COGNITIVE PSYCHOLOGY **36,** 73–137 (1998)
ARTICLE NO. CG980685

# Short-Term Memory for Serial Order: The Start-End Model

Richard N. A. Henson

*Institute of Cognitive Neuroscience, University College London, London, United Kingdom, r.henson@ucl.ac.uk*

Three solutions to the problem of serial order can be identified: chaining, ordinal and positional theories. Error patterns in serial recall from short-term memory fail to support chaining theories, yet provide unequivocal evidence for positional theories. In a new model of short-term memory, the Start-End Model (SEM), the positions of items in a sequence are coded relative to the start and end of that sequence. Simulations confirm SEM's ability to capture the main phenomena in serial recall, such as the effects of primacy, recency, list length, grouping, modality, redundant suffices, proactive interference, retention interval, and phonological similarity. Moreover, SEM is the first model to capture the complete pattern of errors, including transpositions, repetitions, omissions, intrusions, confusions, and, in particular, positional errors between groups and between trials. Unlike other positional models however, SEM predicts that positional errors will maintain relative rather than absolute position, in agreement with recent experiments (Henson, 1977).

This article is concerned with the problem of serial order in short-term memory. More specifically, it addresses the question of how we store and retrieve a novel sequence of items in the correct order. This is the task we face when, for example, attempting to dial a telephone number that we have only recently heard. This task of verbatim or serial recall has a long history of laboratory study, underlying much of the research on the forgetting (e.g., Brown, 1958; Peterson & Peterson, 1959), organization (e.g., Miller, 1956/1994) and structure (e.g., Baddeley, 1986) associated with short-term memory. Surprisingly however, the basic psychological processes underlying the task remain little understood.

The article begins with a review of three theories of how we retain order

This paper is based on part of the author's doctoral dissertation at the University of Cambridge, England, which was supported by a Research Studentship from the Medical Research Council. The author thanks Alan Baddeley, Gordon Brown, Mike Page, and Dennis Norris for reading preliminary drafts. The research was conducted at the MRC Applied Psychology Unit, Cambridge. The author has since moved to University College London.

Address correspondence and reprint requests to Rik Henson, Institute of Cognitive Neuroscience, University College London, Gower Street, London WC1E 6BT, UK.

0010-0285/98 $25.00

in memory and their predictions for the errors people make when they misrecall a sequence. These errors support theories that assume some degree of positional information associated with the items in a sequence. The nature of this positional information is then formalized in a computational model that reproduces the main empirical phenomena associated with serial recall. Finally, it is argued that this model, the Start-End Model (SEM), is not only a considerable improvement on previous models, but is also unique in predicting the pattern of errors found between sequences of different lengths (Henson, 1997).

## THREE THEORIES OF SERIAL ORDER

### *Chaining Theory*

This theory assumes that order is stored by the formation or strengthening of associations between successive elements of a sequence. The order is retrieved by stepping along these associations in a process called *chaining,* where each element cues the recall of its successor.

Chaining is probably the oldest approach to serial order (Ebbinghaus, 1964) and certainly the most intuitive. It is a simple extension of stimulus-response theory, where each response can become the stimulus for the next (Lashley, 1951). In its various guises, it has remained popular in several different models (e.g., Elman, 1990; Jordan, 1986; Lewandowsky & Murdock, 1989; Murdock, 1995; Wickelgren, 1965). However, chaining theory also faces several problems, as discussed below.

The simplest chaining models assume only pairwise associations between adjacent elements of a sequence (e.g., Wickelgren, 1965) and cues that consist entirely of the preceding response (upper illustration in Fig. 1A). There are several immediate objections to such models. For example, how do they handle sequences with a repeated element, in which two different elements will share the same cue? Or how do they allow recovery from error, because once an error has been made, the cue for subsequent responses will be incorrect? This should lead to a cascade of further errors (‘‘a chain is only as strong as its weakest link’’).

These questions are less problematic for models that assume remote associations as well as adjacent ones (e.g., Ebbinghaus, 1964; Slamecka, 1985). In these *compound chaining* models (lower illustration in Fig. 1A), the cue consists of a number of preceding elements, an approach that is popular in recurrent neural networks (e.g., Elman, 1990; Jordan, 1986). These compound cues allow disambiguation of repeated elements, by virtue of the additional context of the elements preceding the repeated elements. They also mean that a single error is less devastating; the additional context allowing recovery from that error.

There are other possible solutions. With respect to the problem of repeated elements, one can appeal to the type/token distinction, so that two occur-

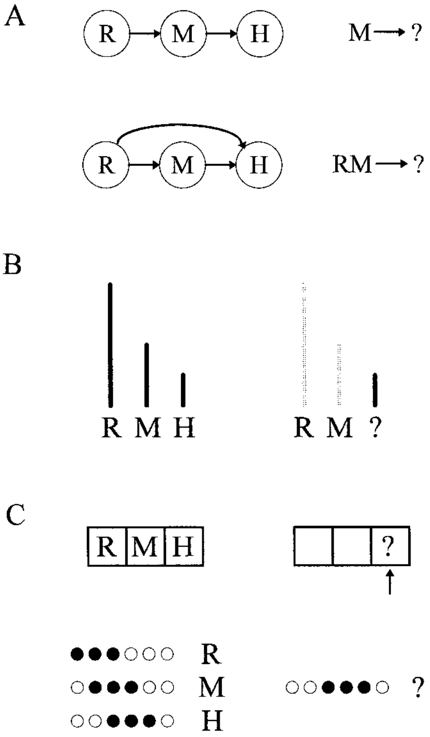


**FIG. 1** Storage and retrieval of a sequence *RMH* according to (A) simple and compound chaining models, (B) the ordinal model of Page and Norris (1996), and (C) the positional models of Conrad (1965) and Burgess and Hitch (1992).

rences of the same type have nonidentical token representations, as with the allophones (or ‘‘wickelphones’’) of Wickelgren (1969). With respect to the problem of errors in recall, the TODAM model (Murdock, 1983) assumes only pairwise associations, but copes with errors by only cueing with the previous response if it is correct. Otherwise, a cue approximating the correct one is used (Lewandowsky & Murdock, 1989).[1]

There are many arguments against the sufficiency of chaining theory as a

---

[1] Though this may be appropriate when feedback of the correct response is provided, it is inappropriate for most situations (such as serial recall), where one does not always know whether each response is correct. Nonetheless, the use of an approximate cue illustrates an important distinction in chaining theory: whether the cue consists of the preceding elements recalled, which may be erroneous, or whether the cue consists of the preceding elements stored, irrespective of whether or not they are recalled correctly. This distinguishes closed-loop chaining models, where responses are fed back as cues, from open-loop models, where there is no feedback (Henson, 1996). Open-loop models do not necessarily suffer from the problem of errors in recall.

general account of sequential behavior (e.g., Johnson, 1972; Lashley, 1951). Though not all these arguments apply to short-term memory, important questions remain. For example, there is the question of what cues the first element in a sequence, in order to ‘‘kickstart’’ the chaining process. Many chaining models appeal to an additional contextual cue (Murdock, 1995) or plan unit (Jordan, 1986). Thus the representation of serial order clearly involves more than a set of interitem associations. Later, it will argued that interitem associations are not even necessary, in the sense that there is no conclusive evidence for chaining in short-term memory.

### *Ordinal Theory*

This theory assumes that elements can be represented along a single dimension, such that order is defined by the relative (rather than absolute) values on that dimension. For example, Grossberg (1978) assumed that order is stored in a primacy gradient of strengths, such that each element is stronger than its successor. The order of elements can be retrieved by an iterative process of selecting the strongest element, and then suppressing it so that it is not selected again (Fig. 1B; suppression indicated by the lighter lines). This idea has been incorporated into the Primacy Model of short-term memory (Page & Norris, in press), where the strengths might represent the level of activation of item representations in memory.

The Perturbation Model was originally an ordinal theory (Estes, 1972), where order was inherent in the cyclic reactivation of elements. Perturbations in the timings of reactivations led to erroneous reorderings of the elements, much like changes in the order of runners racing round a track. Yet another ordinal model is that of Shiffrin and Cook (1978). Their model assumes that each element is associated with a ‘‘node,’’ but only the nodes are associated with one another (unlike chaining models, where it is the elements themselves that are associated with each other). By moving inward from nodes at the start and end of the sequence, the associations between nodes allow the order of items to be reconstructed.

Ordinal models like the Primacy Model require token representations in order to handle repeated elements: The order of repeated elements could not be represented over type representations with a single strength. As regards errors in recall, ordinal models imply that errors will cooccur, in the sense that one error will cause another (because order is defined relationally). For example, if an element becomes stronger than its predecessor in the Primacy Model (owing to random noise), then the two elements will transpose. This is an attractive property, because such paired transpositions of adjacent elements are common in people too.

Ordinal models do not require feedback of responses, and a process like suppression can operate independently of errors occurring at later stages of output. Indeed, the process of selection and suppression in the Primacy

Model is simpler than the reinstatement of positional codes required in positional theory (see below). Ordinal theory therefore escapes some of the criticisms of chaining and positional theories. Nonetheless, it will be argued later that ordinal theory is insufficient as an account of people's short-term memory for serial order.

### *Positional Theory*

This theory assumes order is stored by associating each element with its position in the sequence. The order is retrieved by using each position to cue its associated element. In other words, rather than using the item-item associations of chaining theory, positional theory uses position-item associations.

The simplest example of a positional theory is Conrad's ''box'' model (Conrad, 1965). Conrad assumed that people possess a number of boxes in short-term memory, in which elements of a sequence can be stored (upper illustration in Fig. 1C). The order of elements can be retrieved by stepping through the boxes according to a predetermined routine. This model does not have a problem with repeated elements, because they are stored in separate boxes, nor with recovery from errors, because the retrieval mechanism can continue to the next box irrespective of whether the contents of the previous box were retrieved correctly. This is of course the method by which conventional Von Neumann computers store and retrieve order, through routines accessing separate addresses in memory.

As a psychological model however, Conrad's model is inadequate. First, how many boxes do we possess: five, six, seven, or more? If a new box were created for each element in a sequence, there would be no limit to the length of sequences people could retain in short-term memory, which is clearly not the case. Secondly, the model provides no immediate rationale for the errors people make when they misremember sequences: People are more likely to confuse elements close together in a sequence than elements far apart (e.g., Estes, 1972). There is no reason for this with the perfect coding and retrieval of positions assumed by Conrad (1965).

One explanation for the above errors is that positional codes become confused over time. For example, the Perturbation Model (Lee & Estes, 1977; 1981) assumes the positions of elements are initially coded perfectly, but get perturbed during storage such that nearby elements exchange. Another way to explain such errors is that positional codes are not perfect, but overlap, in that the code for one position is similar to the codes for nearby positions (lower illustration of Fig. 1C). This is the approach taken by Burgess and Hitch (1992). In their model, the circles in Fig. 1C represent nodes in a connectionist network. The filled nodes are active nodes; the unfilled nodes are inactive nodes. The ''window'' of active nodes moves from left to right for each position in a sequence, and is associated with other nodes (not shown) representing each element. However, because there is some overlap in the set of active nodes for nearby positions, elements at these positions can be confused during retrieval.

Positional theory can be extended to a hierarchy of positional codes (e.g., Lee & Estes, 1981). For example, an item can be coded for both its position in a sequence and the position of that sequence in a sequence of sequences. As a general solution to the problem of serial order however, the status of positional theory remains unclear. There is a sense in which the problem is not solved, but circumvented. This sense concerns the question of how the order of the positional codes themselves is stored and retrieved from memory. This question cannot be answered without specifying the nature of the positional codes. One suggestion is that the codes are derived from temporal oscillators in the brain (Brown, Preece & Hulme, in press; Burgess & Hitch, 1996). Elements can be associated with successive states of the oscillators, and these states reconstructed simply by resetting the oscillators. In other words, the oscillators represent a biological clock, which can be rewound in order to retrieve a sequence from memory. However, though there is good evidence for positional information in short-term memory, the main purpose of the present article is to illustrate an alternative specification of positional codes.

## DISTINGUISHING THEORIES EMPIRICALLY

In spite of the various strengths and weaknesses of specific models discussed above, important differences remain between the three theories of serial order. The difference between chaining and positional theories is obvious: The retrieval cue in the former is the previous element; the retrieval cue in the latter is some (abstract) positional code. The difference between positional and ordinal theories is less obvious, but relates to whether the position of an element in a sequence can be defined independently of its surrounding elements. In positional theories, it can; in ordinal (and chaining) theories, it can not. The consequence is that, in ordinal models, the middle element in a sequence can only be retrieved after retrieval of its predecessors (or successors, in the model of Shiffrin & Cook, 1978). In positional models however, it is possible to retrieve the middle element without retrieving its neighbors, by reinstating the appropriate positional cue. This is crucial in explaining a class of positional errors found in serial recall (see below).

Previous means of testing theories of serial order include the serial learning (e.g., Young, 1968) and probed recall (e.g., Murdock, 1968) paradigms. However, these paradigms have had little success in distinguishing the theories, mainly because of methodological problems (see Henson, 1996). A better way to test theories of serial order is to examine the errors people make when they misrecall a list of items. Though a single error may reflect a temporary failure to realize an accurate representation in memory, large numbers of errors show striking patterns in their distribution (Conrad, 1959; Estes, 1972). A detailed classification of such errors is given in Appendix A. Moreover, because serial recall is an everyday cognitive activity (e.g., recalling a telephone number), it less likely to be contaminated by specialized strategies than are other unusual, and perhaps artificial, laboratory tasks (such as probed recall), particularly given that people are often unaware of their errors (Henson, 1996).

### *Evidence against Chaining Theories*

One way to test chaining theories is to ask whether recall of an item depends on the properties of its predecessor. In other words, to look for factors that affect the cueing of an item. One factor might be the phonological confusability of items. It is well established (e.g., Bjork & Healy, 1974; Conrad & Hull, 1964; Estes, 1973; Hintzman, 1968) that serial recall of lists of similar sounding items (e.g., *BDGPTV*) is considerably worse than recall of lists of dissimilar items (e.g., *RMHQVJ*), even when presented visually. Wickelgren (1965) attributed this effect to associations between the phonemes of each item: Because the repeated phonemes of phonologically similar items are associated with more than one successor, there will be some uncertainty in the recall of subsequent items. In other words, Wickelgren predicted an effect of similarity on cueing. This prediction is also holds for distributed (e.g., Lewandowsky & Murdock, 1989) and compound (e.g., Jordan, 1986) chaining models, as soon as they adopt phonological representations.

To test for an effect of similarity on cueing, Henson, Norris, Page and Baddeley (1996) used lists of items that alternated in their phonological confusability (e.g., *BMGQTJ*), following an example by Baddeley (1968). A strong effect of similarity on cueing predicts that most errors should follow the confusable items (i.e., on the nonconfusable items *M*, *Q* and *J*). However, serial position curves showed the opposite, with more errors in recall of confusable items than the alternated nonconfusable items (see ahead to Fig. 13 for a graphical illustration). More importantly, recall of nonconfusable items in the alternating lists did not differ significantly from recall of nonconfusable items on corresponding positions in control lists with no confusable items. In other words, the probability of recalling a nonconfusable item appeared independent of whether or not the previous item was confusable. This failure to find any detectable effect of similarity on cueing is troublesome for any model like Wickelgren's that chains along phonological representations. Further still, the probability of recalling a nonconfusable item appeared independent of whether or not the previous confusable item was recalled correctly. This lack of any effect of errors on cueing is troublesome for any closed-loop chaining model that assumes responses are fed back to cue subsequent items.

One might argue for a model in which interitem associations are made between nonphonological representations, and in which cueing is independent of response feedback. However, if the nonphonological representations are type representations, such a model will still predict an effect of repetition on cueing. In other words, in a list such as *RMHMVJ,* there should be more errors on positions following a repeated item than on corresponding positions in control lists with no repeated items. In fact, chaining theory predicts that these errors are likely to be exchanges between the immediately following items (e.g., *RMHMVJ* recalled as *RMVMHJ*), given that they share the same cue. Wickelgren (1966) called such errors *associative intrusions,* and reported that they were more common in lists with repeated items (repetition lists) than control lists, supporting chaining theory.

However, Wickelgren's evidence was weak, with the differences between repetition lists and control lists only reaching significance in three of his eight conditions. In several recent experiments (Henson, 1996), any such differences were small and failed to reach significance, in spite of a more powerful design and a stricter scoring scheme. More importantly, any differences that are found could have alternative explanations, given that the sheer presence of repeated items in a list can have several effects on recall of that list (Henson, 1996). For example, because there are fewer different items to guess from in a repetition list than a control list (by virtue of the repetition), a simple guessing hypothesis also predicts a higher baseline chance of an associative intrusion in repetition lists. Thus there does not appear to be any conclusive evidence for an effect of repetition on cueing either.

The failure to find any evidence for an effect of similarity, errors or repetition on cueing is problematic for existing chaining models. This is not to deny that there is a specific type of chaining model that can be constructed so as to be consistent with the above data (such a model might chain along nonphonological, token representations for example, independently of response feedback). However, given that there is not, as yet, any positive evidence for chaining, and that there is positive evidence for positional information (see below), it seems reasonable to argue against chaining theory on the grounds of parsimony.

### *Evidence for Positional Theories*

The most common errors in serial recall are order errors, or transpositions. The most striking aspect of these errors is their distribution: Erroneous items are clustered around their correct position, rather than being randomly distributed (e.g., Estes, 1972). This is apparent in transposition gradients, which show, for each position in participants' reports, the proportion of items from each position in the corresponding lists. These proportions peak when the input and output position match (i.e., for correct responses) and decrease as the difference between the input and output position increases (see ahead to Fig. 5 for an example).

Transposition gradients are often taken as evidence for positional theories. They suggest that items are coded for their position in a sequence, but that there is some similarity between these codes that occasionally causes errors. However, they do not force this conclusion, because identical transposition gradients can be produced by ordinal models (as illustrated by the Primacy Model of Page & Norris, in press): Errors in the relative order of nearby items also produce peaked transposition gradients, without any coding of the position of those items. In fact, appropriate transposition gradients can also be produced by compound chaining models (Henson, 1996).

However, there are two types of error that do necessitate the use of positional information. The first of these occurs when lists are grouped. Grouping items by the timing of their occurrence, for example, is well known to improve recall (e.g., Ryan, 1969). Though grouping reduces the overall incidence of errors, one type of error actually increases (Wickelgren, 1967). These *interpositions* (Henson, 1996) are transpositions between groups that maintain their position within groups (Appendix A). With groups of three for example, interpositions are seen as an increase three-apart transpositions (and this increase is not simply the result of whole groups swapping; Lee & Estes, 1981). These errors imply that items can be coded for their position within a group independently of surrounding items.

The second type of positional errors is found between trials. Conrad (1960) reported that an erroneous item in one trial is more likely than chance to have occurred at the same position in the previous trial (see also Estes, 1991). Henson (1996) called the errors caused by such proactive interference of positional information *protrusions.* Interestingly, when recall on the previous trial is incorrect, protrusions are more likely to come from the position of recall than the position of presentation (Henson, 1996). This suggests that the recall episode is itself a learning episode, in that items recalled erroneously are recoded in their output position. In any case, these errors imply that items can be coded for their position within a trial.

In summary, though error analysis provides no evidence for chaining theory, it provides unequivocal evidence for positional theory. This evidence derives from the class of *positional errors:* those substitutions between sequences that maintain their position within a sequence. Such errors cannot be attributed to errors of relative order within a sequence and are therefore inexplicable by ordinal theory. In the next section, an example positional theory is given in the form of a new model of serial recall. This model not only explains positional errors, but also many other phenomena in short-term memory that follow naturally from its assumptions about the nature of positional information.

## THE START-END MODEL

In brief, SEM assumes that position in a sequence is coded relative to the start and end of that sequence. This positional information is encoded during each presentation and rehearsal of an item, creating an episodic token in short-term memory. The order of items is retrieved by cueing with positional codes for each position of recall and selecting the best matching token. Each of these assumptions is examined below (a more precise formalization of SEM is given in Appendix B).

### *Core Assumptions of SEM*

*1. Coding of position.* The start and end of a sequence are normally the most salient aspects of that sequence. As such, they provide potential reference points, or anchors, with which the elements of the sequence can be ordered. With this idea in mind, SEM's coding of position presumes a start marker and an end marker (Houghton, 1990). The start marker is strongest at start of a sequence, and decreases in strength towards end of the sequence. Conversely, the end marker is weakest at the start, and grows in strength towards the end. The relative strengths of the start and end markers therefore provide an approximate two-dimensional code for each position in a sequence.

In the spatial domain, the start and end markers might correspond to the left and right boundaries of a horizontal array, for example. The relative distances from these two extremina therefore code an item's position within that array. In the temporal domain, the start and end markers correspond to the initiation and termination of a sequence. Here, one might wonder how an item's position can be coded with respect to an end marker at its time of presentation, if the end of the sequence has not yet occurred. One possibility is that the strength of the end marker corresponds to the degree of expectation for the end of the sequence. This possibility, together with other interpretations of the start and end marker, is discussed in Henson (1997). For present concerns, start and end markers can be regarded as a simple means with which to formalize positional information.

More specifically, the strength of the start and end markers for position $i = 1, 2, \ldots, N$ in a list of $N$ items, $s(i)$ and $e(i)$ respectively, can be parameterized as:

$$s(i) = S_0 S^{i-1} \qquad e(i) = E_0 E^{N-i} \tag{1}$$

where $S_0, E_0 > 0$ are the maximum strengths of the start and end markers, and $0 < S, E < 1$ are the change in start and end marker strength over positions. The upper panel of Fig. 2 shows example strengths of a start and end marker for each position $i = 1 \ldots 6$ in a sequence of six items.

