), and *noise* is a Gaussian variable with a mean of 0 and $SD = \sigma$. The only parameter that was varied between the simulations of the two experiments is the noise level σ , which was 3 and 1.5 in the free- and cued-recall simulations, respectively (the larger value used in free recall is thought to reflect the increased complexity of this process, especially the output interference, which does not exist in cued recall). All other parameters were fixed to the same values as those used in the previously reported simulations (Usher & Cohen, 1999): $\alpha_1 = 2$, $\alpha_2 = .10$, $\beta = .15$, $I_i = .33$. Each input in the sequence is presented for 550 iterations before being switched off and replaced by the next input. Each data point is obtained by averaging over 400 simulation trials.

NOTE

A1. This type of activation function has been recently proposed to provide a better description of synaptic transmission in the computational neuroscience literature (Tsodyks, Pawelzik, & Markram, 1998).