The code for position $i$ can be represented by the vector $\boldsymbol{p}(i) = (s(i), e(i))$. For example, the first position in Fig. 2 has the code (*1.00, 0.33*), the second position has the code (*0.80, 0.41*), etc. These codes are approximate, in the sense that they share some similarity with one another. This similarity is

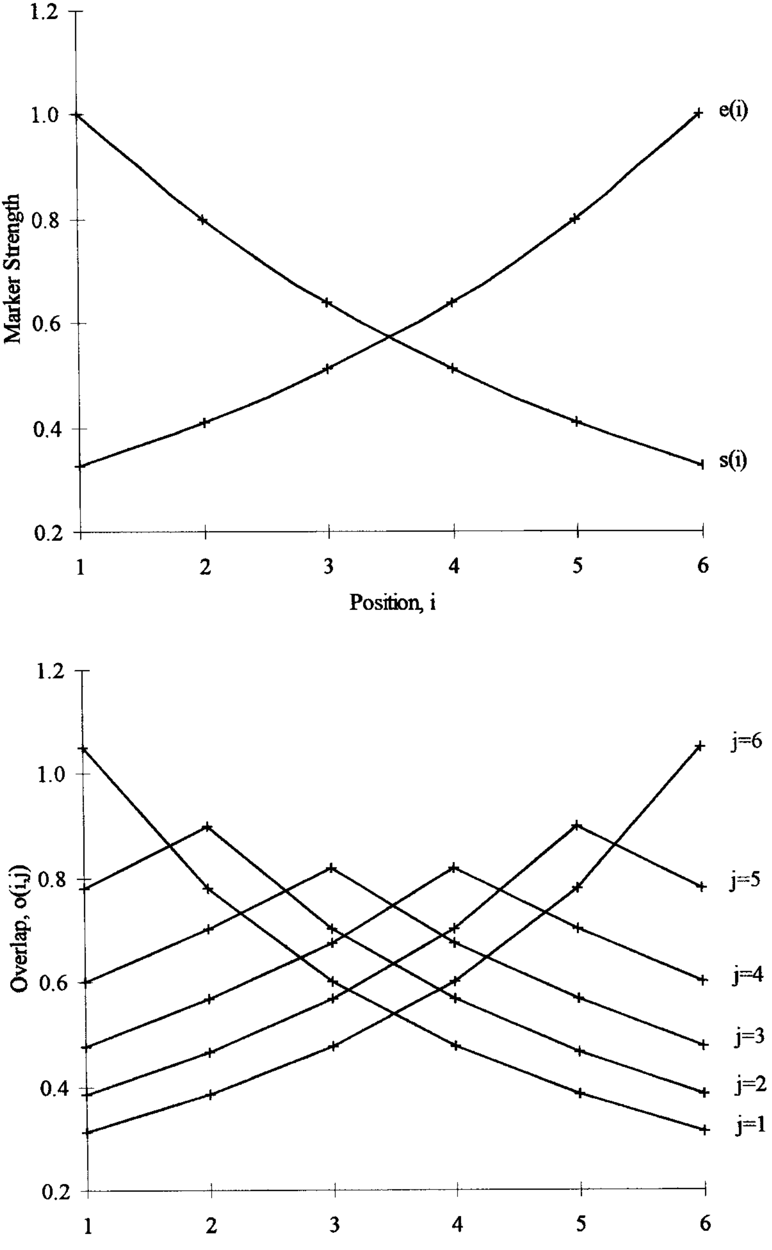


**FIG. 2** Start and end marker strengths, $s(i)$ and $e(i)$ (upper panel), and corresponding positional uncertainty functions (lower panel) for Positions $i, j = 1 \ldots 6$ of a six-item list, $N = 6$, $S_0 = E_0 = 1.00$, $S = E = 0.80$.

defined by the overlap, $o(\boldsymbol{p}(i), \boldsymbol{p}(j))$, between vectors $\boldsymbol{p}(i)$ and $\boldsymbol{p}(j)$ representing two positional codes:

$$o(\boldsymbol{p}(i), \boldsymbol{p}(j)) = \{\boldsymbol{p}(i) \cdot \boldsymbol{p}(j)\}^{1/2} \exp\left\{-\left(\sum_k (p_k(i) - p_k(j))^2\right)^{1/2}\right\} \quad (2)$$

where $k$ indexes the (two) components of each vector. The lower panel of Fig. 2 shows the overlap between positional codes for positions $i, j = 1 \ldots 6$, using the start and end markers in the upper panel. The six peaked functions show that overlap is maximal when positions match ($i = j$) and decreases as positions get further apart (i.e., as $|i - j|$ increases). Each function indicates the *positional uncertainty* associated with a position. The sharper functions for terminal positions mean that there is less positional uncertainty for terminal than medial positions. These positional uncertainty functions resemble the pattern of responses in position-probed item recall (Fuchs, 1969) and item-probed position recall (McNicol, 1975).

The exponential term in Equation 2 is the Euclidean metric of similarity between two vectors (McNicol & Heathcote, 1986; Nosofsky, 1986), which is simply sharpened by the exponentiation. This produces the basic triangular-shape of the positional uncertainty functions. The prior term in Equation 2 is the square-rooted, inner product of two vectors, comprising the combined strength of the start and end markers for the two positions. This premultiplier lowers and widens the positional uncertainty functions for medial positions relative to terminal positions. In general, the height of the positional uncertainty functions is increased by increasing the maximum strengths of the markers (increasing $S_0$, $E_0$), and the sharpness of the functions is increased by increasing the change of marker strength over positions (decreasing $S$, $E$).

*2. Position-sensitive tokens.* Each occurrence of an item is assumed to create a new token in short-term memory (as in multiple-trace theories, e.g., Hintzman, 1986). These tokens are episodic records that a particular item occurred in a particular spatiotemporal context. In other words, memory for an item is ‘‘colored’’ by the context in which it was perceived, such that the representation of an item at the start of a sequence is quite different from the representation of the same item at the end of a sequence. Thus, short-term memory is not viewed as a subset of active long-term memory representations (Cowan, 1993), but as a set of new, episodic tokens. The assumption that order is stored over token rather than type representations allows SEM to represent sequences with repeated items (Henson, 1996, in press).

In SEM, tokens contain several components. Some components represent item information; others represent positional information. For example, the encoding of a list *RMHQVJ* would produce six tokens like those depicted

A

*{ [R] (1.00, 0.33) }*
*{ [M] (0.80, 0.41) }*
*{ [H] (0.64, 0.51) }*
*{ [Q] (0.51, 0.64) }*
*{ [V] (0.41, 0.80) }*
*{ [J] (0.33, 1.00) }*

B

*{ [R] (1.00, 0.64) (1.00, 0.80) }*
*{ [M] (0.80, 0.80) (1.00, 0.80) }*
*{ [H] (0.64, 1.00) (1.00, 0.80) }*
*{ [Q] (1.00, 0.64) (0.80, 1.00) }*
*{ [V] (0.80, 0.80) (0.80, 1.00) }*
*{ [J] (0.64, 1.00) (0.80, 1.00) }*

C

*{ [R] (1.00, 0.64) (1.00, 0.80) (0.85) }*
*{ [M] (0.80, 0.80) (1.00, 0.80) (0.86) }*
*{ [H] (0.64, 1.00) (1.00, 0.80) (0.88) }*
*{ [Q] (1.00, 0.64) (0.80, 1.00) (0.90) }*
*{ [V] (0.80, 0.80) (0.80, 1.00) (0.92) }*
*{ [J] (0.64, 1.00) (0.80, 1.00) (0.94) }*
... ... ... ...
*{ [R] (1.00, 0.64) (1.00, 0.80) (0.96) }*
*{* ***[H]*** *(0.80, 0.80) (1.00, 0.80) (0.98) }*
*{* ***[M]*** *(0.64, 1.00) (1.00, 0.80) (1.00) }*

**FIG. 3** Positional tokens assumed by SEM for (A) an ungrouped list of six items, (B) a list of six items grouped as two groups of three, and (C) recall of the first of two groups of three items, with the addition of contextual change (see text for details).

in Fig. 3A, where the first component [*X*] represents some code for the identity of Item *X*, and the second component represents the positional codes described above.

*3. Response competition.* Tokens in short-term memory are unordered; their ordering occurs during recall. To recall a sequence, SEM cues each response by reinstating the positional code of the position being recalled. For example, the cue for the second response in the previous example can be depicted as {[?] (*0.80, 0.41*)}. This cue is matched against all tokens in parallel, with the overlap between the positional code in the cue and the positional code in the tokens defined by Equation 2. These overlaps determine the strengths with which each item competes for output. This competition is held over a set of type representations, activated in proportion to the maximum overlap between the cue and the tokens of each type. Access to these long-term memory representations is assumed necessary in order to give a categorical response. To model the occasional errors made by people, random noise is added to these activations before choosing the strongest for output.

A further assumption of SEM's recall process is that once an item has been recalled, its type representation is temporarily suppressed. This reduces the probability of recalling an item more than once within the same trial, given that repetitions are rare (Henson, 1996). This assumption is common to most models of serial recall (e.g., Burgess & Hitch, 1992; Lewandowsky & Li, 1994; Page & Norris, in press) and has independent support from the fact that people often fail to recall the second occurrence of a repeated item (the Ranschburg effect; Henson, in press; Jahnke, 1969). Indeed, the suppression of previous actions appears to be a general process in sequential behavior (Houghton & Tipper, 1996).

For the simple case of a list of unique items, the strength with which Item $i$ competes for Response $j$, $c(i, j)$, is given by:

$$c(i, j) = o(\boldsymbol{p}(i), \boldsymbol{p}(j))\,(1 - r(i)) + n \qquad (3)$$

where $o(\boldsymbol{p}(i), \boldsymbol{p}(j))$ is the overlap between positional codes for Positions $i$ and $j$ (Equation 2), $0 <= r(i) <= 1$ is the suppression of Item $i$, and $n$ is a random variable drawn from a Gaussian distribution with a mean of zero and standard deviation given by the parameter $G_C$.[2] In the simplest form of SEM, suppression is absolute in the sense that $r(i) = 1$ as soon Item $i$ is recalled in a given trial, and $r(i) = 0$ otherwise.

## DEMONSTRATIONS

Equations 1 to 3 comprise the most basic form of SEM. Given the probabilistic nature of response selection, and the fact that the probability of recalling an item depends on what has been recalled previously (via the suppression term in Equation 3), analytical solutions to SEM's predictions are difficult to obtain. Consequently, the equations have been implemented in a computer program that simulates serial recall. In fact, the program can simulate recall of the same lists given to participants, producing reports that can be compared directly (in the following simulations, the model was run for 100,000 trials, for which the variability of SEM's predictions is insignificant).

The ability of SEM to capture the important characteristics of short-term memory is illustrated in six demonstrations below. These particular demon-

---

[2] Selection via the addition of Gaussian noise and choice of the strongest competitor closely approximates selection via Luce's Choice Rule (M. Page, personal communication, August, 1996). The present approach is favored because of its mechanistic appeal and its compatibility with a simple output threshold (see Demonstration 1).

strations were chosen to illustrate that SEM can reproduce the full range of errors associated with immediate serial recall. Where possible, approximate quantitative fits to data are given in terms of the root mean square error (RMSE) between SEM's predictions and the mean values of the data. More precise fits to these and other data, which allow for the variance and covariance found in the data, are given in Henson (1996).

As with any model, it is important to minimize the number of parameters that are free to fit the data. As described so far, the basic model has five parameters ($S_0$, $E_0$, $S$, $E$ and $G_C$). To reduce the number of free parameters, the start marker parameters were fixed at $S_0 = 1.00$ and $S = 0.80$. The end marker parameters were redefined in relation to these values, replacing the four marker parameters with two free parameters, $F_0 = E_0/S_0$ and $F = E/S$. In other words, $F_0$ represents the maximum strength of the end marker relative to that of the start marker, and $F$ represents the degree of change of the end marker strength relative to that of the start marker.

In progressing through the subsequent demonstrations, the basic form of SEM is extended and generalized. In particular, Demonstrations 1 to 4 use a single-trial version of SEM, which does not model intertrial effects, while Demonstrations 5 and 6 use a multiple-trial version. This incremental exposition of SEM entails some change in parameters as new assumptions are added. However, each extension of the model subsumes previous versions, with the most general formalization of SEM given in Appendix B, together with the parameter values used in each demonstration.

### *Demonstration 1*

The purpose of this demonstration is to show that SEM can reproduce the basic serial position curve and underlying error distributions in serial recall of a list of six items. With the three free parameters set at $F_0 = F = 0.60$ and $G_C = 0.08$, SEM recalled the list correctly on 57% of occasions (a figure typical of lists of phonologically dissimilar letters). The solid line in the upper panel of Fig. 4 shows the corresponding serial position curve, with the prolonged primacy and last-item recency that are characteristic of immediate serial recall of span-length lists.

The precise form of the serial position curve is best considered from the underlying distribution of transpositions. The upper panel of Fig. 5 shows the transposition gradients produced by SEM for each output position. The six points in each gradient indicate the proportion of responses from input positions 1 to 6, from left to right. The peaks of these transposition gradients are correct responses; the slopes indicate decreasing numbers of transpositions as the transposition distance increases (e.g., Estes, 1972). These gradients reflect the graded nature of positional information in SEM. The lower panel of Fig. 5 shows the corresponding transposition gradients from lists of phonologically dissimilar letters in Experiment 1 of Henson et al. (1996): the RMSE of SEM's fit to the 36 data points is 4.02%.

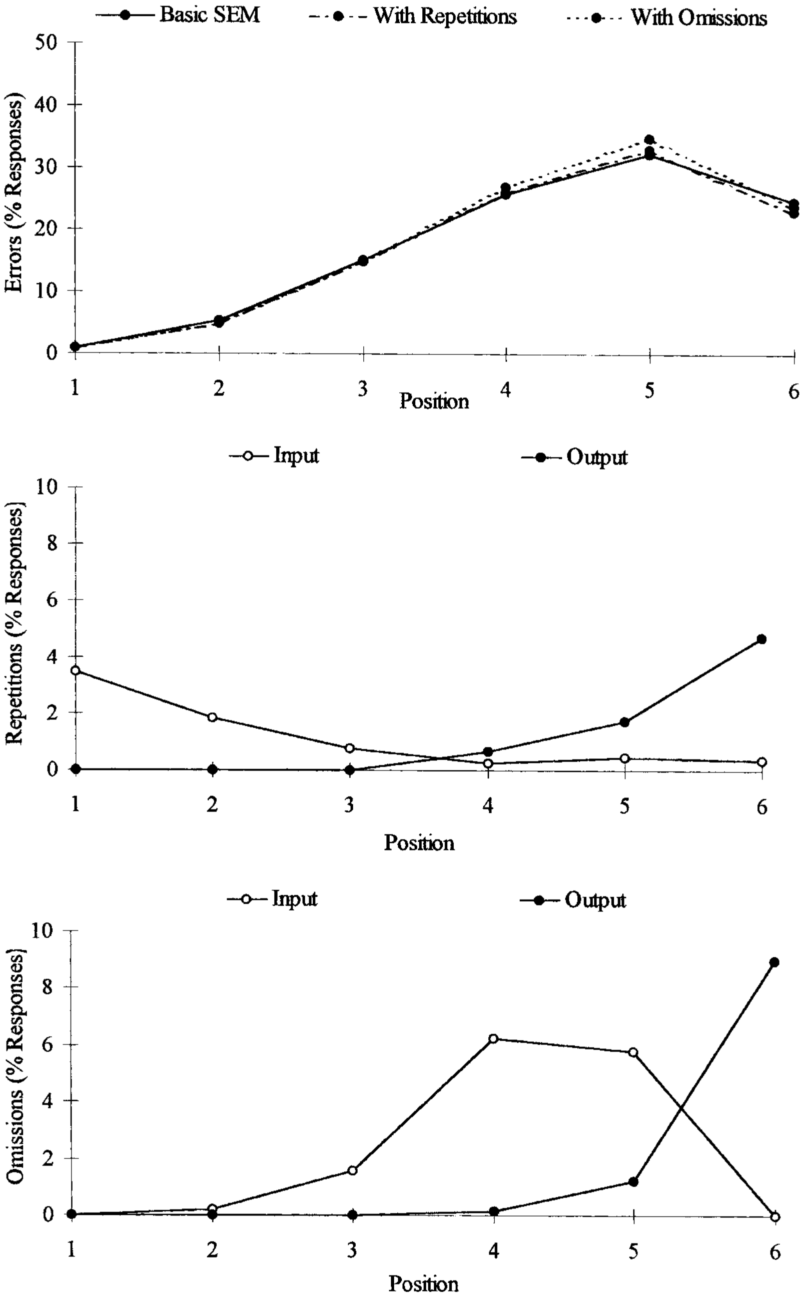


**FIG. 4** Serial position curve (upper panel) and distribution of repetitions (middle panel) and omissions (lower panel) by input and output position produced by SEM in Demonstration 1.

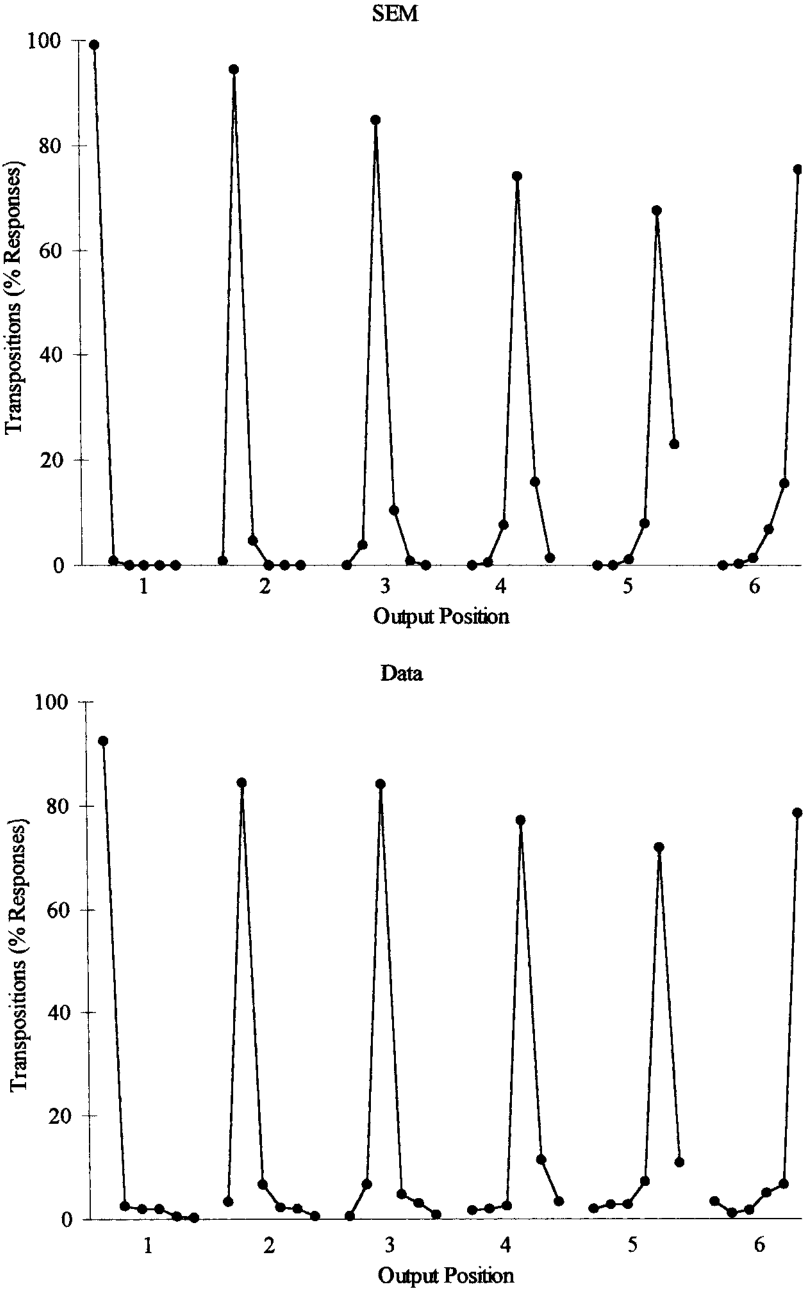


**FIG. 5** Transposition gradients from SEM (upper panel) and from the data (lower panel) in Demonstration 1. The data are replotted from the lists of phonologically dissimilar letters in Experiment 1 of Henson et al. (1996).

The main reason for the primacy and recency effects produced by SEM is the greater positional certainty, or distinctiveness, associated with terminal positions (lower panel of Fig. 2), resulting in fewer transpositions on these positions. The greater extent of primacy than recency arises because the start marker is stronger than the end marker (i.e., $F_0 < 1$), producing greater positional certainty for early items.[3] However, the exact shape of the serial position curve also depends on the complicating effects of suppression, which can enhance recency by virtue of the fact that later items are faced with fewer competitors. Thus, the serial position curve in Fig. 4 is determined by an interaction between the shape of positional uncertainty functions and the sequential act of recall.

*Repetitions.* The basic form of SEM assumes permanent suppression of an item after it is recalled (at least until a new trial begins). A more realistic assumption is that suppression is temporary, wearing off during recall of subsequent items. This entails a new parameter $R_S$, reflecting the rate of exponential decay of suppression (Equation B6 in Appendix B). Suppression is maximal immediately following recall of Item $i$ ($r(i) = 1$), but decreases during subsequent responses, eventually returning to the baseline level ($r(i) = 0$) between trials.

The transient nature of suppression allows SEM to produce an important subclass of transpositions, repetitions. With $R_S = 0.50$ (and the other parameters remaining as above), repetitions comprised 7% of errors. Though rare, they showed clear constraints on their distribution: Most repetitions occurred towards the end of recall and were repetitions of the first few items in the list (middle panel of Fig. 4). Indeed, the two occurrences of an item were generally far apart in a report (3.7 positions on average, cf. a figure of 3.4 in the data, Henson et al., 1996), reflecting the time necessary for appreciable decay of suppression. Though the number of repetitions is often too small for

---

[3] This parametrization not only reproduces the appropriate asymmetry of serial position curves: When combined with the slower change of the start marker relative to the end marker (i.e., $F < 1$), it also provides the appropriate level of *fill-in.* Fill-in is the tendency for items not recalled at their correct position to follow shortly after (Page & Norris, in press). For example, a detailed analysis of transpositions (Henson, 1996) shows that, if Item $i + 1$ is recalled too early on Position $i$, the most likely next response on Position $i + 1$ is Item $i$ (rather than Item $i + 2$). In other words, a paired transposition between adjacent items is more common than a ‘‘slippage’’ of subsequent items. In SEM, fill-in arises because the effect of a stronger and longer-lasting start marker is to skew positional uncertainty functions towards earlier positions, so biasing response competition in favor of earlier over later items. This bias is vital in modeling serial recall, because without it, the tendency for a slippage of items following an error can actually remove the recency effect (for further discussion, see Henson et al., 1996). Another way of viewing this bias is in terms of suppression: The negative skewing of positional uncertainty functions means that the effect of suppression is to enhance recency by reducing the number of competitors in recall of later items. (Conversely, a positive skewing would enhance primacy by virtue of the fact that later items would have a greater opportunity to be recalled too early and suppressed so that they cannot be recalled in their correct position).

more detailed quantitative analysis, the above pattern is consistently found in meta-analyses of serial recall data (Henson, 1996). Thus the addition of the parameter $R_S$ is not justified in order to produce better quantitative fits, but in order to explain an important subclass of errors.

*Omissions.* As it stands, the basic form of SEM produces only transpositions. Yet people will often omit a response if they cannot remember an item. Such omissions are modeled in SEM by the addition of an output threshold, $T_O$. The strongest competitor in Equation 3 is selected as before, but if its strength does not exceed $T_O$, then it is not output, and an omission is indicated instead.

With $T_O = 0.35$ (and the other parameters remaining as above), omissions comprised 10% of errors. When plotted against output position (lower panel of Fig. 4), omissions showed a monotonic increase with position. This increase in omissions towards the end of recall is again typical in meta-analyses of serial recall data (Henson, 1996). When plotted against input position however (i.e., whether the item at each input position was recalled anywhere in the report), omissions showed a recency effect. In other words, the last item was more often recalled somewhere than the penultimate item. This bowed pattern of item errors is typical for lists of span-length or longer (e.g., Drewnowski, 1980; Drewnowski & Murdock, 1980).

How does this somewhat paradoxical pattern of omissions arise? The short answer is that people will occasionally recall the last item too early, and having done so, will often follow this error with omissions. In SEM, this pattern arises because the sharply tuned end marker ($F < 1$) means that the positional uncertainty function for the last position is also very sharp. In other words, only the last item is cued strongly at the last position, and if that item has already been recalled and suppressed, it is less likely that other items, such as the penultimate item, will be cued above the output threshold. The fact that Item $N$ is more likely to be recalled in Position $N - 1$ than Item $N - 1$ is to be recalled in Position $N$ leads to a recency effect when omissions are scored against input position, but not when scored against output position (see Henson, 1996, for further discussion of SEM's account of repetitions and omissions).

No other model appears able to explain this pattern of omissions. The Perturbation Model (Lee & Estes, 1977, 1981) generally assumes that omissions are flat or monotonic across input position, which may be true of short lists (Healy, 1974), but is not true of longer lists (Drewnowski, 1980; Henson, 1996). The Primacy Model (Page & Norris, in press) and the Articulatory Loop Model (Burgess & Hitch, 1992) produce omissions that increase towards the end of recall, but only through more omissions of the last item than any other. In SEM, the pattern of omissions is an emergent property of the dynamics of the recall process and the simple assumption of a weak yet sharply tuned end marker.

Finally, note that the additions of an output threshold and the decay of suppression produce a change in the nature rather than number of errors underlying the serial position curve, the shape of which was little changed by these additions (broken lines in the upper panel of Fig. 4). This reinforces the caution needed in evaluating models of short-term memory: a reasonable fit to serial position curves does not imply that the models are reproducing the appropriate pattern of underlying errors.

### *Demonstration 2*

The purpose of this demonstration is to illustrate SEM's performance as a function of list length. With the parameters maintaining the same values as in Demonstration 1 ($F_0 = F = 0.60$, $G_C = 0.08$, $R_S = 0.50$, $T_O = 0.35$), SEM was tested with lists of 2 to 10 items. The proportion of lists recalled correctly was a sigmoidal function of list length (upper panel of Fig. 6), with a 50%-span of just over 6 items. The RMSE of SEM's fit to 7 data points (adapted from Crannell & Parrish, 1957) is 3.25%.

Serial position curves ‘‘stretched-out’’ as list length increased (middle panel of Fig. 6), with the effect of list length being most apparent on medial positions (often seen as an interaction between list length and position; e.g., Drewnowski & Murdock, 1980).[4] Lengthening the lists also altered the relative proportions of the different error types (lower panel of Fig. 6), with the dominance of transpositions for short lists giving way to approximately equal numbers of transpositions and omissions for longer lists. This shift in error types is reflected in the data, with a RMSE of only 1.65% over 9 data points taken from the incidence of transpositions, omissions and repetitions for lists of seven, eight and nine items in Experiment 2 of Henson (1996; data not shown).

The effects of list length arise because, as the number of positions coded by the start and end marker increases, the resolution of each code decreases. This is illustrated in the upper panel of Fig. 7, which shows the positional uncertainty functions for the middle position in lists of 3, 5, 7 and 9 items. As list length increases, the positional uncertainty functions become flatter and lower (i.e., positions are coded less effectively). The flattening causes more transpositions, through a smaller signal-to-noise ratio, and the lowering causes more omissions, as the positional uncertainty functions approach the output threshold. SEM therefore provides a rationale for the limited capacity of short-term memory: The limit of $7 \pm 2$ items (Miller, 1956/1994) may well reflect a limit on the number of positions that can be reliably distinguished. Note that this property is an automatic consequence of SEM's start

---

[4] List length will often exert a larger effect on the primacy portion of the serial position curve than produced by these simulations of SEM. However, an extended version of SEM that takes into account the additional delay between presentation and recall of each item (i.e., output interference, Demonstrations 5 and 6) produces a greater detrimental effect of list length on recall of the first few items.

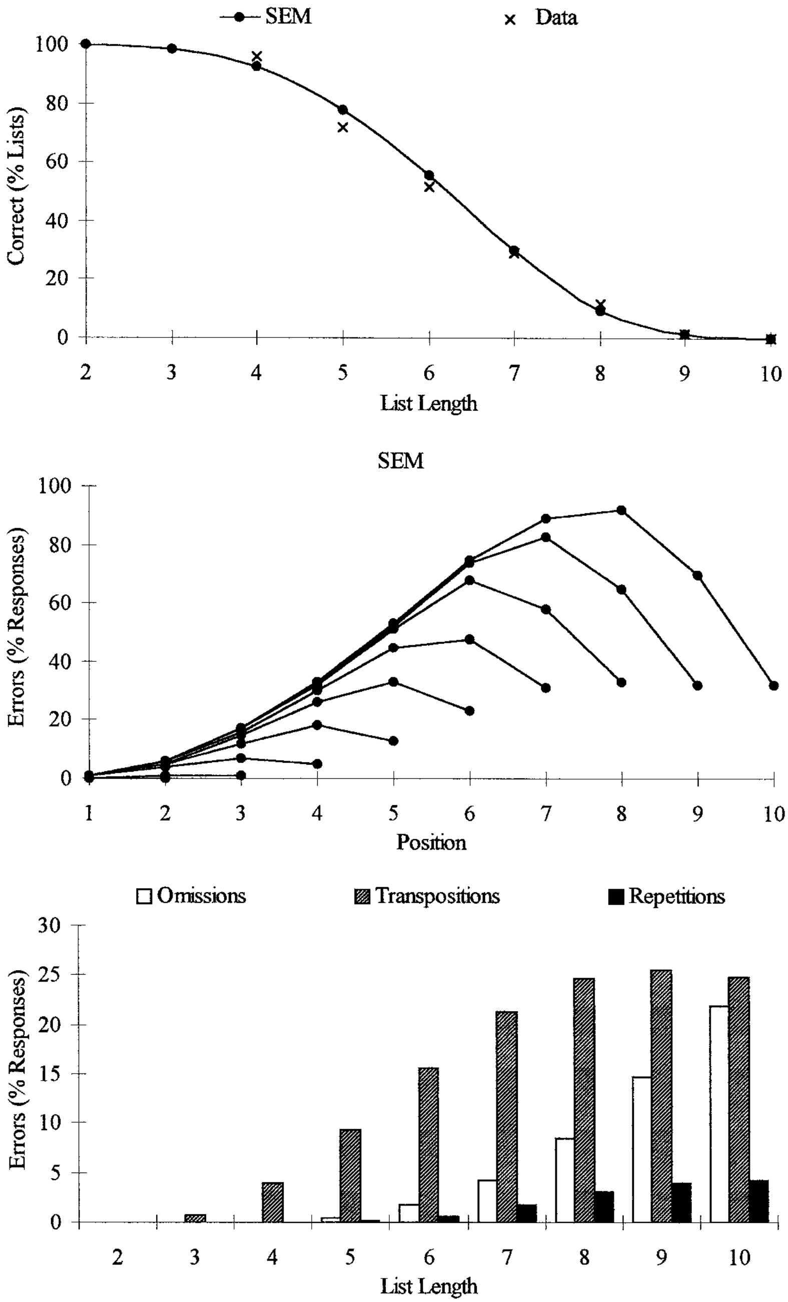


**FIG. 6** Lists correct (upper panel), serial position curves (middle panel) and proportions of different error types (lower panel) as a function of list length in Demonstration 2. The data replotted in the upper panel are averaged across the limited and unlimited letter conditions of Crannell and Parrish (1957).

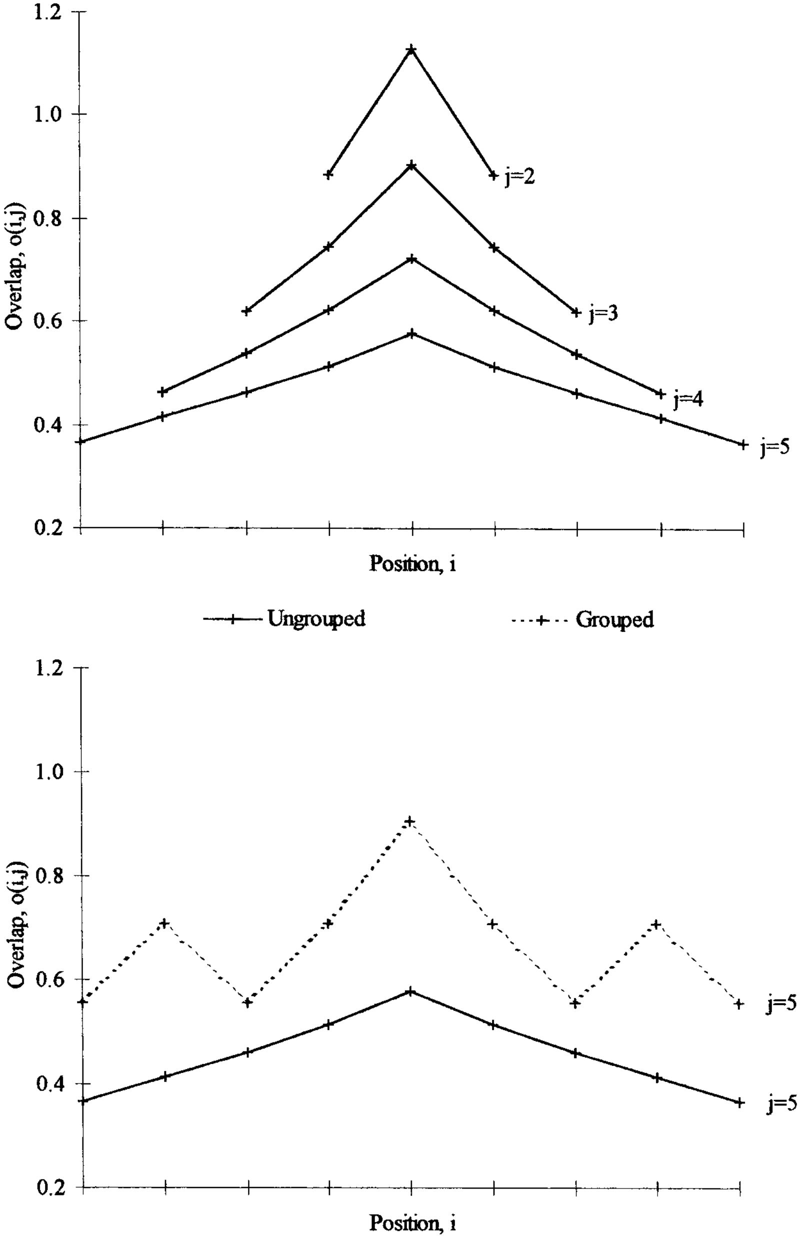


**FIG. 7** Positional uncertainty functions for the middle Position $j$ = 2,3,4,5 of three-, five-, seven- and nine-item ungrouped lists in Demonstration 2 (upper panel) and for the middle Position $j = 5$ of nine-item lists both ungrouped and grouped in threes in Demonstration 3 (lower panel), with $F_0 = F = 1$ as in Figure 2.

and end markers. It is not a property of other positional codes, such as the context signals of Burgess and Hitch (1992; 1996) and Brown et al. (in press), or the unspecified codes of the Perturbation Model (Lee & Estes, 1981).

### *Demonstration 3*

The purpose of this demonstration is to show how SEM models grouping and, in particular, the interpositions between groups that support a positional theory of short-term memory for serial order. The basic idea behind grouping in SEM is that group boundaries allow the insertion of additional markers at the start and end of each group, providing extra anchor points with which to order items. This results in two pairs of start and end markers, one coding the position of an item in a group, and the other coding the position of a group in a list, producing a hierarchy of positional codes (Lee & Estes, 1981). For example, the encoding of a list *RMH QVJ*, grouped as two groups of three, would produce tokens like those depicted in Fig. 3B. The leftmost positional code represents position of item-in-group; the rightmost code represents position of group-in-list. The cue for each response also contains two such codes, and the positional uncertainty function is determined simply by multiplying the overlaps for the item-level and group-level codes (Appendix B).

The effect of a second dimension of positional coding is shown in the lower panel of Fig. 7. The solid curve is the positional uncertainty function for the middle position of an ungrouped nine-item list (identical to that in the upper panel); the broken curve is the corresponding function when the list is grouped as three groups of three. The effect of grouping is to improve the positional resolution of the middle position, particularly with respect to immediately surrounding positions. This is because each pair of start and end markers is only coding three, rather than nine, positions. The smaller peaks for Positions 2 and 8 reflect the fact that these positions share the same code for position of item-in-group (differing only in their code for position of group-in-list). Interpositions are the errors that arise when the differences between these peaks are bridged by the additive noise.

Serial position curves for grouped and ungrouped nine-item lists are shown in the upper panel of Fig. 8. To fit the data, which were taken from Experiment 1 of Hitch, Burgess, Towse and Culpin (1996), the strength of the end of group marker (coding position of item-in-group) was set as $F_{0,\mathrm{I}} = 0.40$ and the strength of the end of list marker (coding position of group-in-list) was set as $F_{0,G} = 0.00$ (i.e., no end of list marker was assumed). The ungrouped lists were treated as single groups of nine items. The noise parameter was increased to $G_C = 0.14$; the remaining parameter values were maintained from Demonstration 2. These values gave a reasonable (though not perfect) fit to the data, with a RMSE over the 18 data points of 13.30%. Most importantly, the effect of grouping was apparent in the ‘‘scalloping’’ of the serial position curve, with mini-primacy and recency effects appearing within each group (e.g., Ryan, 1969).

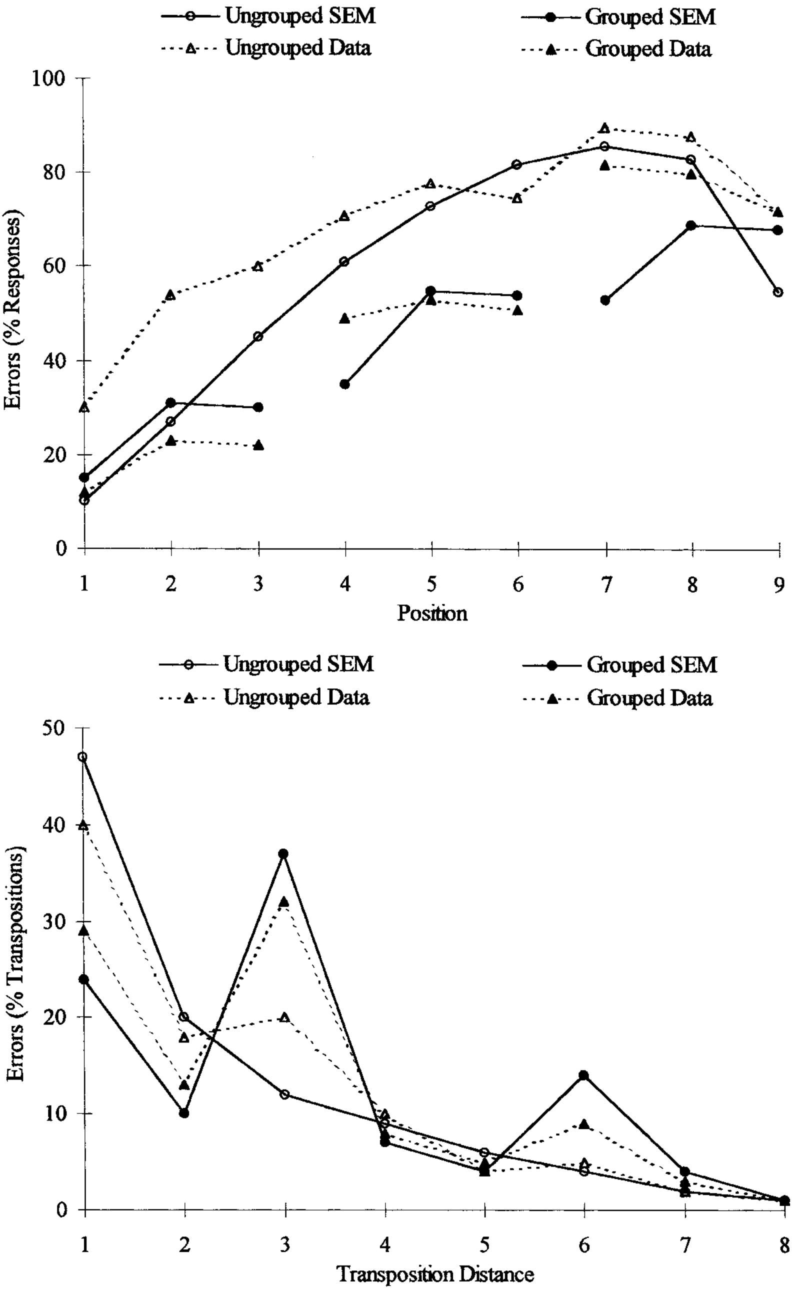


**FIG. 8** Serial position curves (upper panel) and averaged transposition gradient (lower panel) for nine-item lists both ungrouped and grouped in threes in Demonstration 3. The data in the upper panel are replotted from Experiment 1 of Hitch et al. (1996); the data in the lower panel are replotted from Experiment 2 of Henson (1996).

As well as decreasing the total number of transpositions, grouping had a dramatic effect on their distribution (lower panel of Fig. 8). This panel shows the frequency of transpositions averaged over transposition distance (the difference between input and output position). Since the corresponding data for the Hitch et al. experiment were not reported, the data in this panel come from an almost identical experiment by Henson (1996): The RMSE over the 16 data points is 3.62%. For the ungrouped lists, the number of transpositions decreases monotonically with transposition distance. For the grouped lists however, there are peaks for three- and six-apart transpositions, representing transpositions between groups that maintain their position within groups. These are the interpositions that indicate an explicit coding of position in group (Lee & Estes, 1981).

As predicted by previous models of grouping (e.g. Burgess & Hitch, 1996; Frick, 1989; Lee & Estes, 1981), the main advantage of grouping was a decrease in one- and two-apart transpositions across group boundaries, which accounted for 35% of transpositions in ungrouped lists (when scored as if grouped), but only 8% in grouped lists. However, grouping also reduced the proportion of one- and two-apart transpositions within groups, from 41% in ungrouped lists to 26% in grouped lists. This is often found in the data (Henson, 1996), and arises in SEM because grouping improves the discrimination not only of positions between groups, but also positions within groups (owing to the smaller number of positions coded by each start and end marker). Again, this is an automatic property of SEM's positional coding, but not of other models (e.g., Burgess & Hitch, 1996; Lee & Estes, 1981), which attribute the grouping advantage solely to a reduction in transpositions between groups.

The considerable improvement in positional coding entailed by a second pair of start and end markers also provides a rationale for why most people will group lists subjectively, even without any objective grouping (Henson, 1996). The prevalence of such spontaneous grouping is apparent from the fact that most serial position curves for supraspan lists show some degree of scalloping (e.g., the ungrouped data of Hitch et al. in Fig. 8; see also the meta-analysis by Madigan, 1980); those that do not may reflect an averaging over several different grouping strategies (e.g., an eight-item list can be grouped 4-4, 3-3-2, 2-2-2-2, etc.). Thus the ungrouped serial position curves for lists of seven or more items in Fig. 6 and 8 may be more imaginary than real. Indeed, given the limited resolution of start and end markers, a strong prediction of SEM is that people can never recall such long lists correctly without grouping them subjectively.

### *Demonstration 4*

This demonstration illustrates how SEM might be extended to modality and suffix effects. The modality effect is the well-established finding that auditory or vocalized presentation leads to better recall than silent, visual presentation, particularly for the last few items in a list (e.g., Conrad & Hull, 1968; Margrain, 1967). One account of this effect appeals to an additional source of information associated with auditory, but not visual, material, such as the Precategorical Acoustic Store (PAS) of Crowder and Morton (1969). The PAS holds a temporary ‘‘echo’’ of the most recent, auditory events, which can aid recall of the last few items in a list. The limited capacity of the PAS is supported by the suffix effect: the fact that an irrelevant item occurring at the end of an auditory list attenuates the modality effect, perhaps by masking auditory information about the last item (Crowder, 1978).

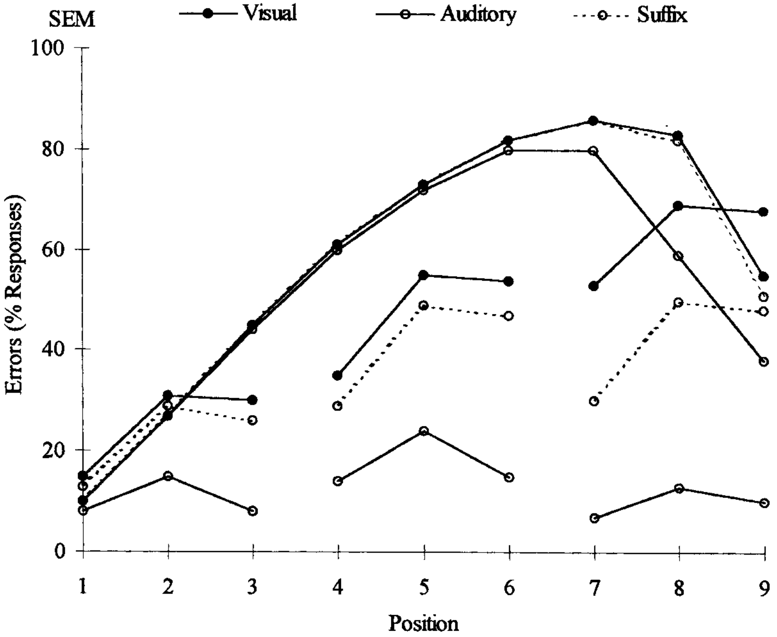
However, the original PAS account cannot explain effects of modality on preterminal items (e.g., Penney, 1989), effects of mouthed or lip-read stimuli (e.g., Campbell & Dodd, 1980) or effects of changing-state stimuli (e.g., Glenberg, 1990). An alternative account of the modality effect appeals to superior coding of the serial order of auditory material (Drewnowski & Murdock, 1980). This might arise through better temporal resolution of auditory than visual material (Glenberg & Swanson, 1986), or simply better positional coding (Greene & Crowder, 1988).

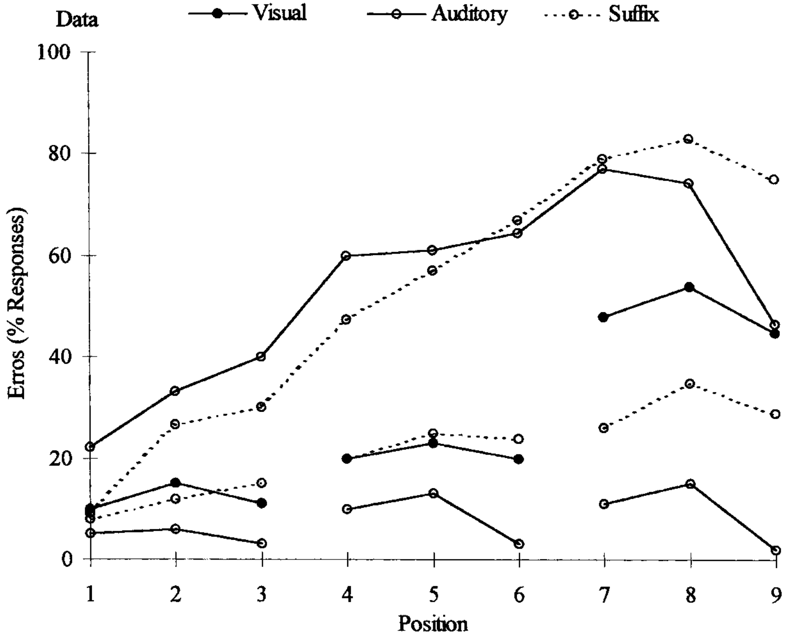
One way to model better positional coding in SEM is to increase the strength or sharpness of SEM’s start and end markers. The inherent temporal properties of auditory information may allow better definition of the start and end of a sequence. The effect of a stronger end marker is illustrated by the serial position curves in the upper panel Fig. 9. The visual ungrouped curve was generated with parameter values identical to those in Demonstration 3; the only difference for the auditory ungrouped curve was the strength of the end marker, which was increased to $F_{O,I} = 0.90$. The stronger end marker produces a modality effect that extends over the last three or four items of ungrouped lists. This advantage reflects not only greater distinctiveness of final positions (Glenberg, 1990), but also improved item memory for the last few items (the stronger end marker increasing the probability that these items are cued suprathreshold).

Stronger end markers can also explain the extended modality effect in grouped lists, which is difficult to explain according to the original formulation of the PAS (Frankish, 1989). As apparent in the grouped data shown in the lower panel of Fig. 9 (taken from Frankish, 1985), auditory presentation of temporally grouped lists improves recall on almost every position, particularly the ends of groups. In the corresponding simulations of SEM

---

**FIG. 9** Effects of modality and redundant suffixes in grouped and ungrouped lists in SEM (upper panel) and in the data (lower panel) for Demonstration 4. The visual grouped data are replotted from Experiment 1 of Frankish (1985); the auditory and suffix grouped data are replotted from Experiment 3 of Frankish (1985); the auditory and suffix ungrouped data are replotted from the Animal Sheep and Human Sheep conditions in Experiment 1 of Neath et al. (1993).


SEM
Visual
Auditory
Suffix
Errors (% Responses)
100
80
60
40
20
0
1
2
3
4
5
6
7
8
9
Position


Data
Visual
Auditory
Suffix
Erros (% Responses)
100
80
60
40
20
0
1
2
3
4
5
6
7
8
9
Position

(upper panel of Fig. 9), auditory presentation was assumed to result in a stronger end marker at both the end of groups and at the end of lists (i.e., $F_{O,G} = F_{O,I} = 0.90$). This produces better coding not only of the final item in groups, but also of the final group in lists (Frankish, 1974), explaining the large difference in the effectiveness of auditory and visual grouping (Frankish, 1985).

The notion of end markers also allows a simple extension to the suffix effect. If it is assumed that the end marker codes the suffix in the last position, instead of the last item, then the positional coding of preceding items will suffer. In other words, a nine-item list followed by a suffix will be coded as if there were ten positions, and recall of the last few items will be impaired as a consequence (even though only the nine list items compete for output). This is illustrated by the broken lines in the upper panel Fig. 9, which show the effects of a suffix at the end of an ungrouped auditory list, and at the end of each group in a grouped auditory list. For these simulations, the strengths of the end markers were set as if the list or group were one position longer (i.e., $F_{O,I} = F_{O,G} = 0.90 \times E_I = 0.43$). The suffix attenuates the modality effect in both cases, in agreement with the data shown in the lower panel of Fig. 9. The overall RMSE to the 45 data points is 14.06%, which is reasonable considering that the data come from three separate experiments (see figure legend).

Whether or not the end marker includes a suffix in coding the last position of a list may depend on the degree to which the suffix is perceptually grouped with the list items (Frankish & Turner, 1984; Kahneman, 1973; LeCompte & Watkins, 1995). Indeed, the ungrouped auditory data in the lower panel of Figure 9 are taken from an experiment by Neath, Surprenant and Crowder (1993), in which the same physical stimulus (the onomatopoetic word *baa*) was heard in two different contexts. In the context of animal sounds produced by a person, the stimulus produced a larger suffix effect than in the context of animal sounds produced by real animals. In other words, a larger suffix effect was obtained when the ambiguous sound was labeled as speech, and presumably therefore grouped with the list items, than when it was labeled as nonspeech, and presumably not grouped with the list items. Similar attenuations of the suffix effect are found when an auditory suffix differs from list items in voicing, location, or rhythm (Frick, 1988). Likewise, the effect of a suffix in the visual modality is attenuated when the suffix is perceptually segregated from the list items (Frick & De Rose, 1986). Though others have argued against a grouping account of the suffix effect (Penney, 1978, 1985; Crowder, 1978), it seems likely that perceptual grouping is at least one of the factors involved (e.g., Morton, 1976).[5] Another factor is suggested in Demonstration 6.

---

[5] Interestingly, SEM provides an alternative explanation for the results of Penney (1978). In her 333(1) condition, Penney presented three groups of three items followed by a suffix

It must be emphasized that the above approach to modality and suffix effects in SEM is only illustrative. The assumption of a stronger end marker for auditory than visual material is somewhat ad hoc, and the magnitude of the suffix effect depends on the particular values chosen for the end marker parameters. Other aspects of the modality effect remain unexplained, such as its interactions with recall order (Broadbent, Cooper, Frankish & Broadbent, 1980), mixed modality lists (Greene, 1989), mixed modality distractors (Marks & Crowder, 1997) and the precategorical properties of items (Frankish, 1996; Surprenant, Pitt & Crowder, 1993). Nonetheless, the assumptions in Demonstration 4 provide a reasonable first step to modeling what have proved surprising complex and subtle effects of modality and redundant suffixes.

### *Demonstration 5*

The purpose of this demonstration is extend SEM to intertrial effects and, in particular, the protrusions between trials that further support a positional theory of short-term memory for serial order. This extension uses a multiple-trial version of SEM that makes two new assumptions.

The first assumption is another addition to SEM's tokens, representing the general context during the creation of each token. This context is nonpositional (i.e., cannot be reinstated at recall) and represents all other intrinsic (e.g., mood) and extrinsic (e.g., environmental) factors that change over the course of an experiment. For mathematical convenience, this context is represented by a single value (a one-dimensional vector), with the current context represented by the value 1, and older contexts represented by values less than 1. Thus, rather than updating the current context, the context of existing tokens in memory is multiplied by a new parameter $E_C < 1$ during each contextual change. For example, assuming that each presentation of an item results in a change of context and that $E_C = 0.98$, then the encoding of a grouped list *RMH QVJ* would produce the six tokens depicted above the dots in Fig. 3C (the dots depict the end of presentation). The rightmost vector represents the general context, with tokens for more recent items having

---

grouped temporally with the last group. This caused an impairment in recall that was restricted mainly to the ninth item, as expected from the perceptual grouping account. However, when the suffix was temporally separated from the last group (in her 3331 condition), an impairment was still found, which Penney used to argue against the perceptual grouping account. In SEM, these data can be explained if a grouped suffix affects the marking of the last item in the group, whereas an ungrouped suffix affects the marking of the last group in the list. In other words, if the suffix in Penney's 3331 condition were coded as a fourth group of one item, the position-in-list coding of the preceding group would suffer. This explains why the impairment in Penney's 3331 condition was found for all items in the last group: Though a grouped suffix only impairs recall of items within that group, an ungrouped suffix impairs recall of whole groups. A similar account may explain why a prefix grouped with a list during presentation impairs recall of the whole list (Crowder, 1967).

vectors closer to (*1.00*). During recall, the general context for each cue is always the current context. The overlap between the general context of the cue and the general context of each token is determined in exactly the same manner as for positional codes (via Equation 2), and the combined positional uncertainty functions are determined by multiplying the overlaps of all three vectors (Appendix B).

The second assumption is that every rehearsal of an item (including recall) creates a new token in memory. Importantly, the item is recorded in its output position, which may or may not be correct, and the general context of the new token is updated to the current context. In other words, the continual updating of contextual information corresponds to maintenance rehearsal in short-term memory. This is also illustrated in Fig. 3C, where the contents of short-term memory are shown after an attempt to recall the first group of three items in the list *RMH QVJ.* The six tokens above the dots are those created during presentation; the three below are those created during recall. Note that erroneous recall of items *H* and *M* means that they are recorded in new (incorrect) positions.

In SEM, contextual changes are assumed to occur across *episodes,* defined for example as the occurrence of a new item or distractor. The parameter $C_P$ is the number of episodes between presentation of each item (e.g., the number of intralist distractors), $C_D$ is the number of episodes during the delay between presentation and recall (e.g., a filled retention interval), $C_R$ is the number of episodes between recall of each item, and $C_I$ is the number of episodes between trials (e.g., a filled intertrial interval). These parameters are not free; they are determined by the experimental design. The last new parameter, $C_A$, represents additional contextual changes between trials. This parameter was free to vary in the simulations below, in order to capture changes in intrinsic context (e.g., shifts of attention) that typically occur between the end of one trial and the start of the next.[6]

To compare the multiple-trial version of SEM with the single-trial version in Demonstration 1, 100,000 lists of six items were created by random selection without replacement from a vocabulary of 12 items. The parameters $C_P = C_R = 1$ and $C_D = C_I = 0$ were set to simulate continuous presentation and recall of each list, and the parameters $E_C$ and $C_A$ were set as $E_C = 0.98$ and $C_A = 5$. The values of remaining parameters were identical to those in Demonstration 1.

SEM recalled 25% of lists correctly, a considerable reduction when com-

---

[6] Such changes reflect situations when, for example, one ‘‘thinks of something else’’, in order to put the previous trial out of mind. The parameter $C_A$ therefore captures the important difference between contextual change and real-time change: A considerable shift in attention may occur in the few seconds between trials, resulting in large differences in intrinsic context over a small length of time. The notion of context used in SEM is not simply a case of relabelling time.

pared with the figure of 57% in Demonstration 1. This reduction resulted from contextual change both during and between trials (as governed by the values of the parameters $C_P$, $C_R$, and $C_A$). In other words, the multiple-trial version of SEM illustrates the combined effects of input, output and proactive interference. The effect of proactive interference was apparent in the 28% of errors that were intrusions. These intrusions replaced some of the omissions and transpositions in Demonstration 1, with only 6% of errors being omissions and the remaining 66% of errors being transpositions.

The nature of the proactive interference is apparent from the *intrusion gradients* in the upper panel of Fig. 10. Taking the subset of erroneous items that also occurred somewhere in the previous report, each gradient shows the proportion of such errors on a given output position that occurred at each position in the previous report. These gradients suggest that an erroneous item was likely to have come from a nearby position in the previous report, with the peaks of the gradients representing protrusions from the same position. The peaks are clearly above chance (one sixth, or 17%) for all positions, though they are highest for terminal positions, where positional coding is sharpest (Fig. 2). The gradients in the lower panel (taken from Experiment 5 of Henson, 1996) are less clear cut, owing to considerable noise in the data. This noise reflects the small numbers of intrusions involved and the confounding effects of spontaneous 3-3 grouping by participants. Nonetheless, the same qualitative pattern is apparent, and the RMSE of SEM's fit to the 36 data points is only 9.02%.

The overall incidence of output protrusions from the previous report (42%) was greater than the incidence of input protrusions from the previous list (38%), in agreement with the data (Henson, 1996). This arises because the tokens created during recall of the previous list are more recent than the tokens created during presentation of that list (their general context therefore having greater overlap with that of the cue). This means that an erroneous item in the previous trial is more likely to protrude in its position of recall than its position of presentation.

The effect of increasing the interval between trials is shown in Fig. 11. In these simulations, the parameter $C_I$ was increased from 0 to 36, to simulate, for example, the number of distractor items between trials. As the intertrial interval increases, recall improves slightly, reaching an asymptotic level of 36% of lists correct (open circles in the upper panel). In other words, with long intertrial intervals, the effects of proactive interference have worn off (the remaining 21% difference from Demonstration 1 resulting from input and output interference within trials). This is mirrored by the decrease in the incidence of output protrusions (filled circles), which returns to chance levels for long intertrial intervals, in agreement with the findings of Conrad (1960). Unfortunately, few studies have systematically varied the intertrial interval in this manner. The four data points at $C_I = 2$ and $C_I = 20$ come from an experiment that used 5 words followed by 3 distractor digits (Henson 1996,

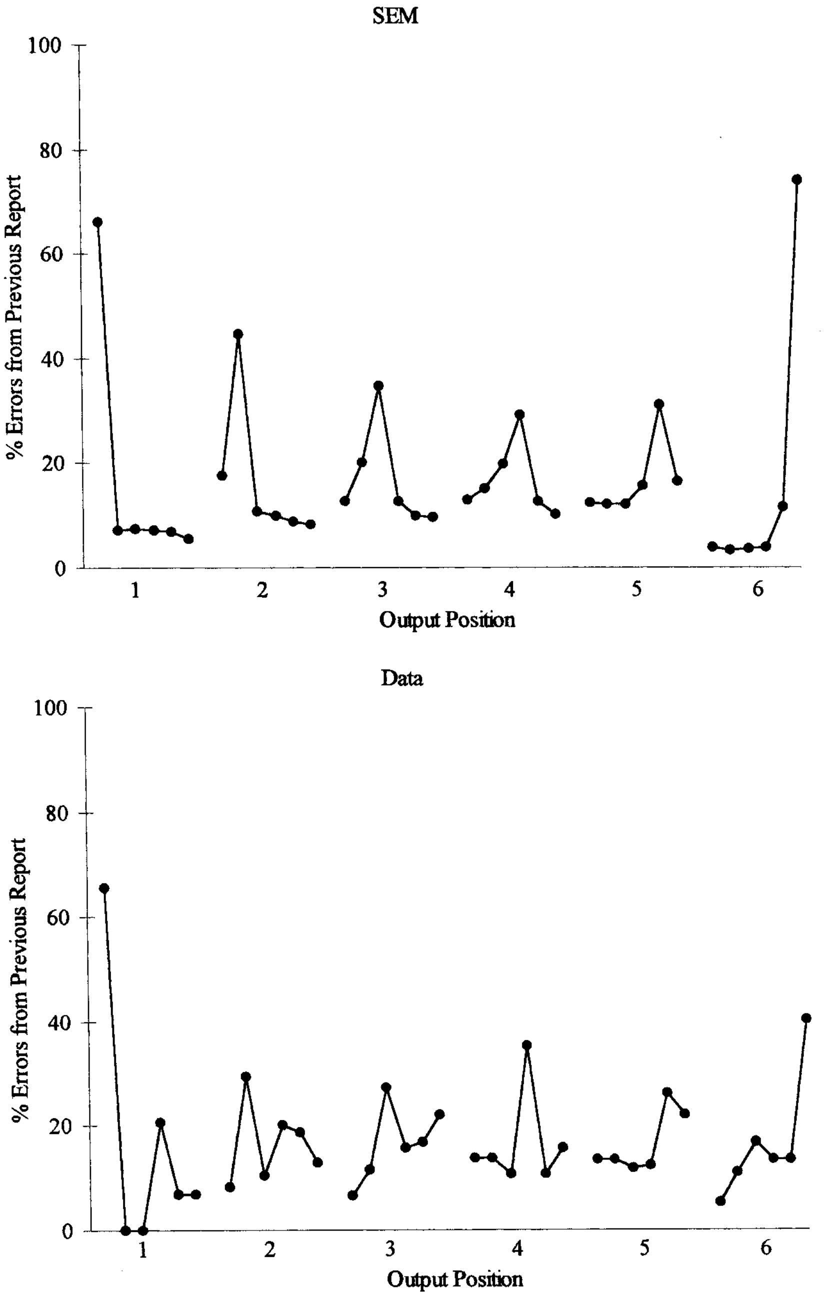


**FIG. 10** Intrusion gradients showing the proportion of erroneous items at each position of a report that occurred at each position of the previous report from SEM (upper panel) and from the data (lower panel) for Demonstration 5. The data are replotted from the Fixed condition in Experiment 5 of Henson (1996).

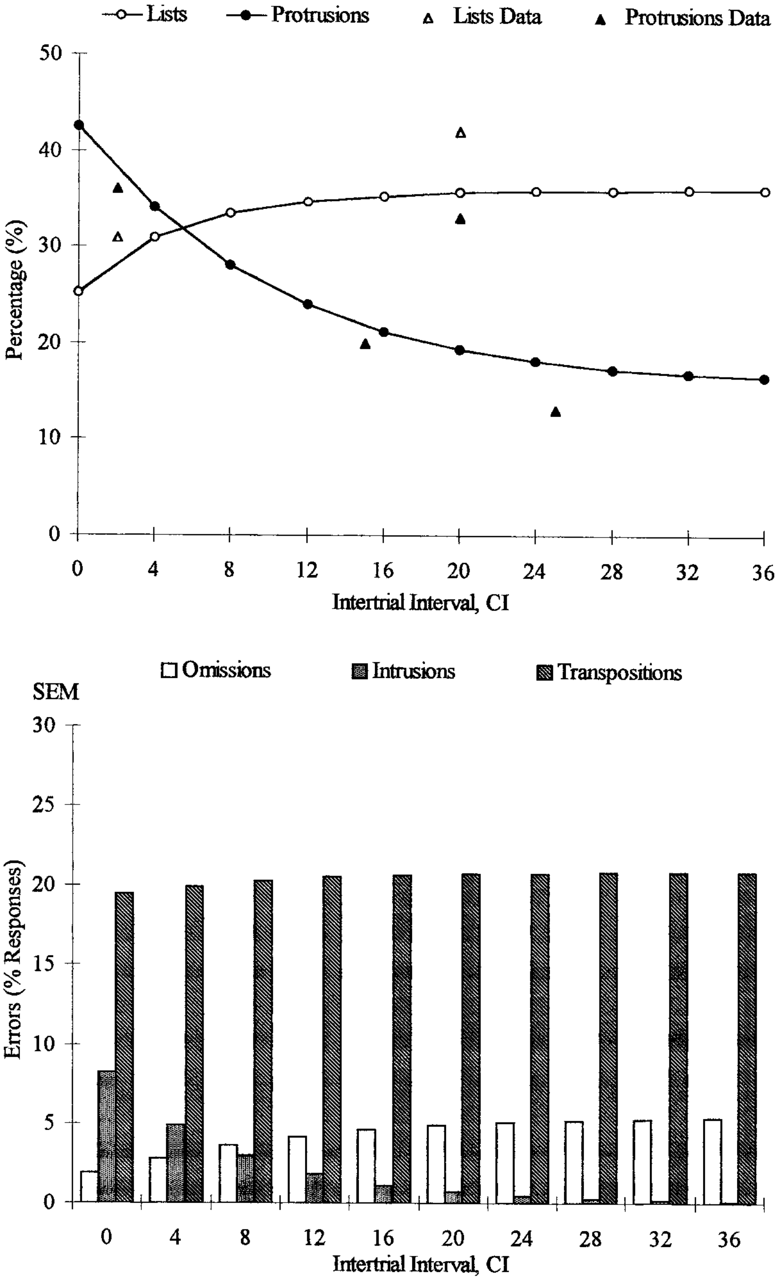


**FIG. 11** Percentage of lists correct and incidence of protrusions (upper panel) and proportions of different error types (lower panel) as a function of intertrial interval in Demonstration 5. The data in the upper panel are replotted from Conrad (1960) and Experiment 3 of Henson (1996); see text for details.

Experiment 3) and the two data points at $C_I = 15$ and $C_I = 25$ come from an experiment that used 8 digits (Conrad, 1960), meaning that the simulation results are not directly comparable with the data. Nonetheless, SEM can be seen to exhibit the main qualitative effects of intertrial interval that are shown in the data.

The decrease in protrusions with longer intertrial intervals is accompanied by a decrease in the overall numbers of intrusions, from over 8% of responses when $C_I = 0$ to under 1% when $C_I = 36$ (lower panel of Fig. 11). This decrease is partially offset by an increase in omissions, whereas the incidence of transpositions is virtually unaffected by the intertrial interval, in agreement with the data (Henson, 1996).

In summary, extending SEM to multiple-trial effects allows it to capture the full range of errors, including intrusions. These intrusions tend to maintain positions between trials, particularly between recall episodes, as found in the data. Few other models reproduce these intertrial effects (most are ‘‘trial-unit’’ rather than ‘‘continuum’’ models, Estes, 1991); not one has explicitly modeled multiple presentation and recall episodes in the same detail as SEM.

### *Demonstration 6*

The purpose of this final demonstration was to extend SEM to the effects of retention interval and phonological similarity. These effects are related, in that the rapid forgetting over the first few seconds of a filled delay is accompanied by a decrease in phonological confusions. In fact, the incidence of phonological confusions approaches chance levels after approximately 30 s of filled delay (Bjork & Healy, 1974; Estes, 1973), suggesting a rapid decay of phonological information in short-term memory (Baddeley, 1986; Tehan & Humphreys, 1995).

Phonological information in SEM is modeled as the transient activation of a set of phonological representations. These representations are assumed to be activated during output, in the retrieval of an item’s phonological form (cf. Levelt, 1989). They are also assumed to be activated during the encoding of auditory items, and recoding of visual items (Baddeley, 1986). More specifically, each presentation and rehearsal of an item activates its phonological representation to a fixed amount which subsequently undergoes exponential decay (Appendix B). Thus immediately after presentation of a list, the phonological activations of list items form a ‘‘recency-gradient’’, with the last item most active in memory, as might be expected from item recognition tasks (e.g., Monsell, 1978; McElree & Dosher, 1989).

The main benefit of the short-lived phonological activations is to provide an additional item memory (akin to the phonological store of Baddeley, 1986), keeping responses above the output threshold and reducing the incidence of intrusions (see below). However, the disadvantage of the phonological activations is that they introduce the potential for phonological confusions. During recall, an item is selected as in previous Demonstrations, but before it is output, a further competition is held over the phonological representations (with noise governed by the parameter $G_P$). Normally, the appropriate phonological form is accessed and the corresponding item is output. Occasionally however, competition can result in access to a similar, but incorrect, phonological form, resulting in a confusion error (see Appendix B). Thus confusions arise at a second stage of response retrieval (Henson et al., 1996; Page & Norris, in press).

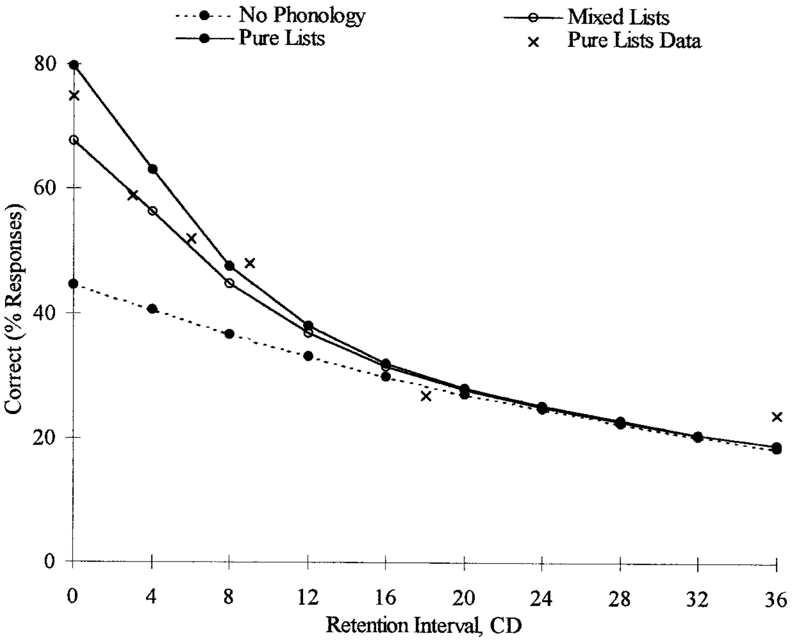
The effects of a filled retention interval can be simulated by increasing $C_D$, corresponding, for example, to the number of distractors between presentation and recall. The filled circles in the upper panel of Fig. 12 show the effect of increasing the retention interval on recall of the lists used in Demonstration 5. The parameter values were identical to those in Demonstration 5, except for the noise parameters $G_C = 0.06$ and $G_P = 0.14$. The broken line results from simulations with no additional phonological activations. This curve shows gradual forgetting as $C_D$ increases, but the rate of forgetting is too slow compared with that found in the data (the single crosses, replotted from Baddeley & Scott, 1971).[7] The solid lines show forgetting curves with phonological activations and the rate of phonological decay, $R_P$, set as $R_P = 0.20$. Immediate recall is improved, but the forgetting rate over the first dozen distractors is now much faster, converging with the broken curve for longer retention intervals. This reflects the rapidly-decaying influence of phonological information. The RMSE for the 6 data points is 6.95%.

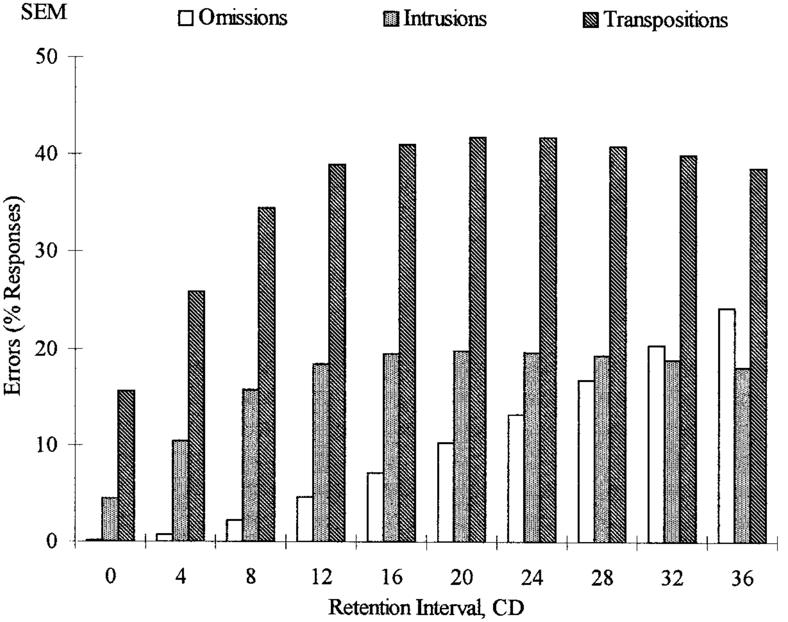
The open circles in the upper panel of Fig. 12 represent *mixed* lists of six items drawn randomly from a vocabulary of six confusable items (e.g., the letters *BDGPTV*) and six nonconfusable items (e.g., the letters *RMHQYJ*). The effect of phonological similarity is apparent in the poorer performance for mixed lists than pure lists when $C_D = 0$. However, this difference between pure and mixed lists disappears by the time $C_D > 20$. Indeed, a more sensitive measure of the proportion of intrusions that were phonological confusions (i.e., one confusable item swapping for another) gave a figure of 58% when $C_D = 0$, but had decreased to chance levels (25%) when $C_D = 20$.[8] By this time, any residual phonological activations are negligible.

---

[7] The data actually represent performance for lists of seven digits (Baddeley & Scott, 1971). For the purposes of this Demonstration however, performance for lists of seven digits is assumed to be comparable with that for lists of six letters (e.g., Crannell & Parrish, 1957).

[8] Interestingly, when expressed as a proportion of intrusions from trials other than the immediately preceding one, the incidence of confusions when $C_D = 0$ was even higher (68%), in agreement with Drewnowski & Murdock (1980). This is because intrusions have at least two sources: proactive interference from items at similar positions in previous trials (during the first stage of response selection), and confusions between items that are phonologically similar (during the second stage of response selection). When discounting the main source of proactive interference (the previous trial), the remaining intrusions are more likely to be confusions.


No Phonology
Pure Lists
Mixed Lists
Pure Lists Data
Correct (% Responses)
0
20
40
60
80
0
4
8
12
16
20
24
28
32
36
Retention Interval, CD


SEM
Omissions
Intrusions
Transpositions
Errors (% Responses)
0
10
20
30
40
50
0
4
8
12
16
20
24
28
32
36
Retention Interval, CD

The effect of retention interval on the different error types for the pure lists is shown in the lower panel of Fig. 12. There is a rise in all error types as $C_D$ increases from 0 to 12. In particular, the fact that the decrease in phonological confusions over this interval is accompanied by an increase in intrusions is consistent with the hypothesis of Tehan and Humphreys (1995). They proposed that short-lived phonological information can overcome the effects of proactive interference. In SEM, this arises because the greater phonological activation of more recent items aids their discrimination from items in previous trials (providing the items can be distinguished phonologically). For longer retention intervals, the incidence of transpositions and intrusions decreases as they become swamped by omissions. By this stage, the phonological activations have decayed to the extent that they no longer provide any additional item information. Interestingly however, of the intrusions remaining after such long delays, the proportion that are protrusions is still above chance. For example, the incidence of output protrusions when $C_D = 36$ was still 27% (cf. Demonstration 5). This is consistent with the long-lasting nature of positional information (Nairne, 1991).

One final property of phonological retrieval in SEM is shown in Fig. 13. The broken lines show serial position curves for further simulations of a special type of mixed list in which the confusable and nonconfusable items alternate in position (Baddeley, 1968; Henson et al., 1996). There are two such alternating curves, A1 and A2, depending on whether the alternation begins with a confusable or nonconfusable item (e.g., the sequences *BMGQTJ* or *MBQGJT* respectively). These curves have a ‘‘sawtooth shape,’’ with the peaks of the sawteeth representing more errors on confusable items than surrounding nonconfusable items. The solid line of closed circles, PN, is the serial position curve for lists of purely nonconfusable items and the solid line of open circles, PC, is the serial position curve for lists of purely confusable items. The phonological similarity effect is again apparent in the higher and flatter PC curves than PN curves. The only change in parameter values for these simulations was a decrease in the noise parameter $G_P = 0.10$. The RMSE of the SEM’s fit (upper panel) over the 24 data points (lower panel), replotted from Experiment 1 of Henson (1996), is 4.77%.

However, the most important aspect of these curves is apparent when the A1 and A2 curves are compared with the PN curve: The troughs of the sawteeth for the alternating lists are virtually coincident with the serial position curve for the pure nonconfusable lists, in agreement with the data (Bad-

---

**FIG. 12** Percentage of correct responses for pure lists with and without phonological activations and mixed lists with phonological activations (upper panel), and proportions of different error types for pure lists with phonological activations (lower panel) as a function of retention interval in Demonstration 6. The data in the upper panel are replotted from the seven item lists in Experiment 2 of Baddeley and Scott (1971).

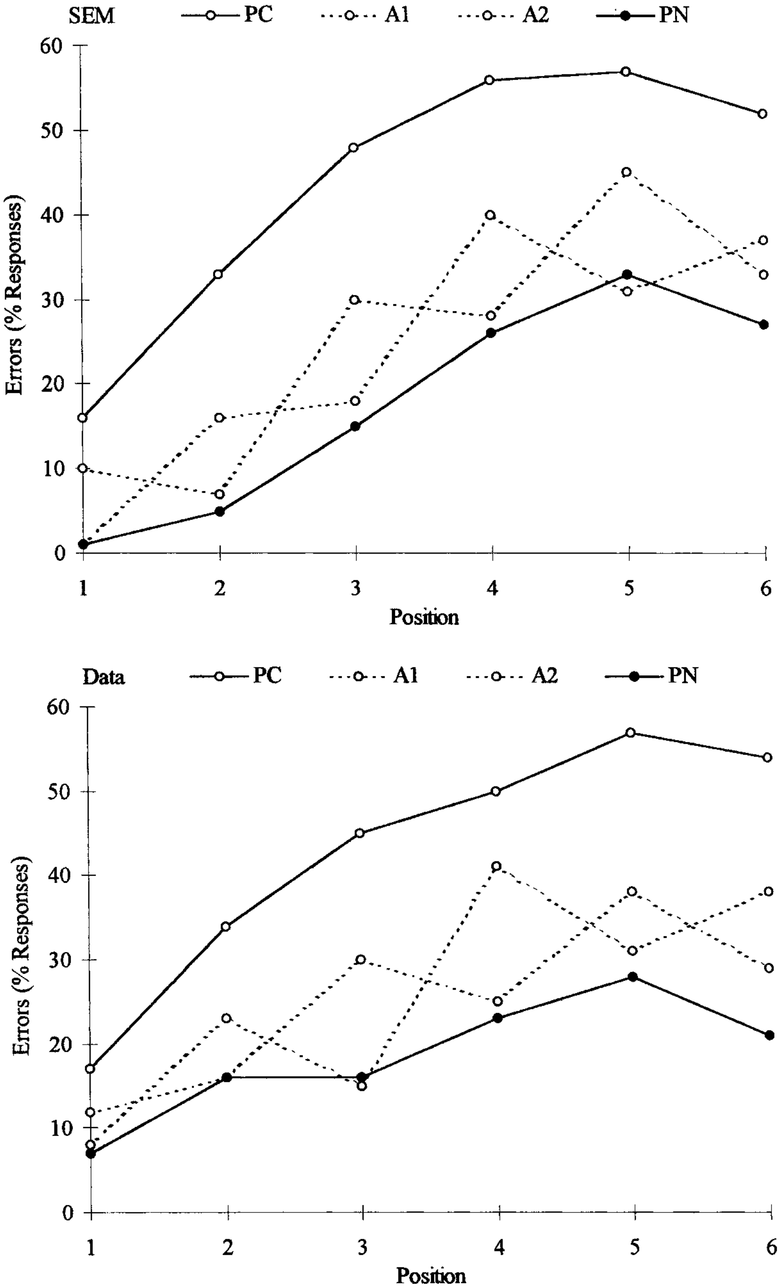


**FIG. 13** Serial position curves for pure nonconfusable (PN), pure confusable (PC) and alternating lists (A1,A2) from SEM (upper panel) and from the data (lower panel) in Demonstration 6. The data are replotted from Experiment 1 of Henson (1996).

deley, 1968; Henson et al., 1996). In other words, the presence of confusable items has little detrimental effect on recall of surrounding nonconfusable items (see also Bjork & Healy, 1974). This is troublesome for most existing models of serial recall, particularly chaining models, as argued earlier. SEM's success stems from three main reasons. Firstly, tokens are stored separately in memory, meaning that phonologically similar items do not interfere with each other. Such interference would occur in distributed memory stores (e.g., Jordan, 1986; Lewandowsky & Li, 1994). The second reason is that items are cued by positional codes that are independent of surrounding items. Thus there is no effect of phonological similarity on cueing, in contrast to models that chain along associations between items (e.g., Lewandowsky & Murdock, 1981; Wickelgren, 1965). Finally, the fact that confusions between similar items arise at a second stage of retrieval means that there is little effect of such errors on recall of subsequent items (see Henson et al., 1996 and Page & Norris, in press, for further elaboration of this point). These assumptions seem vital in order to model what have proved to be extremely constraining, ‘‘benchmark’’ data.

The assumption of rapidly decaying phonological activations also affords SEM closer fits to list-length and word-length effects (Henson, 1996), and indeed any effects, such as those of a redundant prefix or suffix, that can be partially attributed to the additional delay between presentation and recall of each item (Baddeley & Hull, 1979). A further assumption that auditory presentation results in higher initial levels of phonological activation than visual presentation, and that these activations can be masked by subsequent auditory input, might also explain some of the further subtleties of the modality and suffix effects discussed in Demonstration 4. In general, though one might object to the complexity of both contextual change and phonological decay in SEM, the end result is a model that captures an impressive range of experimental phenomena in short-term memory.

## DISCUSSION

In summary, the basic form of SEM produces the appropriate shape of the serial position curve. Though the same is true of many other models, few do justice to the distribution of different error types underlying the curve. With simple assumptions about the refractory nature of suppression and the presence of an output threshold, SEM is unique in reproducing the complete pattern of transpositions, repetitions and omissions. Moreover, the nature of positional coding in SEM provides a rationale for both the limited capacity of short-term memory and the prevalence of spontaneous grouping in memory span tasks. With the addition of contextual change, SEM is the first model to simulate fully the effects of multiple recall trials and the nature of the proactive interference entailed. Finally, SEM provides a unified account of the effects of phonological similarity and retention interval that is often lacking in other models.

### *Extension to Other Phenomena*

SEM can also be extended to other important phenomena in short-term memory. Though demonstrating fits to such data would exceed the present remit, the general approach that SEM might take is outlined below.

*Serial recall.* Other determinants of serial recall include articulation rate, irrelevant sound and articulatory suppression. The effects of these manipulations are well-accounted for, at least in qualitative terms, by the *working memory* framework (Baddeley, 1986). In general terms, SEM can be mapped onto this framework by assuming that its transient phonological activations correspond to Baddeley's phonological store, and that its rehearsal process corresponds to Baddeley's articulatory control process. SEM can then explain many of the same results explained by the working memory framework.

For example, if the rate of rehearsal in SEM is a function of articulation rate, then the time taken to articulate items will affect the amount of decay of phonological activations, and hence be an important determinant of memory span. This can explain why span is smaller for words that take longer to articulate, even when balanced for number of syllables and phonemes (Baddeley, Thomson & Buchanan, 1975). Indeed, the linear relationship often found between span and articulation rate (e.g., Hulme, Maughan & Brown, 1991) can be closely approximated by SEM (Henson, 1996). Moreover, the fact that phonological decay in SEM occurs during presentation and recall means that the effects of word length arise during both input (Page & Norris, in press) and output (Cowan, Day, Saults, Keller, Johnson & Flores, 1992). Thus span is not simply a fixed number of items (or chunks) as might be suggested from Demonstration 2; it also depends on rate with which those items can be rehearsed (Schweikert & Boruff, 1986).

Another factor affecting serial recall is the presence of irrelevant, background sound. For example, concurrent irrelevant speech during a serial recall task impairs performance, to a greater extent than comparable noise levels (Salame & Baddeley, 1982). According to the working memory account, the irrelevant auditory material has automatic access to the phonological store, where it interferes with the relevant material. However, an impairment comparable to that found with speech has also been found with tones (Jones & Macken, 1993), especially if the tones change in pitch, location, or rhythm. This suggests that the effects of irrelevant sound are not restricted to verbal material, and an alternative ‘‘changing-state’’ account has been proposed (Jones & Macken, 1995). In SEM, irrelevant sound might impair the precision of the positional coding of items, with the degree of impairment relating to the amount of change in the irrelevant stream. Irrelevant sound showing rapid changes over time (e.g., abrupt vowel transitions in speech) may interfere with the ability to mark the start and end of sequences. In particular, if irrelevant tones disrupt the ability to group items, some of the results of Jones and Macken (1993, 1995) may have arisen because the tones prevented spontaneous grouping.

A final factor that impairs serial recall is the concurrent articulation of an irrelevant item (e.g., repeating ‘‘the, the, the . . .’’; Murray, 1967). Under visual presentation, such articulatory suppression removes the effects of word-length (Baddeley et al., 1975), phonological similarity (Peterson & Johnson, 1971) and irrelevant sound (Salame & Baddeley, 1982). Under auditory presentation, articulatory suppression removes the effect of word-length (providing it continues throughout presentation and recall), but does not remove the effects of phonological similarity (Baddeley, Lewis & Vallar, 1984) or irrelevant sound (Hanley & Broadbent, 1987). According to the working memory theory, articulatory suppression not only prevents rehearsal, removing the word-length effect, but also prevents the recoding of visual material into the phonological store, removing the effects of phonological similarity and irrelevant sound. By making similar assumptions (i.e., that auditory material automatically activates phonological representations, but visual material requires recoding before doing so), SEM can explain the interactions with articulatory suppression and presentation modality in the same fashion.[9]

One problem faced by the working memory theory however is that some recall of visual material remains possible under articulatory suppression. This cannot be attributed to the phonological store, because recoding of the visual material is prevented. One possibility is to appeal to a second store, such as a visuospatial sketchpad (Baddeley, 1986). SEM does not have to appeal to additional means of storing serial order however. Though the absence of phonological activations impairs recall, items can still be recalled via their positional tokens. A similar account can explain why memory performance remains above chance after several seconds of distractor-filled delay: Though SEM’s phonological activations may have decayed completely, recall can still be supported by contextual and positional cues. This is in contrast with the phonological store, which is unable to support recall when rehearsal is prevented for more than a few seconds (Baddeley, 1986).

In summary, SEM can be extended to much of the data supporting the working memory theory by borrowing some of its assumptions. Furthermore, because SEM does not rely on phonological activations, it can account for short-term memory in situations beyond those explicable by the phonological loop. For example, SEM explains why serial recall, though impoverished, remains possible both under articulatory suppression and after much longer

---

[9] The assumption that irrelevant sound affects positional coding rather than phonological activation does not afford SEM the same explanation of the interaction between articulatory suppression and irrelevant sound with visual material, but interpretation of this interaction is open to debate: see Macken and Jones (1995).

intervals than predicted by the phonological loop. This is attributable to the same long-lasting, nonphonological, positional information that is necessary, for example, to explain protrusions after a filled delay of 20 s between trials (Henson, 1996). By assuming a relationship between positional coding and the rate of change of irrelevant material, SEM may also allow some reconciliation between the working memory and changing state theories of the irrelevant sound effect.

*Interaction with long-term memory.* Other important phenomena in short-term memory concern its interactions with long-term memory. Foremost is the lexicality effect, whereby serial recall of lexical items (e.g., words) is superior to nonlexical items (e.g., nonwords, such as nonsense syllables, or words in an unfamiliar language), even when articulation rate is controlled (Hulme, Roodenrys, Brown & Mercer, 1995). In SEM, long-term memory determines the level at which an ''item'' is defined in short-term memory. For example, each word in a list represents a single item, or chunk (Miller, 1956/1994), whereas each nonword may be better represented as a group of items, where each item is a phoneme. The extra requirement to store both the order of nonwords and the order of their constituent phonemes may contribute to the lexicality effect. The exchange of initial or final phonemes that is common in recall of nonwords (e.g., Treiman & Danis, 1988) would then correspond to interpositions between groups of phonemes. Additional phonotactic constraints clearly play a role however (Hartley & Houghton, 1996), reflecting a general need to relate models like SEM to theories of speech perception and production.[10]

Other influences of long-term memory include the effects of predictability (Baddeley, Conrad & Hull, 1965), semantic similarity (Brooks & Watkins, 1990), word-frequency (Watkins, 1977) and word-likeness (Gathercole & Martin, 1996). The effect of semantic similarity is to allow additional means of organizing items in short-term memory, though such organization is normally secondary to serial organization (Seamon & Chumbley, 1977), and much of the effect may be attributable to guessing strategies (Crowder, 1979). The effect of word frequency might reflect different baseline activations of the categorical or phonological representations in SEM. The effects of predictability and word-likeness are harder to explain. They appear to reflect the number of similar sequences in long-term memory, which is be-

---

[10] For example, SEM's treatment of phonological information is clearly a simplification. Phonological similarity entails more than the simple metric $p$ in Appendix B: It is a function of syllable structure and distinctive phonemic features (Ellis, 1980). Confusions involve the movements of consonants rather than vowels, particularly onsets, and these movements respect position within syllables, so that onsets are only likely to swap with other onsets, to form new syllables (Treiman & Danis, 1988). Moreover, other aspects of phonological similarity, such as its interaction with the modality effect (Murray, 1967; Watkins, Watkins & Crowder, 1974) and grouping (Frick, 1989) require further elaboration of SEM.

yond the current scope of SEM. These more subtle interactions between short- and long-term memory pose problems for most models of serial recall.

In addition to the influence of long-term memory on short-term memory, there is the question of transfer from short- to long-term memory. SEM does not model such long-term learning. Given the episodic nature of SEM's storage, there is no incremental effect of learning the same sequence again and again. In the absence of rehearsal, a long-enough retention interval (i.e., enough contextual change) will cause complete forgetting of sequences. Nonetheless, there is evidence to suggest that such forgetting is typical of short-term memory, and a second process is responsible for long-term learning. For example, the Hebb effect, whereby a list repeated every few trials shows improved recall with each repetition, does not arise simply with repeated presentations, even with vocalization (Cunningham, Healy & Williams, 1984). The effect is contingent on active rehearsal or recall (Kidd & Greenwald, 1988). Though maintenance rehearsal alone can offset forgetting, it does not improve recall (Reitman, 1974; Healy, Fendrich, Cunningham and Till, 1987). Thus SEM's rehearsal process is appropriate for maintenance rehearsal, and, without further active rehearsal, forgetting from short-term memory is consistent with that predicted by SEM.

Unfortunately however, there is little evidence to discern the nature of active rehearsal and long-term learning. It clearly involves the process of chunking subsequences of a repeated list (e.g., Bower & Springston, 1970; Martin, 1974). There may be a role for the strengthening of position-item associations (Burgess & Hitch, 1996; McNicol, 1978), but such an approach faces considerable interference as soon as several sequences of the same items are learned (Henson, 1996). One solution is that positional associations are made to a different pair of start and end markers for each sequence learned (Houghton, 1990). Alternatively, long-term learning may involve a different means of storing serial order. The extension of primacy-gradient ideas (Grossberg, 1978; Page, 1994) would appear to be a promising approach. Since the interest in serial learning has waned (Slamecka, 1985), further data are required to constrain models of this fundamental aspect of human cognition.

In summary, the study of short-term memory suggests that sequences are initially stored by positional codes, but that these codes soon become ineffective in the absence of maintenance rehearsal. Transfer to long-term memory may involve a secondary process of active rehearsal and chunking of these sequences.

*Spatial recall.* Thus far, the data on short-term memory have been confined to the temporal dimension, where serial recall implies recall of temporal order (temporal recall). SEM may also be applicable to recall of spatial order (spatial recall). Like temporal position, spatial position appears to be coded relative to landmarks (Nelson & Chaiklin, 1980) and there is evidence for comparable positional uncertainty functions associated with spatial position

(Hitch, 1974; Nairne & Dutta, 1992). Initial studies of temporal and spatial recall concluded that the two dimensions are independent (Hitch & Morton, 1975; Mandler & Anderson, 1971), suggesting that spatial position might be encoded together with temporal position in SEM's tokens, and one or other cued independently. However, more recent research reveals the relation between spatial and temporal information to be more complex.

Healy (1977) for example reported that spatial recall showed effects of temporal as well as spatial position (a similar result was reported for spatial-probed recall by Murdock, 1969). However, Healy failed to find the phonological confusions in spatial recall that typify temporal recall. This suggests a better distinction is between spatiotemporal and phonological coding: Only spatiotemporal coding applies to spatial recall, whereas both spatiotemporal and phonological coding apply to temporal recall. Furthermore, phonological coding serves mainly to improve item recall (Healy, Cunningham, Gesi, Till & Bourne, 1991), consistent with role of phonological activations in SEM. The more fundamental role of spatiotemporal coding is supported by similarities between temporal recall of verbal items and temporal recall of spatial locations (Jones, Farrand, Stuart & Morris, 1995; Smyth & Scholey, 1996). Nonetheless, the exact nature of the spatiotemporal codes remains unclear (Healy, 1982) and further research is needed.

*Free recall.* There is a large literature on free recall, which exceeds the present remit. However, in relation to SEM and the problem of serial order, two points are worth noting. First, with free recall instructions, actual recall order depends on list length. For lists up to span-length, people will normally default to serial recall (Corballis, 1967). For longer lists, the first and last few items are often recalled first, followed by the middle items (Bjork & Whitten, 1974). In SEM, positional codes are sufficient to support serial recall of short lists, but not long lists, where the codes for middle positions become indistinguishable (Demonstration 2). The ability to distinguish positions reliably may therefore underlie the transition between serial and nonserial recall as list length increases. Second, even when middle positions cannot be distinguished, recall of the first few items may still be mediated by the start marker, and recall of the last few items mediated by the end marker (or by residual phonological activations). Middle items can only be weakly cued by the overlap in general context, and so will not be recalled well, producing bowed serial position curves. Indeed, the assumption of contextual overlap makes SEM compatible with theories that explain primacy and recency in free recall in terms of contextual distinctiveness (e.g., Glenberg & Swanson, 1986; Greene, 1986). Thus, some of SEM's assumptions are applicable to free recall as well as serial recall, if only at the hand-waving level.

*Probed recall.* In the case of item-probed successor recall (i.e., recall of the item that followed a probe, e.g., Murdock, 1968), SEM, possessing no item-item associations, must appeal to indirect processes such as covert serial recall. Nonetheless, this is what the latency data suggest (Sternberg, 1967).

In the case of item-probed position recall (e.g., Jahnke, Davis & Bower, 1989; McNicol, 1975), the probe item may be used to cue the positional code of SEM's corresponding token (i.e., the reverse process to that in serial recall). The case of position-probed item recall is less clear, because a position probe has no necessary relation to the internal positional codes used by people. With a numerical probe for example (e.g., Nairne, Whiteman & Woessner, 1995), an additional translation process will be required to convert the probe into start and end marker values in SEM. With a spatial position probe, there is evidence for more direct access to internal positional codes, at least for the first and last item of a list (Sanders & Willsemsen, 1978) or group (Hendrikx, 1984). Furthermore, though positional codes can be reinstated directly, it may often be easier to reinstate codes for only the first and last positions directly, and reinstate the rest serially (Sanders & Willsemsen, 1978). The problem with spatial positional probing however is to specify how a spatial probe is used to access memory for temporal order (see above).

In item recognition tasks (where the task is simply to state whether the probe item was somewhere in the list), latency data were originally taken to support serial search (Sternberg, 1969). More recent data however demonstrate a recency effect that is better explained by direct access (McElree & Dosher, 1989). In contrast to other probing techniques in SEM, the item recognition task could be achieved simply by checking the activation of the phonological representation of the probe item. This would produce direct access and a recency effect (though the complete story is likely to be more complex, e.g., Monsell, 1978).

*Backward recall.* The difficulty in reinstating positional codes in any order is supported by data on backward recall (e.g., Madigan, 1971). This task is normally harder than forward recall, though once item information is equated, the difference can disappear (Farrand & Jones, 1996). The latter authors argue that their data imply a single process underlying forward and backward recall, though others argue for different processes (Li & Lewandowsky, 1995). These discrepancies may reflect strategic differences in the way people attempt backward recall, the most common strategy depending on procedural details (e.g., whether recall is immediate or delayed, or whether there are intralist distractors, as in Li & Lewandowsky's experiments).

Clearer evidence on backward recall comes from latency measures. Longer latencies in immediate backward than forward recall (Anders & Lillyquist, 1971) suggest that backward recall may involve successive forward searches, reporting the last item after each search (Page & Norris, in press). This would imply that positional codes can only be reinstated in a forward order, from the first through to the last. However, closer inspection of the latency data reveals that participants often group the lists subjectively, and that the extra time in backwards recall comes from the retrieval of groups: There is little extra time required to reverse items within groups (Anders &

Lillyquist, 1971). This suggests that some direct reinstatement of positional codes is possible, at least within groups. Thus the evidence from backward recall, much like that from probed recall, suggests a combination of covert serial search and direct access via positional codes, which is not necessarily problematic for SEM.

*Other tasks.* In the running span task (Pollack, Johnson & Knaff, 1959), people are presented with a long list of items and attempt to recall as many of the most recent items in order as possible. Though lower than conventional spans, running memory spans are at least 3–4 items. Prima facie, this task would appear difficult to model in SEM, since the start of the sequence is undefined. However, there is no reason why people cannot impose their own subjective starts and ends of subsequences, and use these to define position. In other words, they may continually update the start of a group of items they intend to remember. Indeed, such spontaneous grouping is apparent (Pollack et al., 1959). By assuming a subjective start marker, determined by one's rehearsal strategy, the running memory task may be reconciled with SEM.

In the case of recognition of serial order, Jahnke et al. (1989) showed that recognition was poorer for foil sequences differing from the target sequence by an adjacent transposition than for those differing by a remote transposition. These authors fitted their recognition data using the positional uncertainty functions generated by a second task of item-probed position recall. These functions resemble those generated by SEM. A similar result was found by Ratcliff (1981) in perceptual matching of spatial sequences: Performance was better for foil sequences in which items were transposed longer distances. Ratcliff fitted his accuracy and reaction time data using positional uncertainty functions produced by the Perturbation Model (Lee & Estes, 1978). However, positional uncertainty in this model results from perturbations over time, and the same data may be fitted equally well by SEM, in which ''spatial perturbations'' would result directly from the positional uncertainty entailed by markers at the left and right of the sequence. Thus both recognition of temporal sequences and perceptual matching of spatial sequences provide data consistent with the positional coding of SEM.

### *Comparison with Other Theories*

SEM shares many assumptions with previous models. Of its three core assumptions, SEM's positional coding is based on the work of Houghton (1990), its separate storage of tokens is based on multiple-trace theory (Hintzman, 1986), and its processes of response suppression and phonological retrieval are based on the Primacy Model (Page & Norris, in press). Of SEM's further assumptions, the coding of position at multiple levels is based on the work of Lee and Estes (1981), though their notion of trial-level codes differs from SEM's notion of general context. Indeed, SEM's distinction between reinstateable (positional) and non-reinstateable (general) contexts is more akin to the ideas of Hintzman, Block and Summers (1973). The assumption of maintenance rehearsal and rapidly-decaying phonological activations is based on the phonological loop (Baddeley, 1986), as described above.

Forgetting in SEM is both interference-based (Keppel & Underwood, 1962; Melton, 1963), in the retrieval of tokens, and decay-based (Baddeley & Scott, 1971; Conrad, 1967; Reitman, 1974), in the retrieval of phonological forms. Decay occurs during storage, and both proactive and retroactive interference occur during retrieval, from competition between items in similar positions (i.e., an overload of start and end cues, Sanders, 1975). As regards the modal model (Healy & McNamara, 1996), SEM's phonological activations resemble the short-term store of Atkinson and Shiffrin (1968), the decay of which explains the rapid forgetting over the first few seconds of retention. SEM's episodic tokens are longer-lived, subserving memory after longer periods of distraction (though one might not call this a long-term store, in the sense of permanent storage). More generally though, SEM is an example of contextual distinctiveness theories, which emphasize similar principles applying to both short- and long-term memory (Crowder, 1993; Neath, 1993).

However, SEM also differs from previous models in important ways. Firstly, in relation to models of distinctiveness (e.g., Bower, 1971; Johnson, 1991; Murdock, 1960, 1974; Neath, 1993; Neath & Crowder, 1990), SEM is more precise about the nature of positional information. The model of Johnson (1991), for example, assumes that serial position is represented along a single dimension, much like a physical property (e.g., weight). Expressing magnitudes on this dimension in relation to others allows a parameter-free estimation of positional overlap or distinctiveness. Though appealing, given the several parameters SEM uses to characterize positional codes, these models have only been shown to reproduce general phenomena such as primacy and recency. It is not clear that they can provide fits to detailed error patterns in tests of short-term memory. More importantly, these models are descriptive models rather than process models (with the exception perhaps of Bower, 1971). In other words, they only characterize the long-run statistics of recall, and cannot produce an example recall protocol in the way SEM can.

The Perturbation Model (Lee & Estes, 1978, 1981) is better specified than distinctiveness models, and captures positional uncertainty with a single parameter, the perturbation rate. However, the nature of the positional codes remains unclear. For example, the number of positional codes would appear to be unrestricted, providing no rationale for the limited resolution of positional coding (Demonstration 2). The Perturbation Model is also another descriptive model that does not fully simulate the recall process (Page & Norris, in press; see Nairne & Neath, 1994, and Mewhort, Popham & James, 1994, for a similar criticism of TODAM). For example, by assuming that items within a sequence perturb independently, the Perturbation Model predicts impossible situations where more than one item is stored at the same position and provides no mechanism for repetitions (Demonstration 1). Moreover, by assuming that items perturb independently between sequences, it cannot explain the small dependencies found in the data (Nairne, 1991; see Henson, 1996, for a SEM's account of these dependencies). Finally, its assumption that omissions arise when items perturb ''out of the trial'' (Lee & Estes, 1981) is incompatible with the distribution of omissions (Demonstration 1).

The ability of TODAM (Lewandowsky & Murdock, 1989) and its various extensions (Murdock, 1993, 1995) to model serial recall from short-term memory has already been questioned in the discussion of chaining models (though this does not necessarily detract from its application to other aspects of memory). It is unable to fit much of the data fitted by SEM, and, possessing only interitem associations, cannot explain positional errors (Henson, 1996).

Another model is the Feature Model (Nairne, 1990). The strength of this model is its parsimonious account of recency, modality and suffix effects. However, the Feature Model has no explicit representation of serial order, and therefore fails to produce the fundamental locality constraint on transpositions. As Nairne admits: ''This is clearly an unattractive feature of the model because it predicts fewer transposition errors for adjacent items than for remote items in the list; this prediction is counter to the data.'' (Nairne, 1990; p. 257). Nonetheless, SEM and the Feature Model are not incompatible, and it may be fruitful to combine ideas from both. A similar model is the Attribute Model of Drewnowski (1980), which also captures several aspects of short-term memory, including effects of list length and phonological similarity. In this model, multiple attributes of items are coded, such as identity, position, auditory features and interitem relations, and during recall, these attributes are addressed in a predetermined order of priority. Though appealing however, these ideas appear to have little explanatory power. For example, effects of list length are a simple consequence of ''item load'' in memory. Moreover, its assumption of only four positional codes appears incorrect (Henson, 1996), and, like the Feature Model, it does not produce appropriate transposition gradients (Drewnowski, 1980).

The Primacy Model (Page & Norris, in press) is one of the few models to produce quantitative fits to error distributions in immediate serial recall. This model is appealing in its simplicity, and provides a unified account of effects of articulation rate, rehearsal and delay. However, the Primacy Model is yet to be extended to grouping and intertrial effects. More importantly, as an ordinal model, in which order is stored in a primacy gradient (Figure 1B), it cannot explain positional errors. One possibility would be to combine the two models by incorporating a primacy gradient into SEM's phonological activations. This would provide an ordered phonological store that would enhance immediate serial recall, but not recall after a short delay, which would rely on SEM's positional information.

Of all current models however, SEM is most similar to the Articulatory Loop Model of Burgess and Hitch (1992). This model and its revision (Burgess & Hitch, 1996) give reasonable qualitative fits to error data, such as transpositions, omissions, and phonological confusions. It can also provide a qualitative fit to positional errors such as interpositions and protrusions (Burgess & Hitch, 1996), though not to the same level of detail as SEM. Nonetheless, there remains an important difference between SEM and the Articulatory Loop Model. This reflects the nature of the positional codes.

### *Predictions of SEM*

The moving context window assumed by Burgess and Hitch (Fig. 1C) codes absolute position (e.g., first, second, third, etc.), regardless of list length. Indeed, the coding of absolute position would appear to be an assumption of all other positional models of short-term memory (e.g., Anderson & Matessa, 1997; Brown et al., in press; Lee & Estes, 1981). SEM on the other hand codes position relative to both the start and the end of a sequence; a coding that is sensitive to list length.

To test whether position is coded in absolute or relative terms, Henson (1997) examined the positional errors between sequences of different lengths. For example, when lists are grouped as a group of three followed by a group of four, SEM predicts that interpositions will occur between the ends of groups (i.e., between Positions 3 and 7). Other positional models (e.g., Brown et al., in press; Burgess & Hitch, 1996) predict that interpositions will occur between the same absolute position within groups (e.g., between Positions 3 and 6). The data showed that the incidence of transpositions between the ends of groups was higher than that between the third position of groups, favoring SEM's coding of relative rather than absolute position. This prediction was confirmed in a second experiment examining the nature of protrusions between trials of different length. With lists of five, six or seven items, the incidence of protrusions between the ends of reports was greater than the incidence of protrusions between the fifth position of reports. Again, this supports SEM's hypothesis that position is coded relative to the start and the end of a sequence, but is incompatible with other positional models (e.g., Brown et al., in press; Burgess & Hitch, 1996).

### *Problems for SEM*

The main problem for SEM is to specify more precisely the psychological correlates of its start and end markers. This would allow SEM to predict further experimental manipulations that should affect short-term memory for serial order.

One question concerns how the influence of the end marker can extend backwards in time in the coding of temporal order. It was originally proposed that the strength of the end marker might correspond to the degree of expectation for the end of a sequence. This is a reasonable assumption in most experimental situations, where the length of the sequence is known in advance. For example, most experiments use a fixed number of items each trial, and in the case of grouping, the size of the groups is either known advance, or, in the case of subjective grouping, decided by the participant. However, in other situations the length of a sequence is not known in advance. For example, in the second experiment by Henson (1997), the length of the list on each trial was unpredictable. This makes an expectancy interpretation harder to uphold. (Alternative interpretations, including a modified notion of expectancy, are given in Henson, 1997.) Nonetheless, the predictability of the end of a sequence is at least one experimental manipulation that is likely to affect the behavior of SEM's end marker.

Another question concerns whether the start and end marker strengths change only with position, or whether they change along some other correlated dimension. In the temporal domain for example, the strength of the markers might be a function of time, which would make SEM compatible with temporal distinctiveness theories (e.g., Glenberg & Swanson, 1986; Neath & Crowder, 1990, 1996). In the spatial domain, the nature of start and end marker strengths might be tested by varying the physical separation between the items in a sequence, to see whether this manipulation affects recall or recognition of spatial order. Experiments like this will help constrain interpretations of the start and end markers in SEM.

## CONCLUSION

This article introduced a new computational model of short-term memory for serial order, the Start-End Model (SEM). The core assumptions of SEM are: 1) the position of an item in a sequence is coded relative to the start and end of that sequence, 2) items are stored in memory as position-sensitive tokens, and 3) items are retrieved by reinstating positional codes for each response, and letting tokens compete in parallel for output. Additional assumptions that not all context is reinstateable at recall and that response selection is supplemented by transient phonological information, allows SEM to reproduce the main findings in short-term memory research. Furthermore, SEM is readily extendible to other phenomena in serial recall, such as the effects of articulation rate, irrelevant sound and articulatory suppression, and other tasks, such as spatial, probed and free recall. Though SEM shares several assumptions with previous models, it is unique in reproducing the complete pattern of errors in serial recall. Most importantly, it is unique in predicting that positional errors between sequences of different length will maintain relative rather than absolute position. The main challenge for future research is to constrain further the nature of SEM's start and end markers by identifying additional factors that affect the coding of position in short-term memory.

## APPENDIX A

### *Scoring Serial Recall*

One of the first measurements of serial recall was the proportion of lists recalled correctly (e.g., Crannell & Parrish, 1957). This measure underlies the span index of short-term memory: One's memory span is usually defined as the length of list that one can recall correctly 50% of the time. However, this measure ignores differences in the recall of each item in a list. Murdock (1968) pioneered the use of serial position curves, which plot the proportion of errors at each position of a list. These curves are typically bowed, with an advantage in recall for the first and last few items (the *primacy* and *recency* effects respectively). In immediate serial recall, primacy is normally more pronounced than recency.

Later work distinguished two main types of error: order errors and item errors (e.g., Estes, 1972). Order errors are list items recalled in the wrong position; item errors are list items not recalled anywhere in the report. However, few studies actually go beyond measuring the proportion of lists correct or plotting serial position curves, let alone analyzing the distribution of item and order errors. A major theme behind this article is that a great deal more information is available through analyzing error patterns in detail. As Estes (1972) observed: ''When retention is imperfect, the confusion errors that occur are highly systematic'' (p. 161).

### *Classification of Errors*

The classification of errors used in the present article is described below, with examples given in Table A1. This classification distinguishes an item's position in a list, its *input position,* from its position in a participant's report

TABLE A1
Example Errors in Serial Recall of a List RMHQVJ

| Error type | List (input postions) | | | | | | Report (output positions) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Omissions | R | M | H | Q | V | J | R | M | **K** | **V** | — | — |
| Transpositions | R | M | H | Q | V | J | R | M | H | **V** | **Q** | J |
| Intrusions | R | M | H | Q | V | J | R | M | **K** | **V** | **Y** | J |
| Confusions | R | M | H | Q | V | J | R | M | H | Q | **P** | **K** |
| Repetitions | R | M | H | Q | V | J | R | M | H | **R** | V | **M** |
| Interpositions | R | M | H | Q | V | J | R | M | **J** | Q | V | **H** |
| Protrusions | F | P | Y | K | Z | W | F | P | Y | **Z** | **K** | W |
| | R | M | H | Q | V | J | **P** | M | **Y** | Q | **K** | J |

*Note:* Errors are in bold; items corresponding to a particular error type are underlined.

of that list, its *output position.* Though the complete classification of errors might appear somewhat complex, each type of error plays an important role in constraining models of serial recall (Henson, 1996).

When scoring by output position (i.e., taking each response in a participant's report), errors can be broadly categorized into *substitutions* and *omissions.* Substitutions arise when an incorrect item is given; omissions arise when no item is given (participants are often encouraged to guess when unsure, but can omit if no item comes to mind). Substitutions may be either *transpositions* or *intrusions.* Transpositions are list items in the wrong position; intrusions are items that were not present in the list. Intrusions may be items outside the experimental vocabulary (the set of items from which the lists are constructed), but most often they are items appearing on previous trials (Demonstration 5).

A special class of substitutions are phonological *confusions.* These are incorrect items that are phonologically similar to the correct item, and are common in tests of immediate memory (Demonstration 6). A special class of transpositions are *repetitions.* Repetitions are items that occur more than once in a report, even though they only occurred once in the list. The distribution of repetitions sheds light on the retrieval processes underlying serial recall (Demonstration 1).

Two further types of *positional errors* can be identified. *Interpositions* arise when lists are grouped (e.g., by a pause between presentation of every third item). They are transpositions between groups that maintain their position within a group (Demonstration 3). *Protrusions* are similar errors, but maintain position between trials rather than between groups (Demonstration 5). Two types of protrusions can be distinguished: *input protrusions* are items occurring at the same position in the previous list, whereas *output protrusions* are items occurring at the same position in the previous report (which may or may not be correct). Note that the definition of protrusions is orthogonal to that of transpositions and intrusions, in that a protrusion may be either an intrusion or a transposition with respect to the current trial (which is why the term is preferable to Conrad's ''serial order intrusions''; Conrad, 1960).

Additional information is provided by scoring against input position (under which categorization of errors is similar, and usually self-evident; Table A1). For example, when omissions are scored against input position, they represent items that are not recalled anywhere in a report. This distribution of omissions can differ from that plotted against output position (Demonstration 1). The distinction between item and order errors (e.g., Healy, 1974) also refers to input position, though this distinction is rarely employed in the present article.

Transpositions can be scored against both input and output position, producing *transposition gradients* (Demonstration 1). Errors can also be scored against both the output position of the current trial and either the input or output position of the previous trial, producing *intrusion gradients* (Demonstration 5). Note that not all of the errors in intrusion gradients are necessarily intrusions in the strict definition of the term; some may also be transpositions with respect to the current trial. Indeed, when errors involve items occurring in two successive lists, it is not always possible to determine the source of those errors (i.e., whether they reflect interference within or between trials). Nonetheless, assuming there is no systematic relationship between the lists on successive trials, it is still possible to compare intrusion gradients with those expected by chance (normally flat).

## APPENDIX B

There are two versions of SEM: a single-trial version and a multiple-trial version. In the single-trial version, the tokens in short-term memory are restricted to those from the most recent list. The multiple-trial version is more general, including tokens from previous lists and reports, together with the notions of general context, rehearsal and phonological decay. These versions are formalized below.

### Single-Trial Version of SEM

The single-trial version takes a single list of items, and simulates $N_L$ independent attempts at serial recall of that list. Each attempt can be split into presentation and recall.

#### *Presentation*

A token is created for the item at each position $p = 1..N_P$ of the list. Specifically, for each group $g = 1..N_G$ within the list and for each item in group $g$, $i = 1..N_I$(g), a token $t$ is created with positional codes $\boldsymbol{p}_I^{(t)}$ and $\boldsymbol{p}_G^{(t)}$. $N_G$ is determined by the temporal or spatial grouping of the presentation of items, or simply by instruction (for ungrouped lists, $N_G = 1$). The vector $\boldsymbol{p}_I^{(t)} = (s_I(i), e_I(i))$ codes the position of item $i$ within group $g$, where $s_I(i)$ and $e_I(i)$ are the strengths of markers for the start and end of that group:

$$s_I(i) = S_{O,I}(i) S_I^{i-1} \qquad e_I(i) = E_{0,I}\, E_I^{N_I(g)-i} \tag{B1}$$

$S_{O,I}$ and $S_I$ are parameters reflecting the initial strength of the start marker and the change in its strength over position, and $E_{O,I}$ and $E_I$ are parameters reflecting the initial strength of the end marker and the change in its strength over position.

The vector $\boldsymbol{p}_G^{(t)} = (s_G(g), e_G(g))$ codes the position of group $g$ in the list, where $s_G(g)$ and $e_G(g)$ are the strengths of markers for the start and end of that list:

$$s_G(g) = S_{0,G} S_G^{g-1} \qquad e_G(g) = E_{0,G}\, E_G^{N_G-g} \tag{B2}$$

$S_{0,G,}\, S_G$, $E_{O,G}$ and $E_G$ are parameters defined similarly to those for the item-level above.

*Recall*

For each response $j = 1..N_P$, retrieval of an item can be divided into four stages:

*Stage 1: Cueing.* A cue for response $j$ is generated with positional codes $\boldsymbol{p}_I^{(j)}$ and $\boldsymbol{p}_G^{(j)}$, as defined in Equations B1 and B2. These codes are matched against the positional codes $\boldsymbol{p}_I^{(t)}$ and $\boldsymbol{p}_G^{(t)}$, of the $t = 1..N_T (= N_P)$ tokens in short-term memory, cueing each with strength $q^{(t)}(j)$:

$$q^{(t)}(j) = o(\boldsymbol{p}_I^{(t)}, \boldsymbol{p}_I^{(j)}),\ o(\boldsymbol{p}_G^{(t)}, \boldsymbol{p}_G^{(j)}) \tag{B3}$$

where $o(\boldsymbol{p}^{(t)}, \boldsymbol{p}^{(j)})$ is the overlap between positional codes, defined as:

$$o(\boldsymbol{p}^{(t)}, \boldsymbol{p}^{(j)} = (\boldsymbol{p}^{(t)} \cdot \boldsymbol{p}^{(j)})^{1/2} \exp\left\{-\left(\sum_k (p_k^{(t)} - p_k^{(j)})^2\right)^{1/2}\right\}$$

where the summand $k$ is over the (two) components of the vectors $\boldsymbol{p}^{(t)}$ and $\boldsymbol{p}^{(j)}$.

Note that SEM is readily extendible to any number of subgroupings of a sequence by assuming that each boundary between groups can be marked by start and end markers. For example, with $m = 1 \ldots L$ levels of grouping and positional codes given by $p_1, p_2, \ldots p_L$, the strength with which token $t$ is cued for response $j$ is simply:

$$q^{(t)}(j) = \mathrm{II}_m o(p_m^{(j)}, p_m^{(t)})$$

*Stage 2: Categorical selection.* Items compete for selection with a strength proportional to their most strongly cued token. Specifically, the categorical (type) representations of all items $u = 1..N_V$ in the experimental vocabulary compete with strength $c_c^{(u)}$, where:

$$c_c^{(u)} = \max\{q^{(t)}|_{(i(t)=u)}\}\ (1 - r^{(u)}) + n_C \tag{B4}$$

where $i(t)$ is the identity of (the item corresponding to) token $t$, $r^{(u)}$ is the suppression of item $u$, and $n_C$ is random noise drawn from a zero-mean Gaussian distribution with standard deviation $G_C$ for each item $u$. The strongest item $v$ is passed to Stage 3.

*Stage 3: Phonological selection.* The item $v$ selected from Stage 2 is matched against a set of phonological representations in order to output a response. Specifically, a second competition is held in which each item competes with strength, $c_P^{(u)}$:

$$c_P^{(u)} = c_c^{(u)} + p(u, v)a_P^{(v)}(1 - r^{(u)}) + n_P \tag{B5}$$

where $p(u, v)$ is the phonological similarity between items $u$ and $v$, $a_P^{(v)}$ is the activation of the phonological representation of item $v$, $r^{(u)}$ is the suppression of item $u$, and $n_p$ is random noise drawn from a zero-mean Gaussian distribution with standard deviation $G_P$ for each item $u$. The strongest item $w$ is passed to Stage 4.

The value of $p(u, v)$ is such that $p(u, v) = 1$ if $u = v$, $p(u, v) = \mathrm{P}_s$ if items $u$ and $v$ are phonologically similar (i.e., confusable, such as the letters $B$ and $G$), and $p(u, v) = \mathrm{P}_D$ if they are dissimilar (i.e., if one is nonconfusable, such as the letters $B$ and $H$). The value of $a_P^{(v)}$ is such that $a_P^{(v)} = A_P$ if item $v$ was in the most recent list, and $a_P^{(v)} = 0$ otherwise.

Note that this stage was effectively bypassed in Demonstrations 1 to 5 by setting $A_P = G_P = 0$. For lists of nonconfusable items, this does not affect the predictions. In Demonstration 6, the stage was simplified by fixing the values, $A_P = 1$, $P_s = 1$ and $P_D = 0$.

*Stage 4: Thresholding and suppression.* If the strength of the item $w$ is below an output threshold, $T_O$, such that $c_P^{(w)} < T_O$, then no item is recalled and an omission is indicated instead. Otherwise, item $w$ is output as response $j$, and suppressed such that $r^{(w)} = 1$. Meanwhile, the suppression of all other items $u$ wears off according to the update rule:

$$r^{(u)} \rightarrow r^{(u)} \exp(-R_s) \tag{B6}$$

where $R_s$ is the rate of decay of suppression (i.e., the decay is discretized over responses for simplicity).

Stages 1 to 5 are then repeated for response $j + 1$.

## Multiple-Trial Version of SEM

In this version of SEM, short-term memory is assumed to contain tokens from previous trials as well as the most recent trial (i.e., $N_T > N_P$). These tokens include a new component which represents general (nonpositional) context, which cannot be reinstated at recall. Each item recalled is also recoded as a new token (coded in its recall position, regardless of whether that is correct), and every presentation and rehearsal of an item activates its phonological representation. Finally, these activations decay over time, reflecting the transient nature of phonological information in short-term memory.

The multiple-trial version takes $N_L$ different lists and recalls each one once. Recall of each list $q = 1..N_L$, with positions $p = 1..N_P(l)$, can be split into presentation, retention, recall and intertrial intervals. Only the differences between the multiple-trial version and the single-trial version are formalized below.

### *Presentation*

Each token $t$ has three components $\boldsymbol{p}_I^{(t)}$, $\boldsymbol{p}_G^{(t)}$ and $\boldsymbol{p}_C^{(t)}$, where $\boldsymbol{p}_I^{(t)}$ and $\boldsymbol{p}_G^{(t)}$ are the positional codes defined in Equations B1 and B2, and $\boldsymbol{p}_C^{(t)}$ is a one-dimensional vector representing the general context when token $t$ was created. For mathematical convenience, the current context is represented by the value 1, such that each token is created with $\boldsymbol{p}_C^{(t)} = (1.00)$, and during subsequent contextual changes (e.g., during the presentation of other items), the general context of all tokens is updated according to:

$$\boldsymbol{p}_C^{(t)} \rightarrow \boldsymbol{p}_C^{(t)} E_C \tag{B7}$$

where $E_C < 1$ represents the rate of contextual change, discretized over units of ‘‘episodes’’. The number of episodes between presentation of each item is given by the parameter $C_P$.

Presentation of an item $v$ also activates its phonological representation such that $a_P^{(v)} = A_P$, while the activation of the other $u = 1..N_v$ phonological representations decays during each episode as follows:

$$a_P^{(u)} \rightarrow a_P^{(u)} \exp(-R_P) \tag{B8}$$

where $R_P$ is the rate of decay of phonological activations.

### *Retention Interval*

During the retention interval, the general context of all tokens is updated according to Equation B7, and the phonological activations of items decay according to Equation B8, for each of the $C_D$ episodes during the delay before recall.

### *Recall*

For each response $j = 1..N_P(1)$, a cue is generated with positional codes $\boldsymbol{p}_I^{(j)}$, $\boldsymbol{p}_G^{(j)}$ and general context $\boldsymbol{p}_C^{(j)}$, where $\boldsymbol{p}_C^{(j)}$ is always the current context *(1.00).*

The multiple-trial and single-trial versions differ in Stage 1 and Stage 4 of recall:

*Stage 1: Cueing.* The positional and general contexts of the cue are matched against those of the $t = 1..N_T$ tokens in short-term memory, cueing each with strength $q^{(t)}(j)$:

$$q^{(t)}(j) = o(p_I^{(t)}, p_I^{(j)}) \cdot o(p_G^{(t)}, p_G^{(j)}) \cdot o(p_C^{(t)}, p_C^{(j)}) \tag{B9}$$

Note that, in theory, SEM does not ncessarily assume a limit on the number of tokens in short-term memory. In practice however, only the most recent tokens can ever be retrieved, assuming continual context drift. The computer implementation of SEM therefore stores only the last $N_T$ tokens, such that the oldest token is overwritten by the newest. With typical parameter values, a storage limit of $N_T > 4N_P$ has negligible effects on SEM's predictions. Demonstrations reported here used $N_T = 30$.

*Stage 4: Thresholding and suppression.* The item $w$ selected in Stage 4 is output as response $j$, as before. In addition however, it is recoded as a new token in memory, with positional and general context given by $\boldsymbol{p}_I^{(i)}$, $\boldsymbol{p}_G^{(i)}$ and $\boldsymbol{p}_C^{(i)}$ (i.e., those of the cue for response $i$), and its phonological representation is reactivated, such that $a_P^{(w)} = A_P$. Finally, the general context of all tokens is updated according to Equation B7, and the phonological activations decay according to Equation B8, for each of the $C_R$ episodes between recall of each item.

### Intertrial interval

During the intertrial interval, the general context of all tokens is updated according to Equation B7 for $C_A + C_I$ contextual changes. The parameter $C_A$ reflects a baseline change in context between trials, given the discrete fashion with which trials are typically presented (see Footnote 6). The parameter $C_I$ represents the additional delay between trials, owing, for example, to the presence of a distraction task. The activations of phonological representations also decay according to Equation B8 for each of the $C_I$ episodes.

Note that, in general, the decay of phonological activations (and the decay of suppression) might be uncoupled from contextual change by further parameterizing presentation rates, recall rates, etc. in terms of time (introducing new parameters in addition to $C_P$, $C_D$, $C_R$ and $C_I$). This is beyond the scope of the present model.

### Model Parameters

The full, multiple-trial version of SEM has many parameters. However, the majority of these are determined by the experimental design (e.g., $N_P$, $N_G$, $N_V$) or given a fixed value (e.g., $S_{O,I}$, $S_I$, $S_{O,G}$, $S_G$). The values of parameters in Demonstrations 1 to 6 that were fitted to the data are given in Table B1

TABLE B1
Parameter Values Used in Demonstrations 1–6

| Demo | $F_{O,I}$ | $F_I$ | $F_{O,G}$ | $F_G$ | $G_C$ | $G_P$ | $R_S$ | $R_P$ | $T_O$ | $E_C$ | $C_P$ | $C_D$ | $C_R$ | $C_I$ | $C_A$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. | 0.60 | 0.60 | - | - | 0.08 | - | 0.50 | - | 0.35 | - | - | - | - | - | - |
| 2. | 0.60 | 0.60 | - | - | 0.08 | - | 0.50 | - | 0.35 | - | - | - | - | - | - |
| 3. | 0.40 | 0.60 | 0.00 | 0.60 | 0.14 | - | 0.50 | - | 0.35 | - | - | - | - | - | - |
| 4. | * | 0.60 | * | 0.60 | 0.14 | - | 0.50 | - | 0.35 | - | - | - | - | - | - |
| 5. | 0.60 | 0.60 | - | - | 0.08 | - | 0.50 | - | 0.35 | 0.98 | 1 | 0 | 1 | * | 5 |
| 6. | 0.60 | 0.60 | - | - | 0.06 | * | 0.50 | 0.20 | 0.35 | 0.98 | 1 | * | 1 | 0 | 5 |

*Note:* Hyphens indicate that the parameter was not relevant, or set ot zero; asterics indicate that the parameter was varied within the Demonstration.

(note that $F_{0,I} = E_{0,I}/S_{0,I}$, $F_I = \mathrm{E}_I/S_I$, $F_{O,G} = E_{O,G}/S_{O,G}$, $F_G = E_G/S_G$ where $S_{O,I} = S_{O,G} = 1.00$ and $\mathrm{S}_I = S_G = 0.80$). Some parameters were constant across Demonstrations (e.g., $R_S$, $T_O$). Some of the changes in parameter values across Demonstrations were necessary because of the incremental exposition of SEM (e.g., the increase in $G_C$ with the introduction additional positional codes in Demonstration 3). The remaining changes were necessary to fit different data sets (e.g., changes of $G_P$ in Demonstration 6).

## REFERENCES

Anders, T. R., & Lillyquist, T. D. (1971). Retrieval time in forward and backward recall. *Psychonomic Science,* **22,** 205–206.

Anderson, J. R. & Matessa, M. (1997). A production system theory of serial memory. *Psychological Review,* **104,** 728–748.

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

Baddeley, A. D. (1968). How does acoustic similarity influence short-term memory? *The Quarterly Journal of Experimental Psychology,* **20,** 249–263.

Baddeley, A. D. (1986). *Working memory.* Oxford: Clarendon Press.

Baddeley, A. D., Conrad, R., & Hull, J. (1965). Predictability and immediate memory for consonant sequences. *The Quarterly Journal of Experimental Psychology,* **17,** 175–177.

Baddeley, A. D., & Hull, A. (1979). Prefix and suffix effects: Do they have a common basis? *Journal of Verbal Learning and Verbal Behavior,* **18,** 129–140.

Baddeley, A. D., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology,* **36A,** 233–252.

Baddeley, A. D., & Scott, D. (1971). Short term forgetting in the absence of proactive interference. *The Quarterly Journal of Experimental Psychology,* **23,** 275–283.

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

Bjork, E. L., & Healy, A. F. (1974). Short-term order and item retention. *Journal of Verbal Learning and Verbal Behavior,* **13,** 80–97.

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

Bower, G. H. (1971). Adaptation-level coding of stimuli and serial position effects. In M. H. Appley (Ed.), *Adaptation-level theory* (pp. 175–201). New York: Academic Press.

Bower, G. H., & Springston, F. (1970). Pauses as recording points in letter strings. *Journal of Experimental Psychology,* **83,** 421–430.

Broadbent, D. E., Cooper, P. J., Frankish, C. R., & Broadbent, M. H. P. (1980). Modality differences in relation to grouping in immediate recall. *British Journal of Psychology,* **71,** 475–485.

Brooks, B. M., & Watkins, M. J. (1990). Further evidence for the intricacy of memory span. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **16,** 1134–1141.

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

Brown, G. D. A., Preece, T., & Hulme, C. (in press). Oscillator-based memory for serial order. *Psychological Review.*

Burgess, N., & Hitch, G. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language,* **31,** 429–460.

Burgess, N., & Hitch, G. (1996). *Memory for serial order: A network model of the phonological loop and its timing.* Manuscript submitted for publication.

Campbell, R., & Dodd, B. (1980). Hearing by eye. *Quarterly Journal of Experimental Psychology,* **32,** 85–89.

Conrad, R. (1959). Errors of immediate memory. *British Journal of Psychology,* **50,** 349–359.

Conrad, R. (1960). Serial order intrusions in immediate memory. *British Journal of Psychology,* **51,** 45–48.

Conrad, R. (1965). Order error in immediate recall of sequences. *Journal of Verbal Learning and Verbal Behavior,* **4,** 161–169.

Conrad, R. (1967). Interference or decay over short-retention intervals? *Journal of Verbal Learning and Verbal Behavior,* **6,** 49–54.

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology,* **55,** 429–432.

Conrad, R., & Hull, A. J. (1968). Input modality and the serial position curve in short-term memory. *Psychonomic Science,* **10,** 135–136.

Corballis, M. (1967). Serial order in recognition and recall. *Journal of Experimental Psychology,* **74,** 99–105.

Cowan, N. (1993). Activation, attention and short-term memory. *Memory & Cognition,* **21,** 162–167.

Cowan, N., Day, L., Saults, J. S., Keller, T. A., Johnson, T., & Flores, L. (1992). The role of verbal output time in the effects of word length on immediate memory. *Journal of Memory and Language,* **31,** 1–17.

Crannell, C. W., & Parrish, J. M. (1957). A comparison of immediate memory span for digits, letters and words. *The Journal of Psychology,* **44,** 319–327.

Crowder, R. G. (1967). Prefix effects in immediate memory. *Canadian Journal of Psychology,* **21,** 450–461.

Crowder, R. G. (1978). Mechanisms of auditory backward masking in the stimulus suffix effect. *Psychological Review, 85,* 502–524.

Crowder, R. G. (1979). Similarity and 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.

Crowder, R. G., & Morton, J. (1969). Precategorical acoustic storage. *Perception & Psychophysics,* **5,** 365–373.

Cunningham, T. F., Healy, A. F., & Williams, D. M. (1984). Effects of repetition on short-term retention of order information. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **10,** 575–597.

Drewnowski, A. (1980). Attributes and priorities in short-term recall: a new model of memory span. *Journal of Experimental Psychology: General,* **109,** 208–250.

Drewnowski, A., & Murdock, B. B. (1980). The role of auditory features in memory span for words. *Journal of Experimental Psychology: Human Learning and Memory,* **6,** 319–332.

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology.* New York: Dover.

Ellis, A. (1980). Errors in speech and short-term memory: the effects of phonemic similarity and syllable position. *Journal of Verbal Learning and Verbal Behavior,* **19,** 624–634.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science,* **14,** 179–211.

Estes, W. K. (1972). An associative basis for coding and organisation in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 161–190). Washington, DC: Winston & Sons.

Estes, W. K. (1973). Phonemic coding and rehearsal in short-term memory for letter strings. *Journal of Verbal Learning and Verbal Behavior,* **12,** 360–372.

Estes, W. K. (1991). On types of item coding and sources of recall in short-term memory. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: In honor of Bennet B. Murdock* (pp. 175–194). Hillside, NJ:Erlbaum.

Farrand, P., & Jones, D. (1996). Direction of report in spatial and verbal serial short-term memory. *The Quarterly Journal of Experimental Psychology,* **49A,** 140–158.

Frankish, C. R. (1974). *Organisational factors in short-term memory.* Unpublished doctoral dissertation, University of Cambridge.

Frankish, C. R. (1985). Modality-specific grouping effects in short-term memory. *Journal of Memory and Language,* **24,** 200–209.

Frankish, C. R. (1989). Perceptual organisation and precategorical acoustic storage. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **15,** 469–479.

Frankish, C. R. (1996). Auditory short-term memory and the perception of speech. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 101–128). UK: Psychology Press.

Frankish, C. R., & Turner, J. (1984). Delayed suffix effect at very short delays. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **10,** 767–777.

Frick, R. W. (1988). The role of memory in attenuations of the suffix effect. *Memory & Cognition,* **16,** 15–22.

Frick, R. W. (1989). Explanations of grouping in immediate ordered recall. *Memory & Cognition,* **17,** 551–562.

Frick, R. W., & De Rose, A. (1986). The suffix effect and preattentive unit-formation in visual short-term memory. *Canadian Journal of Psychology,* **40,** 97–108.

Fuchs, A. H. (1969). Recall for order and content of serial word lists in short-term memory. *Journal of Experimental Psychology,* **82,** 14–21.

Gathercole, S. E., & Martin, A. J. (1996). Interactive processes in phonological memory. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 73–100). UK: Psychology Press.

Glenberg, A. M. (1990). Common processes underlie enhanced recency effects for auditory and changing state stimuli. *Memory & Cognition,* **18,** 638–650.

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

Greene, R. L. (1986). Sources of recency effects in free recall. *Psychological Bulletin,* **99,** 221–228.

Greene, R. L. (1989). Immediate serial recall of mixed-modality lists. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **15,** 266–274.

Greene, R. L., & Crowder, R. G. (1988). Memory for serial position: effects of spacing, vocalisation and stimulus suffixes. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **14,** 740–748.

Grossberg, S. (1978). Behavioral contrast in short-term memory: serial binary memory models or parallel continuous memory models? *Journal of Mathematical Psychology,* **17,** 199–219.

Hanley, J. R., & Broadbent, C. (1987). The effect of unattended speech on serial recall following auditory presentation. *British Journal of Psychology,* **78,** 287–297.

Hartley, T., & Houghton, G. (1996). A linguistically constrained model of short-term memory for nonwords. *Journal of Memory and Language,* **35,** 1–31.

Healy, A. F. (1974). Separating item from order information in short-term memory. *Journal of Verbal Learning and Verbal Behavior,* **13,** 644–655.

Healy, A. F. (1977). Pattern coding of spatial order information in short-term memory. *Journal of Verbal Learning and Verbal Behavior,* **16,** 419–437.

Healy, A. F. (1982). Short-term memory for order information. In G. H. Bower (Ed.), *The psychology of learning and motivation, vol.* **16** (pp. 191–238). New York: Academic Press.

Healy, A. F., Cunningham, T. F., Gesi, A. T., Till, R. E., & Bourne, L. E. (1991). Comparing short-term recall of item, temporal and spatial information in children and adults. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: In honor of Bennet B. Murdock* (pp. 175–194). Hillside, NJ: Erlbaum.

Healy, A. F., Fendrich, D. W., Cunningham, T. F., & Till, R. E. (1987). Effects of cuing on short-term retention of order information. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **13,** 413–425.

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

Hendrikx, A. J. P. (1984). Effects of grouping on recall latency. *Acta Psychologica,* **55,** 1–18.

Henson, R. N. A. (in press). Item repetition in short-term memory: Ranschburg repeated. *Journal of Experimental Psychology: Learning, Memory and Cognition.*

Henson, R. N. A. (1996). *Short-term memory for serial order.* Unpublished doctoral dissertation, University of Cambridge, England.

Henson, R. N. A. (1997). *Positional information in short-term memory: Relative or absolute*? Manuscript submitted for publication.

Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology,* **49A,** 80–115.

Hintzman, D. L. (1968). Articulatory coding in short-term memory. *Psychonomic Science,* **3,** 161–162.

Hintzman, D. L. (1986). ‘‘Schema abstraction’’ in a multiple-trace memory model. *Psychological Review,* **93,** 411–428.

Hintzman, D. L., Block, R. A., & Summers, J. L. (1973). Contextual associations and memory for serial position. *Journal of Experimental Psychology,* **97,** 220–229.

Hitch, G. J. (1974). Short-term memory for spatial and temporal information. *Quarterly Journal of Experimental Psychology,* **26,** 503–513.

Hitch, G. J., Burgess, N., Towse, J. N. & Culpin, V. (1996). Temporal grouping effects in immediate recall: a working memory analysis. *Quarterly Journal of Experimental Psychology,* **49A,** 116–140.

Hitch, G. J., & Morton, J. (1975). The unimportance of explicit spatial information in serial recall of visually presented lists. *The Quarterly Journal of Experimental Psychology,* **27,** 161–164.

Houghton, G. (1990). The problem of serial order: A neural network model of sequence learning and recall. In R. Dale, C. Mellish & M. Zock, (Eds.), *Current research in natural language generation* (pp. 287–319). London: Academic Press.

Houghton, G., & Tipper, S. P. (1996). Inhibitory mechanisms of neural and cognitive control: applications to selective attention and sequential action. *Brain & Cognition,* **30,** 20–43.

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language,* **30,** 685–701.

Hulme, C., Roodenrys, S., Brown, G., & Mercer, R. (1995). The role of long-term memory mechanisms in memory span. *British Journal of Psychology,* **86,** 527–536.

Jahnke, J. C. (1969). The Ranschburg effect. *Psychological Review,* **76,** 592–605.

Jahnke, J. C., Davis, S. T., & Bower, R. E. (1989). Position and order information in recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **15,** 859–867.

Johnson, G. J. (1991). A distinctiveness model of serial learning. *Psychological Review,* **98,** 204–217.

Johnson, N. (1972). Organisation and the concept of a memory code. In A. W. Melton, & E. Martin, (Eds.), *Coding process in human memory* (pp. 125–159). Washington, DC: Winston & Sons.

Jones, D. M., Farrand, P., Stuart, G., & Morris, N. (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **21,** 1008–1018.

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **19,** 369–381.

Jones, D. M., & Macken, W. J. (1995). Organisational factors in the effect of irrelevant speech: the role of spatial location and timing. *Memory & Cognition,* **23,** 192–200.

Jordan, M. I. (1986). *Serial order: a parallel distributed approach* (ICS Report 8604). San Diego: University of California, Institute for Cognitive Science.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

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

Kidd, G. R., & Greenwald, A. G. (1988). Attention, rehearsal, and memory for serial order. *American Journal of Psychology,* **101,** 259–279.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior. The Hixon symposium* (pp. 112–136). New York: Wiley.

LeCompte, D. C., & Watkins, M. J. (1995). Grouping in primary memory: the case of the compound suffix. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **21,** 96–102.

Lee, C. L., & Estes, W. K. (1977). Order and position in primary memory for letter strings. *Journal of Verbal Learning and Verbal Memory,* **16,** 395–418.

Lee, C. L., & Estes, W. K. (1981). Item and order information in short-term memory: evidence for multilevel perturbation processes. *Journal of Experimental Psychology: Human Learning and Memory,* **7,** 149–169.

Levelt, W. J. (1989). *Speaking: From intention to articulation.* London: MIT Press.

Lewandowsky, S., & Li, S. (1994). Memory for serial order revisited. *Psychological Review,* **101,** 539–543.

Lewandowsky, S., & Murdock, B. B., Jr. (1989). Memory for serial order. *Psychological Review,* **96,** 25–57.

Li, S., & Lewandowsky, S. (1995). Forward and backward recall: different retrieval processes. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **21,** 837–847.

Macken, W. J., & Jones, D. M. (1995). Functional characteristics of the inner voice and the inner ear: single or double agency? *Journal of Experimental Psychology: Learning Memory and Cognition,* **21,** 436–448.

Madigan, S. A. (1971). Modality and recall order interactions in short-term memory for serial order. *Journal of Experimental Psychology,* **87,** 294–296.

Madigan, S. A. (1980). The serial position curve in immediate serial recall. *Bulletin of the Psychonomic Society,* **15,** 335–338.

Mandler, G., & Anderson, S. E. (1971). Temporal and spatial cues in seriation. *Journal of Experimental Psychology,* **90,** 128–135.

Margrain, S. A. (1967). Short-term memory as a function of input modality. *The Quarterly Journal of Experimental Psychology,* **19,** 109–114.

Marks, A. R. & Crowder, R. G. (1997). Temporal distinctiveness and modality. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **23,** 164–180.

Martin, E. (1974). Serial learning: a multilevel access analysis. *Memory & Cognition,* **2,** 322–328.

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

McNicol, D. (1975). Fixed and random address models for storing order in short-term memory. *The Quarterly Journal of Experimental Psychology,* **27,** 273–288.

McNicol, D. (1978). Initial codes for order in memory. *The Quarterly Journal of Experimental Psychology,* **30,** 681–691.

McNicol, D., & Heathcote, A. (1986). Representation of order information: an analysis of grouping effects in short-term memory. *Journal of Experimental Psychology: General,* **115,** 76–95.

Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior,* **2,** 1–21.

Mewhort, D. J. K., Popham, D. & James, G. (1994). On serial recall: A critique of chaining in the Theory of Distributed Associative Memory. *Psychological Review,* **101,** 534–538.

Miller, G. A. (1956/1994). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review,* **101,** 343–352.

Monsell, S. (1978). Recency, immediate recognition memory and reaction time. *Cognitive Psychology,* **10,** 465–501.

Morton, J. (1976). Two mechanisms in the stimulus suffix effect. *Memory & Cognition,* **4,** 144–149.

Murdock, B. B., Jr. (1960). The distinctiveness of stimuli. *Psychological Review,* **67,** 16–31.

Murdock, B. B. Jr. (1968). Serial order effects in short-term memory. *Journal of Experimental Psychology Monograph Supplement,* **76,** 1–15.

Murdock, B. B. Jr. (1969). Where or when: Modality effects as a function of temporal or spatial distribution of information. *Journal of Verbal Learning and Verbal Behavior,* **8,** 378–383.

Murdock, B. B., Jr. (1974). *Human memory: theory and data.* Potomac, MD: Erlbaum.

Murdock, B. B., Jr. (1983). A distributed model for serial-order information. *Psychological Review,* **90,** 316–338.

Murdock, B. B., Jr. (1993). TODAM2: A model for the storage and retrieval of item, associative, and serial-order information. *Psychological Review,* **100,** 183–203.

Murdock, B. B., Jr. (1995). Developing TODAM: Three models for serial order information. *Memory & Cognition,* **23,** 631–645.

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

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition,* **18,** 251–269.

Nairne, J. S. (1991). Positional uncertainty in long-term memory. *Memory & Cognition,* **19,** 332–340.

Nairne, J. S. & Dutta, A. (1992). Spatial and temporal uncertainty in long-term memory. *Journal of Memory & Language,* **31,** 396–407.

Nairne, J. S., & Neath, I. (1994). Critique of the retrieval/deblurring assumptions of the theory of distributed associative memory. *Psychological Review,* **101,** 528–533.

Nairne, J. S., & Whiteman, H. L., & Woessner, L. (1995). Symmetrical cuing effetcs for item and position information. *American Journal of Psychology,* **108,** 345–358.

Neath, I. (1993). Contextual and distinctive processes in the serial position function. *Journal of Memory and Language,* **32,** 820–840.

Neath, I., & Crowder, R. G. (1990). Schedules of presentation and temporal distinctiveness in human memory. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **16,** 316–327.

Neath, I., & Crowder, R. G. (1996). Distinctiveness and very short-term serial position effects. *Memory,* **4,** 225–242.

Neath, I., Surprenant, A. M. & Crowder, R. G. (1993). The context-dependent stimulus suffix effect. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **19,** 698–703.

Nelson, T. O., & Chaiklin, S. (1980). Immediate memory for spatial location. *Journal of Experimental Psychology: Human Learning and Memory,* **6,** 529–545.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorisation relationship. *Journal of Experimental Psychology: General,* **115,** 39–57.

Page, M. P. A. (1994). Modelling the perception of musical sequences with self-organising neural networks. *Connection Science,* **6,** 223–246.

Page, M. P. A., & Norris, D. G. (in press). The primacy model: a new model of immediate serial recall. *Psychological Review.*

Penney, C. G. (1978). Suffix effect in lists of temporally grouped words. *The Quarterly Journal of Experimental Psychology,* **30,** 235–250.

Penney, C. G. (1985). Elimination of the suffix effect on preterminal items with unpredictable list length: evidence for a dual model of suffix effects. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **11,** 229–247.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory and Cognition,* **17,** 398–422.

Peterson, L. R., & Johnson, S. T. (1971). Some effects of minimizing articulation on short-term retention. *Journal of Verbal Learning and Verbal Behavior,* **10,** 346–354.

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

Pollack, I., Johnson, L. B., & Knaff, P. R. (1959). Running memory span. *Journal of Experimental Psychology,* **57,** 137–146.

Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review,* **88,** 552–572.

Reitman, J. S. (1974). Without surreptitious rehearsal, information in short-term memory decays. *Journal of Verbal Learning and Verbal Behavior,* **13,** 365–377.

Ryan, J. (1969). Grouping and short-term memory: different means and patterns of grouping. *The Quarterly Journal of Experimental Psychology,* **21,** 137–147.

Salame, P., & Baddeley, A. (1982). Disruption of short-term memory by unattended speech: implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior,* **21,** 150–164.

Sanders, A. F. (1975). Some remarks on short-term memory. In P. M. A. Rabbit & S. Dornic (Eds.), *Attention and Performance V* (pp. 241–267). Hillside, NJ: Erlbaum.

Sanders, A. F., & Willemsen, E. M. (1978). Proactive interference in immediate serial recall. *Acta Psychologica,* **42,** 29–38.

Schweikert, R., & Boruff, B. (1986). Sort-term memory capacity: magic number seven or magic spell? *Journal of Experimental Psychology: Learning, Memory and Cognition,* **16,** 419–425.

Seamon, J. G., & Chumbley, J. I. (1977). Retrieval process for serial order information. *Memory & Cognition,* **5,** 709–715.

Shiffrin, R., & Cook, J. (1978). Short-term forgetting of item and order information. *Journal of Verbal Learning and Verbal Behavior,* **17,** 189–218.

Slamecka, N. (1985). Ebbinghaus: some associations. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **11,** 414–435.

Smyth, M. M., & Scholey, K. A. (1996). Serial order in spatial immediate memory. *The Quarterly Journal of Experimental Psychology,* **49A,** 159–177.

Sternberg, S. (1967). Retrieval of contextual information from memory. *Psychonomic Science,* **8,** 55–56.

Sternberg, S. (1969). Memory scanning: mental processes revealed by reaction time experiments. *American Scientist,* **57,** 421–457.

Surprenant, A. M., Pitt, M. A., & Crowder, R. G. (1993). Auditory recency in immediate memory. *The Quarterly Journal of Experimental Psychology,* **46A,** 193–223.

Tehan, G., & Humphreys, M. S. (1995). Transient phonemic codes and immunity to proactive interference. *Memory & Cognition,* **23,** 181–191.

Treiman, R., & Danis, C. (1988). Short-term memory errors for spoken syllables are affected by linguistic structure of the syllables. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **14,** 145–152.

Watkins, M. J. (1977). The intricacy of memory span. *Memory & Cognition,* **5,** 529–534.

Watkins, M. J., Watkins, O. C., & Crowder, R. (1974). The modality effect in free and serial recall as a function of phonological similarity. *Journal of Verbal Learning and Verbal Behavior,* **13,** 430–447.

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

Wickelgren, W. A. (1966). Associative intrusions in short-term recall. *Journal of Experimental Psychology,* **72,** 853–858.

Wickelgren, W. A. (1967). Rehearsal grouping and hierarchical organization of serial position cues in short-term memory. *Quarterly Journal of Experimental Psychology,* **19,** 97–102.

Wickelgren, W. A. (1969). Context-sensitive coding, associative memory, and serial order in (speech) behavior. *Psychological Review,* **76,** 1–15.

Young, R. (1968). Serial learning. In T. R. Dixon & D. L. Horton (Eds.), *Verbal Behavior and General Behavior Theory* (pp. 122–148). Englewood Cliffs, NJ: Prentice-Hall.

Accepted April 2, 1998